



# A Survey on the Eigenvalues Local Behavior of Large Complex Correlated Wishart Matrices

Walid Hachem, Adrien Hardy, Jamal Najim

## ► To cite this version:

Walid Hachem, Adrien Hardy, Jamal Najim. A Survey on the Eigenvalues Local Behavior of Large Complex Correlated Wishart Matrices. ESAIM: Proceedings and Surveys, 2015, 51, pp.150-174. hal-01199904

**HAL Id: hal-01199904**

**<https://hal.science/hal-01199904>**

Submitted on 16 Sep 2015

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

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

# A Survey on the Eigenvalues Local Behavior of Large Complex Correlated Wishart Matrices

Walid Hachem <sup>\*</sup>, Adrien Hardy <sup>†</sup>, Jamal Najim <sup>‡</sup>

September 16, 2015

## Abstract

The aim of this note is to provide a pedagogical survey of the recent works [30, 31] concerning the local behavior of the eigenvalues of large complex correlated Wishart matrices at the edges and cusp points of the spectrum: Under quite general conditions, the eigenvalues fluctuations at a soft edge of the limiting spectrum, at the hard edge when it is present, or at a cusp point, are respectively described by mean of the Airy kernel, the Bessel kernel, or the Pearcey kernel. Moreover, the eigenvalues fluctuations at several soft edges are asymptotically independent. In particular, the asymptotic fluctuations of the matrix condition number can be described. Finally, the next order term of the hard edge asymptotics is provided.

## 1 The matrix model and assumptions

Consider the  $N \times N$  random matrix defined as

$$\mathbf{M}_N = \frac{1}{N} \mathbf{X}_N \boldsymbol{\Sigma}_N \mathbf{X}_N^* \quad (1.1)$$

where  $\mathbf{X}_N$  is an  $N \times n$  matrix with independent and identically distributed (i.i.d.) entries with zero mean and unit variance, and  $\boldsymbol{\Sigma}_N$  is a  $n \times n$  deterministic positive definite Hermitian matrix. The random matrix  $\mathbf{M}_N$  has  $N$  non-negative eigenvalues, but which may be of different nature. Indeed, the smallest  $N - \min(n, N)$  eigenvalues are deterministic and all equal to zero, whereas the other  $\min(n, N)$  eigenvalues are random. The problem is then to describe the asymptotic behavior of the random eigenvalues of  $\mathbf{M}_N$ , as both dimensions of  $\mathbf{X}_N$  grow to infinity at the same rate. Let us mention that the  $n \times n$  random covariance matrix

$$\widetilde{\mathbf{M}}_N = \frac{1}{N} \boldsymbol{\Sigma}_N^{1/2} \mathbf{X}_N^* \mathbf{X}_N \boldsymbol{\Sigma}_N^{1/2},$$

which is also under consideration, has exactly the same random eigenvalues as  $\mathbf{M}_N$ , and hence results on the random eigenvalues can be carried out from one model to the other immediately.

---

<sup>\*</sup>CNRS LTCI; Télécom ParisTech, 46 rue Barrault, 75046 Paris Cedex 13, France. Email: walid.hachem@telecom-paristech.fr

<sup>†</sup>Department of Mathematics, KTH Royal Institute of Technology, Lindstedtsvägen 25, 10044 Stockholm, Sweden. Email: ahardy@kth.se

<sup>‡</sup>CNRS LIGM; Université Paris-Est, Cité Descartes, 5 Boulevard Descartes, Champs sur Marne, 77 454 Marne la Vallée Cedex 2, France. Email: najim@univ-mlv.fr

The global behavior of the spectral distribution of  $\widetilde{\mathbf{M}}_N$  in the large dimensional regime is known since the work of Marčenko and Pastur [42], where it is shown that this spectral distribution converges to a deterministic probability measure  $\mu$  that can be identified. In this paper, we will be interested in the local behavior of the eigenvalues of  $\widetilde{\mathbf{M}}_N$  near the edge points and near the so-called cusp points of the support of  $\mu$ . The former will be called the extremal eigenvalues of  $\widetilde{\mathbf{M}}_N$ .

The random matrices  $\mathbf{M}_N$  and  $\widetilde{\mathbf{M}}_N$  are ubiquitous in multivariate statistics [4], mathematical finance [39, 46], electrical engineering and signal processing [18], etc. Indeed, in multivariate statistics, the performance study of the Principal Component Analysis algorithms [36] requires the knowledge of the fluctuations of the extremal eigenvalues of  $\widetilde{\mathbf{M}}_N$ . In mathematical finance,  $\widetilde{\mathbf{M}}_N$  represents the empirical covariance matrix obtained from a sequence of asset returns. In signal processing,  $\widetilde{\mathbf{M}}_N$  often stands for the empirical covariance matrix of a spatially correlated signal received by an array of antennas, and source detection [12, 38] or subspace separation [45] algorithms also rely on the statistical study of these extremal eigenvalues.

In this article, except when stated otherwise, we restrict ourselves to the case of complex Wishart matrices. Namely, we make the following assumption.

**Assumption 1.** The entries of  $\mathbf{X}_N$  are i.i.d. standard complex Gaussian random variables.

Concerning the asymptotic regime of interest, we consider here the large random matrix regime, where the number of rows and columns of  $\mathbf{M}_N$  both grow to infinity at the same pace. More precisely, we assume  $n = n(N)$  and  $n, N \rightarrow \infty$  in such a way that

$$\lim_{N \rightarrow \infty} \frac{n}{N} = \gamma \in (0, \infty) . \quad (1.2)$$

This regime will be simply referred to as  $N \rightarrow \infty$  in the sequel.

Turning to  $\Sigma_N$ , let us denote by  $0 < \lambda_1 \leq \dots \leq \lambda_n$  the eigenvalues of this matrix and let

$$\nu_N = \frac{1}{n} \sum_{j=1}^n \delta_{\lambda_j} \quad (1.3)$$

be its spectral measure. Then we make the following assumption.

**Assumption 2.**

1. The measure  $\nu_N$  weakly converges towards a limiting probability measure  $\nu$  as  $N \rightarrow \infty$ , namely

$$\frac{1}{n} \sum_{j=1}^n f(\lambda_j) \xrightarrow{N \rightarrow \infty} \int f(x) \nu(dx)$$

for every bounded and continuous function  $f$ .

2. For  $N$  large enough, the eigenvalues of  $\Sigma_N$  stay in a compact subset of  $(0, +\infty)$  independent of  $N$ , i.e.

$$0 < \liminf_{N \rightarrow \infty} \lambda_1, \quad \sup_N \lambda_n < +\infty. \quad (1.4)$$

Under these assumptions, a comprehensive description of the large  $N$  behavior of the eigenvalues of  $\mathbf{M}_N$  can be made. To start with, we recall in Section 2 some classical results describing the global asymptotic behavior of these eigenvalues, as a necessary step for studying their local behavior. We review the results of Marčenko-Pastur [42] and those of Silverstein-Choi [50], which show among other things that the spectral measure of  $\mathbf{M}_N$  converges to a limit probability measure  $\mu$ , that  $\mu$  has a density away from zero, that the support of  $\mu$  can be delineated, and that the behavior of the density of  $\mu$  near the positive endpoints (soft edges) of this support can be characterized. We moreover complete the picture by describing the behavior of the limiting density near the origin when it is positive there (hard edge), and also when it vanishes in the interior of the support (cusp point). The latter results are extracted from [31].

Next, in Section 3 we turn to the eigenvalues local behavior. More precisely, we investigate the behavior of the random eigenvalues after zooming around several points of interest in the support, namely the soft edges, the hard edge when existing, and the cusp points. In a word, it is shown in the works [30, 31] that the Airy kernel, the Bessel kernel, and the Pearcey kernel describe the local statistics around the soft edges, the hard edge, and the cusp points respectively, provided that a regularity condition holds true. In particular, the extremal eigenvalues fluctuate according to Tracy-Widom laws.

In Section 4, we provide sketches of proofs. We first recall an important expression for the kernel  $K_N$  associated to the (random) eigenvalues of  $\mathbf{M}_N$  and then outline how one can prove asymptotic convergence towards Airy, Pearcey or Bessel kernels by zooming around the points of interest: either a soft edge, a cusp point or the hard edge.

In Section 5, we provide a list of open questions, directly related to the results of the paper.

**Acknowledgements.** WH is pleased to thank the organizers of the *Journées MAS 2014* where the project of this note was initiated. During this work, AH was supported by the grant KAW 2010.0063 from the Knut and Alice Wallenberg Foundation. The work of WH and JN was partially supported by the program “modèles numériques” of the French Agence Nationale de la Recherche under the grant ANR-12-MONU-0003 (project DIONISOS).

## 2 Global behavior

Since the seminal work of Marčenko and Pastur [42], it is known that under Assumptions 1 and 2 the spectral measure of  $\mathbf{M}_N$  almost surely (a.s.) converges weakly towards a limiting probability measure  $\mu$  with a compact support. Namely we have

$$\frac{1}{N} \text{Tr} f(\mathbf{M}_N) \xrightarrow[N \rightarrow \infty]{a.s.} \int f(x) \mu(dx) \quad (2.1)$$

for every bounded and continuous function  $f$ . As a probability measure,  $\mu$  can be characterized by its Cauchy transform: this is the holomorphic function defined by

$$m(z) = \int \frac{1}{z - \lambda} \mu(d\lambda), \quad z \in \mathbb{C}_+ = \{z \in \mathbb{C} : \text{Im}(z) > 0\},$$

and which takes its values in  $\mathbb{C}_- = \{z \in \mathbb{C} : \text{Im}(z) < 0\}$ . More precisely, for any open interval  $I \subset \mathbb{R}$  with neither endpoints on an atom of  $\mu$ , we have the inversion formula

$$\mu(I) = -\frac{1}{\pi} \lim_{\varepsilon \rightarrow 0} \int_I \text{Im}(m(x + i\varepsilon)) dx.$$

For every  $z \in \mathbb{C}_+$ , the Cauchy transform  $m(z)$  of  $\mu$  happens to be the unique solution  $m \in \mathbb{C}_-$  of the fixed-point equation

$$m = \left( z - \gamma \int \frac{\lambda}{1 - m\lambda} \nu(d\lambda) \right)^{-1}, \quad (2.2)$$

where  $\gamma$  and  $\nu$  were introduced in (1.2) and Assumption 2.

Moreover, using the free probability terminology [2, 32], the limiting distribution  $\mu$  is also known to be the free multiplicative convolution of the Marčenko-Pastur law (2.3)–(2.4) with  $\nu$ , and equation (2.2) is a consequence from the subordination property of the multiplicative free convolution [19].

For example, in the case where  $\nu = \delta_1$ , which happens e.g. when  $\Sigma_N = I_n$ , this equation has an explicit solution and the measure  $\mu$  can be recovered explicitly.

$$\mu(dx) = (1 - \gamma)^+ \delta_0 + \rho(x)dx, \quad (2.3)$$

where  $x^+ = \max(x, 0)$  and the density  $\rho$  has the expression

$$\rho(x) = \frac{1}{2\pi x} \sqrt{(\mathfrak{b} - x)(x - \mathfrak{a})} \mathbf{1}_{[\mathfrak{a}, \mathfrak{b}]}(x), \quad \mathfrak{a} = (1 - \sqrt{\gamma})^2, \quad \mathfrak{b} = (1 + \sqrt{\gamma})^2. \quad (2.4)$$

This is the celebrated Marčenko-Pastur law.

When  $\nu$  has a more complicated form, it is in general impossible to obtain an explicit expression for  $\mu$ , except in a few particular cases. Nonetheless, it is possible to make a detailed analysis of the properties of this measure, and this analysis was done by Silverstein and Choi in [50]. These authors started by showing that  $\lim_{z \in \mathbb{C}_+ \rightarrow x} m(z) \equiv m(x)$  exists for every  $x \in \mathbb{R}^* = \mathbb{R} - \{0\}$ . Consequently, the function  $m(z)$  can be continuously extended to  $\mathbb{C}_+ \cup \mathbb{R}^*$ , and furthermore,  $\mu$  has a density on  $\mathbb{R}^*$  defined as  $\rho(x) = -\pi^{-1} \text{Im}(m(x))$ . We still have the representation

$$\mu(dx) = (1 - \gamma)^+ \delta_0 + \rho(x)dx \quad (2.5)$$

with this new  $\rho$ , making  $\rho(x)dx$  the limiting distribution of the random eigenvalues of  $\mathbf{M}_N$ . As is common in random matrix theory, we shall refer to the support of  $\rho(x)dx$  as the **bulk**; we will denote (with a slight abuse of notation) its support by  $\text{Supp}(\rho)$ . Silverstein and Choi also showed that  $\rho$  is real analytic wherever it is positive, and they moreover characterized the compact support  $\text{Supp}(\mu)$  following the ideas of [42]. More specifically, one can see that the function  $m(z)$  has an explicit inverse (for the composition law) on  $m(\mathbb{C}_+)$  defined by

$$g(m) = \frac{1}{m} + \gamma \int \frac{\lambda}{1 - m\lambda} \nu(d\lambda), \quad (2.6)$$

and that this inverse extends to  $\mathbb{C}_- \cup D$  and is real analytic on  $D$ , where  $D$  is the open subset of the real line

$$D = \{x \in \mathbb{R} : x \neq 0, x^{-1} \notin \text{Supp}(\nu)\}. \quad (2.7)$$

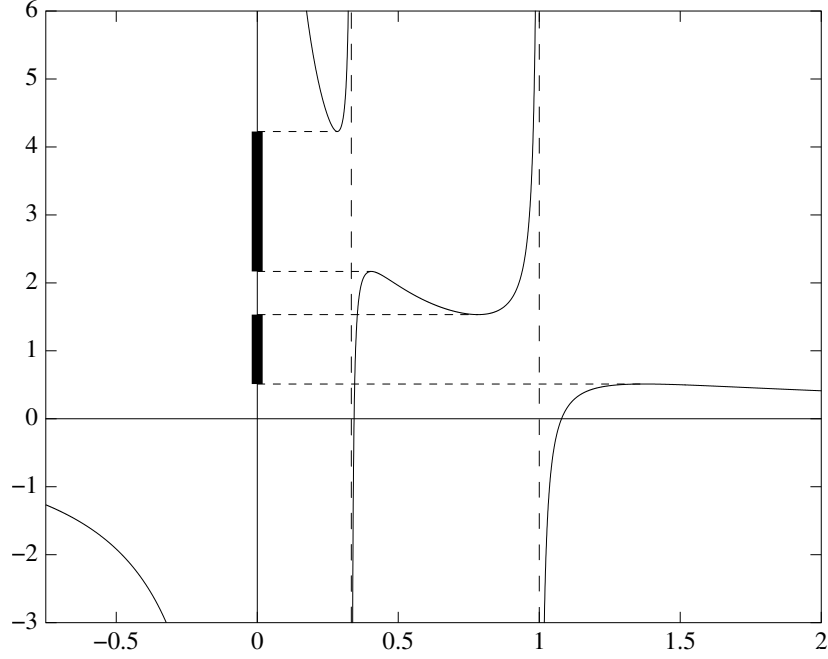


Figure 1: Plot of  $g : D \rightarrow \mathbb{R}$  for  $\gamma = 0.1$  and  $\nu = 0.7\delta_1 + 0.3\delta_3$ . In this case,  $D = (-\infty, 0) \cup (0, \frac{1}{3}) \cup (\frac{1}{3}, 1) \cup (1, \infty)$ . The two thick segments on the vertical axis represent  $\text{Supp}(\rho)$ .

It was proved in [50] that

$$\mathbb{R} - \text{Supp}(\rho) = \{g(m) : m \in D, g'(m) < 0\}.$$

An illustration of these results is provided by Figures 1 and 2.

Of interest in this paper are the left edges, the right edges and the cusp points of  $\text{Supp}(\rho)$ . A **left edge** is a real number  $\mathfrak{a}$  satisfying for every  $\delta > 0$  small enough

$$\int_{\mathfrak{a}-\delta}^{\mathfrak{a}} \rho(x) dx = 0, \quad \int_{\mathfrak{a}}^{\mathfrak{a}+\delta} \rho(x) dx > 0.$$

A **right edge** is a real number  $\mathfrak{a}$  satisfying for every  $\delta > 0$  small enough

$$\int_{\mathfrak{a}-\delta}^{\mathfrak{a}} \rho(x) dx > 0, \quad \int_{\mathfrak{a}}^{\mathfrak{a}+\delta} \rho(x) dx = 0.$$

A **cusp point** is a real number  $\mathfrak{a}$  such that  $\rho(\mathfrak{a}) = 0$  and, for every  $\delta > 0$  small enough,

$$\int_{\mathfrak{a}-\delta}^{\mathfrak{a}} \rho(x) dx > 0 \quad \text{and} \quad \int_{\mathfrak{a}}^{\mathfrak{a}+\delta} \rho(x) dx > 0.$$

Of course all edges and cusp points are positive numbers, except perhaps the leftmost edge. When the leftmost edge is the origin, it is common in random matrix theory to refer to it as the **hard edge**. In contrast, any positive edge is also called a **soft edge**.

The results of [50] summarized above show that the study of the map  $g$  on the closure  $\overline{D}$  of  $D$  provides a complete description for the edges and the cusp points. First, a right

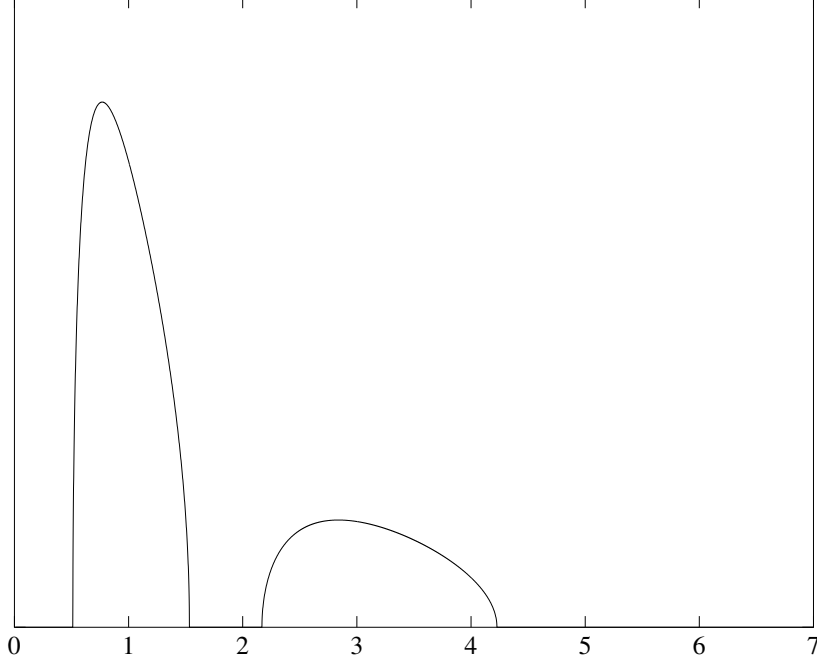


Figure 2: Plot of the density  $\rho$  in the framework of Figure 1.

edge is either a local minimum of  $g$  reached in  $D$ , or belongs to  $g(\partial D)$ , which means there is  $\mathfrak{c} \in \partial D = \overline{D} \setminus D$  such that  $\lim_{x \rightarrow \mathfrak{c}, x \in D} g(x)$  exists, is finite, and equals to that edge. In the former case,  $\rho(x)$  behaves like a square root near the edge.

**Proposition 2.1.** *If  $\mathfrak{a}$  is a right edge, then either there is a unique  $\mathfrak{c} \in D$  such that*

$$g(\mathfrak{c}) = \mathfrak{a}, \quad g'(\mathfrak{c}) = 0, \quad g''(\mathfrak{c}) > 0, \quad (2.8)$$

*or  $\mathfrak{a} \in g(\partial D)$ . In the former case, we have*

$$\rho(x) = \frac{1}{\pi} \left( \frac{2}{g''(\mathfrak{c})} \right)^{1/2} (\mathfrak{a} - x)^{1/2} (1 + o(1)), \quad x \rightarrow \mathfrak{a}_-. \quad (2.9)$$

*Conversely, if  $\mathfrak{c} \in D$  satisfies (2.8), then  $\mathfrak{a}$  is a right edge and (2.9) holds true.*

The case where an edge lies in  $g(\partial D)$  turns out to be quite delicate. In the forthcoming description of the eigenvalues local behavior near the edges, we shall restrict ourselves to the edges arising as local minima of  $g$ , see also Section 5 for further discussion. Notice also that if  $\nu$  is a discrete measure, as exemplified by Figures 1 and 2, then  $g$  is infinite on  $\partial D$  and in particular a right edge cannot belong to  $g(\partial D)$ : the right edges are in this case in a one-to-one correspondence with the local minima of  $g$  on  $D$ .

The situation is similar for the soft left edges, except that they correspond to local maxima.

**Proposition 2.2.** *If  $\mathfrak{a} > 0$  is a left edge, then either there is a unique  $\mathfrak{c} \in D$  such that*

$$g(\mathfrak{c}) = \mathfrak{a}, \quad g'(\mathfrak{c}) = 0, \quad g''(\mathfrak{c}) < 0, \quad (2.10)$$

or  $\mathfrak{a} \in g(\partial D)$ . In the former case, we have

$$\rho(x) = \frac{1}{\pi} \left( \frac{2}{-g''(\mathfrak{c})} \right)^{1/2} (x - \mathfrak{a})^{1/2} (1 + o(1)) , \quad x \rightarrow \mathfrak{a}_+ . \quad (2.11)$$

Conversely, if  $\mathfrak{c} \in D$  satisfies (2.10), then  $\mathfrak{a}$  is a right edge and (2.11) holds true.

Propositions 2.1 and 2.2 have been established in [50]. We state below their counterparts for the hard edge and a cusp point.

The hard edge setting turns out to be similar to the soft left edge one, except that  $\mathfrak{c}$  is now located at infinity, and  $\rho(x)$  behaves like an inverse square root near the hard edge. More precisely, observe that the map  $g$  is holomorphic at  $\infty$  and  $g(\infty) = 0$ , in the sense that the map  $z \mapsto g(1/z)$  is holomorphic at zero and vanishes at  $z = 0$ . We also denote by  $g'(\infty)$  and  $g''(\infty)$  the first and second derivatives of the latter map evaluated at  $z = 0$ .

**Proposition 2.3.** *The bulk presents a hard edge if and only if*

$$g(\infty) = 0, \quad g'(\infty) = 0, \quad g''(\infty) < 0, \quad (2.12)$$

or equivalently if  $\gamma = 1$ . In this case, we have

$$\rho(x) = \frac{1}{\pi} \left( \frac{2}{-g''(\infty)} \right)^{-1/2} x^{-1/2} (1 + o(1)) , \quad x \rightarrow 0_+ . \quad (2.13)$$

More precisely, we have the explicit formulas  $g'(\infty) = 1 - \gamma$  and  $g''(\infty) = -2\gamma \int \lambda^{-1} \nu(d\lambda)$ . In particular the statement  $g''(\infty) < 0$  is always true, and so is  $g(\infty) = 0$  as explained above; we included them in (2.12) to stress the analogy with (2.10).

A simple illustration of Propositions 2.1 to 2.3 is provided by the Marčenko-Pastur law. From (2.4), one immediately sees that  $\rho(x) \sim (\mathfrak{b} - x)^{1/2}$  as  $x \rightarrow \mathfrak{b}_-$ , and that a similar square root behavior near  $\mathfrak{a}$  holds if and only if  $\mathfrak{a} > 0$ , that is  $\gamma \neq 1$ . If  $\gamma = 1$ , i.e.,  $\mathfrak{a} = 0$ , then  $\rho(x) \sim x^{-1/2}$  as  $x \rightarrow 0_+$  instead.

We now turn to the cusp points. Those who will be of interest here correspond to inflexion points of  $g$  where this function is non decreasing. Moreover, a cubic root behavior for the density  $\rho(x)$  is observed near such a cusp point, hence justifying the terminology (we recall cusp usually refers to the curve defined by  $y^2 = x^3$ ).

**Proposition 2.4.** *Let  $\mathfrak{a}$  be a cusp point, set  $\mathfrak{c} = m(\mathfrak{a})$  and assume  $\mathfrak{c} \in D$ . Then*

$$g(\mathfrak{c}) = \mathfrak{a}, \quad g'(\mathfrak{c}) = 0, \quad g''(\mathfrak{c}) = 0, \quad \text{and} \quad g'''(\mathfrak{c}) > 0. \quad (2.14)$$

Moreover,

$$\rho(x) = \frac{\sqrt{3}}{2\pi} \left( \frac{6}{g'''(\mathfrak{c})} \right)^{1/3} |x - \mathfrak{a}|^{1/3} (1 + o(1)) , \quad x \rightarrow \mathfrak{a} . \quad (2.15)$$

Conversely, if  $\mathfrak{c} \in D$  satisfies  $g'(\mathfrak{c}) = g''(\mathfrak{c}) = 0$ , then the real number  $\mathfrak{a} = g(\mathfrak{c})$  is a cusp point,  $g'''(\mathfrak{c}) > 0$  and (2.15) holds true.

Propositions 2.3 and 2.4 appear in [31]. Proposition 2.4 is illustrated in Figures 3 and 4.



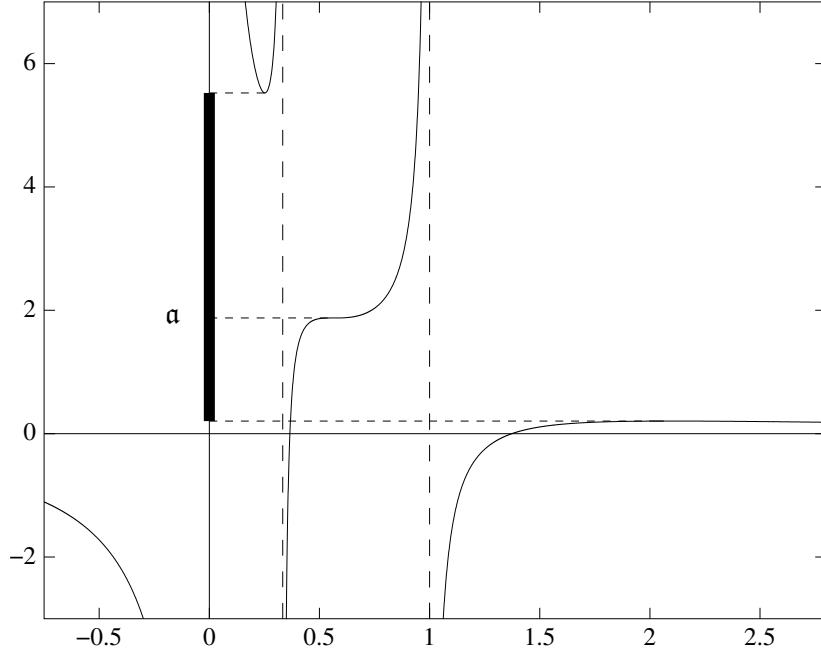


Figure 3: Plot of  $g$  for  $\gamma \simeq 0.336$  and  $\nu = 0.7\delta_1 + 0.3\delta_3$ . The thick segment on the vertical axis represents  $\text{Supp}(\mu)$ . The point  $\mathfrak{a}$  is a cusp point.

### 3 Local behavior

The study of the eigenvalues local behavior of random matrices is a central topic in random matrix theory. When dealing with large Hermitian random matrices, it is recognized that the local correlation of the eigenvalues around an edge where the density vanishes like a square root should be described by a particular point process involving the Airy kernel (see below), whose maximal particle's distribution is known as the Tracy-Widom law. For instance, this has been established for unitary invariant random matrices and for Wigner matrices as well, see e.g. the surveys [21, 27] and references therein. Similarly, the Bessel kernel is expected to describe the fluctuations around an hard edge where the density vanishes like an inverse square root, and the Pearcey kernel around a cusp with cubic root behavior. Let us also mention that the sine kernel is expected around a point where the density is positive, and that more sophisticated behaviors have also been observed in matrix models where the density vanishes like a rational power of different order, but we will not further investigate these aspects here. Another interesting feature not covered by this survey is the study of the random eigenvectors, see e.g. [10, 13].

The purpose of this section is to present the central results of [30, 31], where such typical local behaviors arise for the complex correlated Wishart matrices under consideration at every edges and cusp points satisfying a certain regularity condition.

In fact, the free parameter family  $(\nu_N)$  may have a deep impact on the limiting local fluctuations and one may not recover the expected fluctuations without further conditions. A first manifestation of this phenomenon is the Baik-Ben Arous-Péché (BBP) phase transition, that we present in Section 3.1. In a nutshell, this phase transition yields that slight variations on the family  $(\nu_N)$  may modify the fluctuations at a soft edge and may no longer be described by

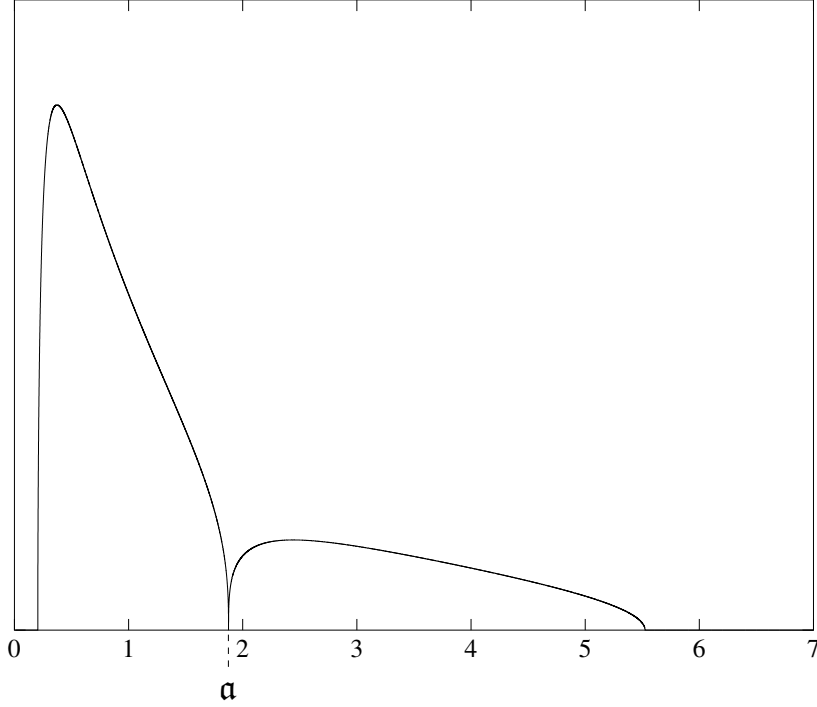


Figure 4: Plot of the density of  $\mu$  in the framework of Figure 3.

the Tracy-Widom law. Such phenomenas motivate the introduction of a regularity condition which essentially rules out this kind of behaviors.

In Section 3.2, we provide the existence of finite  $N$  approximations of the edges or cusp points under study and satisfying the regularity condition; the reader not interested in these precise definitions may skip this section.

Next, in Section 3.3 we introduce the Airy kernel, the Tracy-Widom law, and state our results concerning the soft edges. In Section 3.4, we introduce the Bessel kernel and describe the fluctuations at the hard edge. As an application, we provide in Section 3.5 a precise description for the asymptotic behavior of the condition number of  $\mathbf{M}_N$ . Finally, in Section 3.6 we introduce the Pearcey kernel, and state our result concerning the asymptotic behavior near a cusp point.

### 3.1 The BBP phase transition and the regularity assumption

First, assume  $\Sigma_N$  is the identity matrix, so that the (limiting) spectral distribution  $\nu$  of  $\Sigma_N$  is  $\delta_1$ , and hence the limiting density  $\rho(x)$  is provided by (2.4). If  $x_{\max}$  stands for the maximal eigenvalue of  $\mathbf{M}_N$ , then it has been established that  $x_{\max}$  converges a.s. towards the right edge and fluctuates at the scale  $N^{2/3}$  according to the Tracy-Widom law [34]. Next, following Baik, Ben Arous and P      [7], assume instead  $\Sigma_N$  is a finite rank additive perturbation of the identity, meaning that the rank of the perturbation is independent on  $N$ . Thus we still have  $\nu = \delta_1$  and the limiting density  $\rho(x)$  remains unchanged. They established that if the strength of the perturbation is limited, then the behavior for  $x_{\max}$  is the same as in the non-perturbed case, see [7, Theorem 1.1(a),  $k = 0$ ]. On the contrary, if the perturbation is strong enough, then  $x_{\max}$  converges a.s. outside of the bulk and the fluctuations are of different nature, see [7,



of the fixed-point equation

$$m = \left( z - \frac{n}{N} \int \frac{\lambda}{1 - m\lambda} \nu_N(d\lambda) \right)^{-1}.$$

The probability measure  $\mu_N$  should be thought of as a deterministic approximation of the distribution of the eigenvalues of  $\mathbf{M}_N$  at finite  $N$ , and is referred to as the **deterministic equivalent** of the spectral measure of  $\mathbf{M}_N$ . The measure  $\mu_N$  reads

$$\mu_N(dx) = \left(1 - \frac{n}{N}\right)^+ \delta_0 + \rho_N(x)dx,$$

and one can apply all the results stated in Section 2 to describe  $\rho_N$ , after replacement of  $g$  with

$$g_N(z) = \frac{1}{z} + \frac{n}{N} \int \frac{\lambda}{1 - z\lambda} \nu_N(d\lambda). \quad (3.2)$$

Recalling  $D$  has been introduced in (2.7), the following proposition encodes the essential consequence of the regularity condition.

**Proposition 3.4.** *If  $\mathfrak{c} \in D$  satisfies the regularity condition (3.1), then there exists  $\delta > 0$  such that  $g_N$  is holomorphic on  $\{z \in \mathbb{C} : |z - \mathfrak{c}| < \delta\} \subset D$  for every  $N$  large enough and converges uniformly towards  $g$  there.*

It is an easy consequence of Montel's theorem. Now, if a sequence of holomorphic function  $h_N$  converges uniformly to a (holomorphic) function  $h$  on an open disc, then a standard result from complex analysis provides that the  $k$ th order derivative  $h_N^{(k)}$  also converges uniformly to  $h^{(k)}$  there, for every  $k \geq 1$ . Moreover, Hurwitz's theorem states that, if  $h$  has a zero  $\mathfrak{c}$  of multiplicity  $\ell$  in this disc, then  $h_N$  has exactly  $\ell$  zeros, including multiplicity, converging towards  $\mathfrak{c}$  as  $N \rightarrow \infty$ .

Thus, as a consequence of the previous proposition, by applying Hurwitz's theorem to  $g'_N$  (and the symmetry  $g'_N(\bar{z}) = g'_N(z)$ ), it is easy to obtain the following statement.

**Proposition 3.5.** *Assume  $\mathfrak{c} \in D$  satisfies the regularity condition (3.1) and moreover*

$$g'(\mathfrak{c}) = 0, \quad g''(\mathfrak{c}) < 0, \quad \text{resp.} \quad g''(\mathfrak{c}) > 0.$$

*Then there exists a sequence  $(\mathfrak{c}_N)$ , unique up to a finite number of terms, converging to  $\mathfrak{c}$  and such that, for every  $N$  large enough, we have  $\mathfrak{c}_N \in D$  and*

$$\lim_{N \rightarrow \infty} g_N(\mathfrak{c}_N) = g(\mathfrak{c}), \quad g'_N(\mathfrak{c}_N) = 0, \quad g''_N(\mathfrak{c}_N) < 0, \quad \text{resp.} \quad g''_N(\mathfrak{c}_N) > 0.$$

Having in mind Propositions 2.2 and 2.1, this proposition thus states that if one considers a regular left (resp. right) soft edge  $\mathfrak{a}$ , and thus  $\mathfrak{a} = g(\mathfrak{c})$  with  $\mathfrak{c} \in D$  by Remark 3.2, then there exists a sequence, unique up to a finite number of terms, of left (resp. right) soft edges  $\mathfrak{a}_N = g_N(\mathfrak{c}_N)$  for the deterministic equivalent  $\mu_N$  converging towards  $\mathfrak{a}$ . These soft edges  $(\mathfrak{a}_N)$  are finite  $N$  approximations of the edge  $\mathfrak{a}$ , while the  $\mathfrak{c}_N$ 's are finite  $N$  approximations of the preimage  $\mathfrak{c}$ .

When dealing with regular cusp points, the situation is slightly more delicate. The reason for this is that if  $\mathfrak{a} = g(\mathfrak{c})$  is a regular cusp point, then  $\mathfrak{c}$  is now a zero of multiplicity two for  $g'$ . By applying Hurwitz's theorem to  $g'_N$  as above, one would obtain two sequences of non-necessarily real zeros for  $g'_N$  converging towards  $\mathfrak{c}$ . It is actually more convenient to apply Hurwitz's theorem to  $g''_N$  instead, in order to get the following statement.

**Proposition 3.6.** Assume  $\mathfrak{c} \in D$  satisfies the regularity condition (3.1) and moreover

$$g'(\mathfrak{c}) = 0, \quad g''(\mathfrak{c}) = 0, \quad (\text{hence } g'''(\mathfrak{c}) > 0 \text{ by Prop. 2.4}).$$

Then there exists a sequence  $(\mathfrak{c}_N)$ , unique up to a finite number of terms, converging to  $\mathfrak{c}$  and such that, for every  $N$  large enough, we have  $\mathfrak{c}_N \in D$  and

$$\lim_{N \rightarrow \infty} g_N(\mathfrak{c}_N) = g(\mathfrak{c}), \quad \lim_{N \rightarrow \infty} g'_N(\mathfrak{c}_N) = 0, \quad g''_N(\mathfrak{c}_N) = 0, \quad g'''_N(\mathfrak{c}_N) > 0.$$

Notice that Proposition 3.6 doesn't guarantee that  $g'_N(\mathfrak{c}_N) = 0$ . Hence, a cusp point is not necessarily the limit of cusp points of the deterministic equivalents  $\mu_N$ . As we shall see in Section 3.6, the speed at which  $g'_N(\mathfrak{c}_N)$  goes to zero will actually influence the local behavior around the cusp.

**Definition 3.7.** Given a soft left edge, resp. right edge, resp. cusp point  $\mathfrak{a}$  which is regular, and thus  $\mathfrak{a} = g(\mathfrak{c})$  with  $g'(\mathfrak{c}) = 0$  and  $g''(\mathfrak{c}) < 0$ , resp.  $g''(\mathfrak{c}) > 0$ , resp.  $g'(\mathfrak{c}) = 0$  and  $g'''(\mathfrak{c}) > 0$ , the **sequence associated with  $\mathfrak{a}$**  is the sequence  $(\mathfrak{c}_N)$  provided by Propositions 3.5 and 3.6.

Equipped with Propositions 3.5 and 3.6, we are now in position to state the results concerning the local asymptotics.

### 3.3 The Airy kernel and Tracy-Widom fluctuations at a soft edge

Given a function  $K(x, y)$  from  $\mathbb{R} \times \mathbb{R}$  to  $\mathbb{R}$  satisfying appropriate conditions, one can consider its associated determinantal point process, which is a simple point process on  $\mathbb{R}$  having as correlation functions the determinants  $\det[K(y_i, y_j)]$ . More precisely, it is a probability distribution  $\mathbb{P}$  over the configurations  $(y_i)$  of real numbers (the particles), namely over discrete subsets of  $\mathbb{R}$  which are locally finite, characterized in the following way: For every  $k \geq 1$  and any test function  $\Phi : \mathbb{R}^k \rightarrow \mathbb{R}$ ,

$$\mathbb{E} \left[ \sum_{y_{i_1} \neq \dots \neq y_{i_k}} \Phi(y_{i_1}, \dots, y_{i_k}) \right] = \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \Phi(y_1, \dots, y_k) \det [K(y_i, y_j)]_{i,j=1}^k dy_1 \dots dy_k,$$

where the sum runs over the  $k$ -tuples of pairwise distinct particles of the configuration  $(y_i)$ . Hence the correlation between the particles  $y_i$ 's is completely encoded by the kernel  $K(x, y)$ . In particular, the inclusion-exclusion principle yields a closed formula for the gap probabilities in terms of Fredholm determinants. Namely, for any interval  $J \subset \mathbb{R}$ , the probability that no particle lies in  $J$  reads

$$\mathbb{P}((y_i) \cap J = \emptyset) = 1 + \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \int_J \dots \int_J \det [K(y_i, y_j)]_{i,j=1}^k dy_1 \dots dy_k,$$

and the latter is the Fredholm determinant  $\det(I - K)_{L^2(J)}$  of the integral operator acting on  $L^2(J)$  with kernel  $K(x, y)$ , provided it makes sense. We refer to [33, 35] for further information on determinantal point processes.

Consider the **Airy point process**  $\mathbb{P}_{\text{Airy}}$  which is defined as the determinantal point process on  $\mathbb{R}$  associated with the Airy kernel

$$K_{\text{Airy}}(x, y) = \frac{\text{Ai}(x)\text{Ai}'(y) - \text{Ai}(y)\text{Ai}'(x)}{x - y}, \quad (3.3)$$

where the Airy function

$$\text{Ai}(x) = \frac{1}{\pi} \int_0^\infty \cos\left(\frac{u^3}{3} + ux\right) du$$

is a solution of the differential equation  $f''(x) = xf(x)$ .

The configurations  $(y_i)$  generated by the Airy point process a.s. involve an infinite number of particles but have a largest particle  $y_{\max}$ . The distribution of  $y_{\max}$  is the **Tracy-Widom law** (see e.g. [35, Section 2.2]), and its distribution function reads, for every  $s \in \mathbb{R}$ ,

$$\mathbb{P}_{\text{Airy}}(y_{\max} \leq s) = \mathbb{P}_{\text{Airy}}((y_i) \cap (s, +\infty) = \emptyset) = \det(I - K_{\text{Airy}})_{L^2(s, \infty)}. \quad (3.4)$$

Tracy and Widom [51] established the famous representation

$$\mathbb{P}_{\text{Airy}}(y_{\max} \leq s) = \exp\left(-\int_s^\infty (x-s)q(x)^2 dx\right),$$

where  $q$  is the Hastings-McLeod solution of the Painlevé II equation, namely the unique solution of  $f''(x) = f(x)^3 + xf(x)$  with boundary condition  $f(x) \sim \text{Ai}(x)$  as  $x \rightarrow +\infty$ .

Recalling that  $g_N$  has been introduced in (3.2), we are now in position to describe the eigenvalues local behavior around regular soft edges. In the three upcoming theorems, we denote by  $\tilde{x}_1 \leq \dots \leq \tilde{x}_n$  the ordered eigenvalues of  $\mathbf{M}_N$ . We also use the notational convention  $\tilde{x}_0 = 0$  and  $\tilde{x}_{n+1} = +\infty$ .

**Theorem 1.** *Let  $\mathbf{a}$  be a right edge and assume it is regular. Writing  $\mathbf{a} = g(\mathbf{c})$ , let  $\varphi(N) = \max\{j : \lambda_j^{-1} > \mathbf{c}\}$ . Then, almost surely,*

$$\tilde{x}_{\varphi(N)} \xrightarrow[N \rightarrow \infty]{} \mathbf{a}, \quad \liminf_{N \rightarrow \infty} (\tilde{x}_{\varphi(N)+1} - \mathbf{a}) > 0. \quad (3.5)$$

Moreover, let  $(\mathbf{c}_N)_N$  be the sequence associated with  $\mathbf{a}$  as in Definition 3.7. Set

$$\mathbf{a}_N = g_N(\mathbf{c}_N), \quad \sigma_N = \left(\frac{2}{g_N''(\mathbf{c}_N)}\right)^{1/3}, \quad (3.6)$$

so that  $\mathbf{a}_N \rightarrow \mathbf{a}$ ,  $\mathbf{c}_N \rightarrow \mathbf{c}$ , and  $\sigma_N \rightarrow (2/g''(\mathbf{c}))^{1/3} > 0$  as  $N \rightarrow \infty$ . Then, for every  $s \in \mathbb{R}$ ,

$$\lim_{N \rightarrow \infty} \mathbb{P}\left(N^{2/3}\sigma_N(\tilde{x}_{\varphi(N)} - \mathbf{a}_N) \leq s\right) = \mathbb{P}_{\text{Airy}}(y_{\max} \leq s). \quad (3.7)$$

**Remark 3.8.** Let us stress that the sequence  $\varphi(N)$  may be non-trivial even when considering the rightmost edge: As explained in Section 3.1, it is indeed possible that a certain amount of eigenvalues (possibly infinite) will converge outside of the limiting support. Thus, if we assume the rightmost edge is regular, then Theorem 1 states that there exists a maximal eigenvalue  $x_{\varphi(N)}$  which actually converge to the rightmost edge and fluctuates according to the Tracy-Widom law.

Let us comment on the history of this theorem. The Tracy-Widom fluctuations have been first obtained by Johansson [34] for the maximal eigenvalue when  $\Sigma_N$  is the identity. Baik, Ben Arous and P      then proved this still holds true when  $\Sigma_N$  is a finite rank perturbation of the identity, provided the perturbation is small enough. Assuming a condition which is

equivalent to the regularity condition (3.1) and that the maximal eigenvalue converges towards the rightmost edge, El Karoui [26] established the Tracy-Widom fluctuations for the maximal eigenvalue for general  $\Sigma_N$ 's assuming  $\gamma \leq 1$ , and Onatski [48] got rid of the last restriction. The statement on the existence of extremal eigenvalues converging to each regular right edge is [30, Theorem 2] and essentially relies on the exact separation results of Bai and Silverstein [5, 6]; the definition of the sequence  $\varphi(N)$  indexing these extremal eigenvalues relies on these results. Finally, the Tracy-Widom fluctuations for the extremal eigenvalues associated to any right regular edge is [30, Theorem 3-(b)].

We now provide a similar statement for the left soft edges.

**Theorem 2.** *Let  $\mathbf{a}$  be a left edge of the bulk. If  $\gamma > 1$  and  $\mathbf{a}$  is the leftmost edge of the bulk, set  $\varphi(N) = n - N + 1$ . Otherwise, assume that  $\mathbf{a} > 0$  is regular, write  $\mathbf{a} = g(\mathbf{c})$  and set  $\varphi(N) = \min\{\lambda_j : \lambda_j^{-1} < \mathbf{c}\}$ . Then, almost surely,*

$$\tilde{x}_{\varphi(N)} \xrightarrow[N \rightarrow \infty]{} \mathbf{a}, \quad \liminf_{N \rightarrow \infty} (\mathbf{a} - \tilde{x}_{\varphi(N)-1}) > 0. \quad (3.8)$$

Moreover, let  $(\mathbf{c}_N)_N$  be the sequence associated with  $\mathbf{a}$  as in Definition 3.7. Set

$$\mathbf{a}_N = g_N(\mathbf{c}_N), \quad \sigma_N = \left( \frac{2}{-g_N''(\mathbf{c}_N)} \right)^{1/3}, \quad (3.9)$$

so that  $\mathbf{a}_N \rightarrow \mathbf{a}$ ,  $\mathbf{c}_N \rightarrow \mathbf{c}$ , and  $\sigma_N \rightarrow (-2/g''(\mathbf{c}))^{1/3} > 0$  as  $N \rightarrow \infty$ . Then, for every  $s \in \mathbb{R}$ ,

$$\lim_{N \rightarrow \infty} \mathbb{P}\left(N^{2/3}\sigma_N(\mathbf{a}_N - \tilde{x}_{\varphi(N)}) \leq s\right) = \mathbb{P}_{\text{Airy}}(y_{\max} \leq s). \quad (3.10)$$

Prior to this result, which is a combination of Theorem 2 and Theorem 3-(a) from [30], the Tracy-Widom fluctuations for the smallest random eigenvalue when  $\Sigma_N$  is the identity has been obtained by Borodin and Forrester [15].

Let us also mention that when  $\nu$  is the sum of two Dirac masses, a local uniform convergence to the Airy kernel (which is a weaker statement than the Tracy-Widom fluctuations) at every (regular) right and soft left edges follows from [41], see also [43, 44].

Finally, we state our last result, concerning the asymptotic independence of the Tracy-Widom fluctuations at a finite number of regular soft edges. For a more precise statement, we refer to [30, Theorem 4].

**Theorem 3.** *Let  $(\mathbf{a}_j)_{j \in J}$  be a finite collection of soft edges, and assume all these edges are regular. For each  $j \in J$ , consider the rescaled eigenvalue  $N^{2/3}\sigma_{N,j}(\tilde{x}_{\varphi_j(N)} - \mathbf{a}_{N,j})$  associated with the soft edge  $\mathbf{a}_j$  provided by (3.5)–(3.6) and (3.8)–(3.9). Then the family of random variables  $\{N^{2/3}\sigma_{N,j}(\tilde{x}_{\varphi_j(N)} - \mathbf{a}_{N,j})\}_{j \in J}$  becomes asymptotically independent as  $N \rightarrow \infty$ .*

The asymptotic independence has been previously established for the smallest and largest eigenvalues when  $\Sigma_N$  is the identity by Basor, Chen and Zhang [9].

**Remark 3.9.** The results presented in this survey rely on the fact that the entries of  $\mathbf{X}_N$  are complex Gaussian random variables, a key assumption in order to take advantage of the determinantal structure of the eigenvalues of the model under study. A recent work [37] by Knowles and Yin enables to transfer the results of Theorems 1, 2 and 3 to the case of complex, but not necessarily Gaussian, random variables. Indeed, by combining the local convergence to the limiting distribution established in [37] together with Theorems 1, 2 and

3, one obtains Tracy-Widom fluctuations and asymptotic independence in this more general setting, provided that the entries of matrix  $\mathbf{X}_N$  fulfill some moment condition. Let us stress that the case of real Gaussian random variables (except the largest one covered in [40]), of important interest in statistical applications, remains open.

We now turn to the hard edge and the Bessel point process.

### 3.4 The Bessel point process at the hard edge

The **Bessel point process**  $\mathbb{P}_{\text{Bessel}}^{(\alpha)}$  of parameter  $\alpha \in \mathbb{Z}$  is the determinantal point process on  $\mathbb{R}_+$  associated with the kernel

$$K_{\text{Bessel}}^{(\alpha)}(x, y) = \frac{\sqrt{y} J_\alpha(\sqrt{x}) J'_\alpha(\sqrt{y}) - \sqrt{x} J'_\alpha(\sqrt{x}) J_\alpha(\sqrt{y})}{2(x - y)}, \quad (3.11)$$

where the Bessel function of the first kind  $J_\alpha$  with parameter  $\alpha$  is defined for  $x \geq 0$  by

$$J_\alpha(x) = \left(\frac{x}{2}\right)^\alpha \sum_{n=0}^{\infty} \frac{(-1)^n}{n! \Gamma(n + \alpha + 1)} \left(\frac{x}{2}\right)^{2n} \quad (3.12)$$

and satisfies the differential equation  $x^2 f''(x) + x f'(x) + (x^2 - \alpha^2) f(x) = 0$ .

The configurations  $(y_i)$  generated by the Bessel point process a.s. have an infinite number of particles  $y_i$  but have a smallest particle  $y_{\min}$ . The law of  $y_{\min}$  is characterized, for every  $s > 0$ , by

$$\mathbb{P}_{\text{Bessel}}^{(\alpha)}(y_{\min} \geq s) = \mathbb{P}_{\text{Bessel}}^{(\alpha)}((y_i) \cap (0, s) = \emptyset) = \det(I - K_{\text{Bessel}}^{(\alpha)})_{L^2(0, s)}. \quad (3.13)$$

When  $\alpha = 0$ , this reduces to an exponential law of parameter 1, namely  $\mathbb{P}_{\text{Bessel}}^{(0)}(y_{\min} \geq s) = e^{-s}$ , as observed by Edelman [24]. In the general case, Tracy and Widom obtained the representation [52],

$$\mathbb{P}_{\text{Bessel}}^{(\alpha)}(y_{\min} \geq s) = \exp\left(-\frac{1}{4} \int_0^s (\log s - \log x) q(x)^2 dx\right),$$

where  $q$  is the solution of a differential equation which is reducible to a particular case of the Painlevé V equation (involving  $\alpha$  in its parameters) and boundary condition  $q(x) \sim J_\alpha(\sqrt{x})$  as  $x \rightarrow 0_+$ .

Recalling the  $\lambda_j$ 's are the eigenvalues of  $\Sigma_N$  and their distributional limit  $\nu$  has a compact support in  $(0, \infty)$ , we now provide our statement concerning the eigenvalues local behavior around the hard edge.

**Theorem 4.** Assume that  $n = N + \alpha$  with  $\alpha \in \mathbb{Z}$  independent of  $N$  and set

$$\sigma_N = -2g_N''(\infty) = \frac{4}{N} \sum_{j=1}^n \frac{1}{\lambda_j}, \quad \zeta_N = -\frac{4}{3}g_N'''(\infty) = \frac{8}{N} \sum_{j=1}^n \frac{1}{\lambda_j^2}.$$

Thus  $\sigma_N \rightarrow 4 \int \lambda^{-1} \nu(d\lambda) > 0$  and  $\zeta_N \rightarrow 8 \int \lambda^{-2} \nu(d\lambda) > 0$  as  $N \rightarrow \infty$ .

Let  $x_{\min}$  be the smallest random eigenvalue of  $\mathbf{M}_N$ . Then, for every  $s > 0$ , we have

$$\lim_{N \rightarrow \infty} \mathbb{P}(N^2 \sigma_N x_{\min} \geq s) = \mathbb{P}_{\text{Bessel}}^{(\alpha)}(y_{\min} \geq s). \quad (3.14)$$



Furthermore, we have the expansion as  $N \rightarrow \infty$ ,

$$\mathbb{P}\left(N^2 \sigma_N x_{\min} \geq s\right) = \mathbb{P}_{\text{Bessel}}^{(\alpha)}(y_{\min} \geq s) - \frac{1}{N} \left( \frac{\alpha \zeta_N}{\sigma_N^2} \right) s \frac{d}{ds} \mathbb{P}_{\text{Bessel}}^{(\alpha)}(y_{\min} \geq s) + O\left(\frac{1}{N^2}\right). \quad (3.15)$$

When  $\Sigma_N$  is the identity, the convergence (3.14) has been established by Forrester [28]. As for the next order term (3.15), it has been obtained when  $\Sigma_N$  is the identity by Perret and Schehr [49] and Bornemann [14], motivated by a question raised by Edelman, Guionnet and P      in [25]. The statement (3.14) has been first obtained in [30], while the stronger statement (3.15) is [31, Theorem 6].

### 3.5 Application to condition numbers

In this subsection, we study the fluctuations of the ratio

$$\kappa_N = \frac{x_{\max}}{x_{\min}}$$

of the largest to the smallest random eigenvalue of  $\mathbf{M}_N$ . Notice that if  $n \geq N$ , then  $\kappa_N$  is the condition number of  $\mathbf{M}_N$  while if  $n \leq N$ , then  $\kappa_N$  is the condition number of  $\widetilde{\mathbf{M}}_N$ . The condition number is a central object of study in numerical linear algebra [47, 29]. Using our previous results, we can obtain an asymptotic description for  $\kappa_N$ . Let us emphasize that the leftmost edge  $\mathbf{a}$  of the support of  $\rho$  is positive if and only if  $\gamma \neq 1$ , see [31, Proposition 3].

**Proposition 3.10.** *Assume  $\gamma \neq 1$ . Denote by  $\mathbf{a}$  the leftmost edge and by  $\mathbf{b}$  the rightmost one. Assume that  $\mathbf{a}, \mathbf{b}$  are regular,  $x_{\min} \rightarrow \mathbf{a}$  and  $x_{\max} \rightarrow \mathbf{b}$  a.s. as  $N \rightarrow \infty$  (that  $\mathbf{a}$  is regular and  $x_{\min} \rightarrow \mathbf{a}$  a.s. is always true when  $\gamma > 1$ ). Write  $\mathbf{a} = g(\mathbf{c})$ ,  $\mathbf{b} = g(\mathbf{d})$ , consider the sequences  $(\mathbf{c}_N)$  and  $(\mathbf{d}_N)$  associated with  $\mathbf{c}$  and  $\mathbf{d}$  respectively (see Definition 3.7) and set*

$$\mathbf{a}_N = g_N(\mathbf{c}_N), \quad \sigma = \left( \frac{2}{-g''(\mathbf{c})} \right)^{1/3}, \quad \mathbf{b}_N = g_N(\mathbf{d}_N), \quad \delta = \left( \frac{2}{g''(\mathbf{d})} \right)^{1/3}.$$

Then,

$$\kappa_N \xrightarrow[N \rightarrow \infty]{a.s.} \frac{\mathbf{b}}{\mathbf{a}} \quad \text{and} \quad N^{2/3} \left( \kappa_N - \frac{\mathbf{b}_N}{\mathbf{a}_N} \right) \xrightarrow[N \rightarrow \infty]{\mathcal{D}} \frac{X}{\delta \mathbf{a}} + \frac{\mathbf{b} Y}{\sigma \mathbf{a}^2}$$

where  $\xrightarrow{\mathcal{D}}$  stands for the convergence in distribution and where  $X$  and  $Y$  are two independent random variables with the Tracy-Widom distribution.

We now handle the case where  $\gamma = 1$ .

**Proposition 3.11.** *Assume  $n = N + \alpha$ , where  $\alpha \in \mathbb{Z}$  is independent of  $N$ , and moreover  $x_{\max} \rightarrow \mathbf{b}$  a.s. for some  $\mathbf{b} > 0$ . Then,*

$$\frac{1}{N^2} \kappa_N \xrightarrow[N \rightarrow \infty]{\mathcal{D}} \frac{4\mathbf{b}}{X} \left( \int \lambda^{-1} \nu(d\lambda) \right)$$

where  $\mathbb{P}(X \geq s) = \mathbb{P}_{\text{Bessel}}^{(\alpha)}(y_{\min} \geq s)$  for every  $s > 0$ .

**Remark 3.12.** Interestingly, in the square case where  $\gamma = 1$ , the fluctuations of the largest eigenvalue  $x_{\max}$  have no influence on the fluctuations of  $\kappa_N$  as these are imposed by the limiting distribution of  $x_{\min}$  and the a.s. limit  $\mathbf{b}$  of  $x_{\max}$ .

Finally, we turn to the eigenvalues local behavior near a cusp point.

### 3.6 The Pearcey kernel at a cusp point

Given any  $\tau \in \mathbb{R}$ , following [53] we introduce the Pearcey-like integral functions

$$\varphi(x) = \frac{1}{2i\pi} \oint_{\Sigma} e^{xz - \tau z^2/2 + z^4/4} dz, \quad \psi(y) = \frac{1}{2i\pi} \int_{-i\infty}^{i\infty} e^{-yw + \tau w^2/2 - w^4/4} dw,$$

where the contour  $\Sigma$  consists in two rays going from  $\pm e^{i\pi/4}\infty$  to zero, and two rays going from zero to  $\pm e^{-i\pi/4}$ . They satisfy the respective differential equations

$$\varphi'''(x) - \tau\varphi'(x) + x\varphi(x) = 0, \quad \psi'''(y) - \tau\psi'(y) - y\psi(y) = 0.$$

The **Pearcey point process**  $\mathbb{P}_{\text{Pearcey}}^{(\tau)}$  is the determinantal point process associated with the Pearcey kernel

$$K_{\text{Pearcey}}^{(\tau)}(x, y) = \frac{\varphi''(x)\psi(y) - \varphi'(x)\psi'(y) + \varphi(x)\psi''(y) - \tau\psi(x)\psi(y)}{x - y}. \quad (3.16)$$

This process has been first introduced by Brézin and Hikami [16, 17] when  $\tau = 0$ , and subsequent generalizations have been considered by Tracy and Widom [53].

The configurations  $(y_i)$  generated by the Pearcey point process are a.s. infinite and do not have a largest nor smallest particle. With this respect, the quantities of interest here are the gap probabilities of the Pearcey point process, defined for every  $0 < s < t$  by

$$\mathbb{P}_{\text{Pearcey}}^{(\tau)}\left((y_i) \cap [s, t] = \emptyset\right) = \det(I - K_{\text{Pearcey}}^{(\tau)})_{L^2(s, t)}. \quad (3.17)$$

Seen as a function of  $s, t$  and  $\tau$ , the log of the righthand side of (3.17) is known to satisfy a system of PDEs, see [53, 11, 1].

In [31, Theorem 5], we prove the following statement.

**Theorem 5.** *Let  $\mathbf{a} = g(\mathbf{c})$  be a cusp point such that  $\mathbf{c} \in D$ , and assume it is regular. Let  $(\mathbf{c}_N)$  be the sequence associated with  $\mathbf{a}$  as in Definition 3.7. Assume moreover the following decay assumption holds true: There exists  $\kappa \in \mathbb{R}$  such that*

$$\sqrt{N} g'_N(\mathbf{c}_N) \xrightarrow{N \rightarrow \infty} \kappa. \quad (3.18)$$

We set

$$\mathbf{a}_N = g_N(\mathbf{c}_N), \quad \sigma_N = \left(\frac{6}{g_N^{(3)}(\mathbf{c}_N)}\right)^{1/4}, \quad \tau = -\kappa \left(\frac{6}{g^{(3)}(\mathbf{c})}\right)^{1/2}, \quad (3.19)$$

so that  $\mathbf{a}_N \rightarrow \mathbf{a}$  and  $\sigma_N \rightarrow (6/g'''(\mathbf{c}))^{1/4} > 0$  as  $N \rightarrow \infty$ . Then, for every  $s > 0$ , we have

$$\lim_{N \rightarrow \infty} \mathbb{P}\left((N^{3/4}\sigma_N(x_i - \mathbf{a}_N)) \cap [-s, s] = \emptyset\right) = \mathbb{P}_{\text{Pearcey}}^{(\tau)}\left((y_i) \cap [-s, s] = \emptyset\right), \quad (3.20)$$

where the  $x_i$ 's are the random eigenvalues of  $\mathbf{M}_N$ .

This result has been obtained by Mo when  $\Sigma_N$  has exactly two distinct eigenvalues [44].

As advocated in Section 3.2, the precise decay for  $g'_N(\mathbf{c}_N) \rightarrow 0$  does influence the eigenvalues local behavior near a cusp (see Proposition 3.6 and the discussion below). Our assumption (3.18) covers the general case where  $\sqrt{N}g'_N(\mathbf{c}_N) \rightarrow 0$ , and hence the limiting kernel is  $K_{\text{Pearcey}}^{(0)}(x, y)$  introduced by Brézin and Hikami, and the limiting regime where  $\sqrt{N}g'_N(\mathbf{c}_N)$  has a limit as well.

**Remark 3.13. (erosion of a valley)** In the case where this limit  $\kappa$  in (3.18) is positive, the deterministic equivalent measure  $\mu_N$  will not feature a cusp but rather a valley that will become deeper as  $N \rightarrow \infty$ , see the thin curve in Figure 5. The density of  $\mu_N$  will always be positive near the cusp and the condition

$$g'_N(\mathbf{c}_N) \sim \frac{\kappa}{\sqrt{N}}$$

should be thought of as a speed condition of the erosion of the valley.

**Remark 3.14. (moving cliffs)** In the case where  $\kappa < 0$  in (3.18),  $g'_N(\mathbf{c}_N)$  is always negative for  $N$  large enough. In particular, there exists a small  $N$ -neighborhood of  $\mathbf{c}_N$  whose image by  $g_N$  is outside the support of  $\mu_N$ : There is a small hole in the support of  $\mu_N$  but the two connected components move towards one another (moving cliffs), see the dotted curve in Figure 5. In this case, the condition

$$g'_N(\mathbf{c}_N) \sim \frac{\kappa}{\sqrt{N}}$$

can also be interpreted as a speed condition at which the cliffs approach one another.

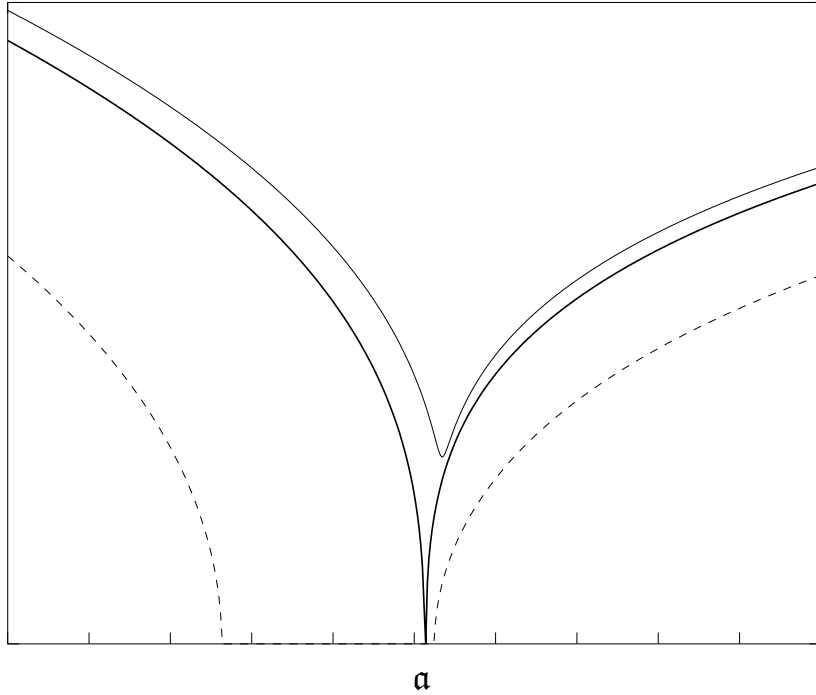


Figure 5: Zoom of the density of  $\mu_N$  near the cusp point  $\mathbf{a}$ . The thick curve is the density of  $\mu$  in the framework of Figure 4. The thin curve (resp. the dotted curve) is the density of  $\mu_N$  when  $\sqrt{N}g'_N(\mathbf{c}_N) > 0$  (resp.  $\sqrt{N}g'_N(\mathbf{c}_N) < 0$ ).

**Remark 3.15. (slow decay)** The slow decay setting where

$$\sqrt{N}g'_N(\mathbf{c}_N) \rightarrow \pm\infty$$

is not covered by our results. In this case, we do not expect the Pearcey point process to arise anymore, and refer to Section 5 for further discussion.

## 4 Sketches of the proofs

In this section, we provide an outline for the proofs of the results presented in Section 3.

### 4.1 The random eigenvalues of $\mathbf{M}_N$ form a determinantal point process

The key input on which all our proofs are based on, is that when the elements of  $\mathbf{X}_N$  are complex Gaussian (Assumption 1), the configuration of the random eigenvalues  $x_i$ 's of  $\mathbf{M}_N$  form a determinantal point process with an explicit kernel. More precisely Baik, Ben Arous and P      provided in [7] a formula for that kernel, to which they give credit to Johansson. It is given by the following double complex integral

$$K_N(x, y) = \frac{N}{(2i\pi)^2} \oint_{\Gamma} dz \oint_{\Theta} dw \frac{1}{w - z} e^{-Nx(z - \mathfrak{q}) + Ny(w - \mathfrak{q})} \left(\frac{z}{w}\right)^N \prod_{j=1}^n \left(\frac{w - \lambda_j^{-1}}{z - \lambda_j^{-1}}\right), \quad (4.1)$$

where the  $\mathfrak{q} \in \mathbb{R}$  is a free parameter (see [30, Remark 4.3]) and we recall the  $\lambda_j$ 's are the eigenvalues of  $\Sigma_N$ . The contours  $\Gamma$  and  $\Theta$  are disjoint and closed, both oriented counterclockwise, such that  $\Gamma$  encloses all the  $\lambda_j^{-1}$ 's whereas  $\Theta$  encloses the origin.

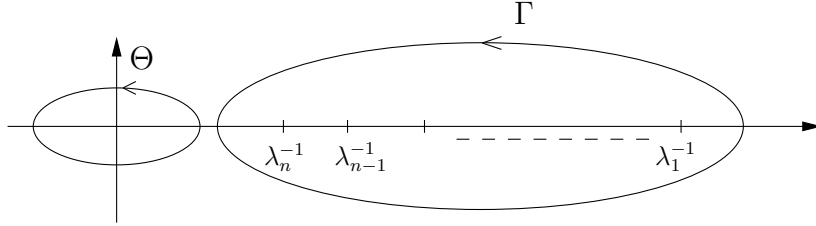


Figure 6: The contours of integration

**Remark 4.1.** The main ingredient to obtain this determinantal representation is the Harish-Chandra-Itzykson-Zuber integral formula, which allows to write a particular integral over the unitary group in terms of determinants, see [7, Section 2.1]. The analogue of this integral formula does not seem to exist for correlated Wishart matrices with real or quaternionic entries, and thus the determinantal structure seems only available in the complex setting.

If we consider a random configuration of the form

$$(N^\beta \sigma_N (x_i - \mathfrak{a}_N)) ,$$

then a change of variables yields it is a determinantal point process with kernel

$$\frac{1}{N^\beta \sigma_N} K_N \left( \mathfrak{a}_N + \frac{x}{N^\beta \sigma_N}, \mathfrak{a}_N + \frac{y}{N^\beta \sigma_N} \right), \quad (4.2)$$

where  $K_N$  is as in (4.1). Hence, the study of the eigenvalues local behavior boils down to the asymptotic analysis as  $N \rightarrow \infty$  of kernels of the form (4.2) with different choices for the scaling parameters  $\beta, \sigma_N, \mathfrak{a}_N$ .

## 4.2 Modes of convergence

In order to prove the convergence (3.14) at the hard edge, it is enough to establish a local uniform convergence on  $\mathbb{R}_+ \times \mathbb{R}_+$  for the kernel (4.2) to the Bessel kernel  $K_{\text{Bessel}}^{(\alpha)}(x, y)$ , after choosing appropriately the scaling parameters. Similarly, the local uniform convergence on  $\mathbb{R} \times \mathbb{R}$  to the Pearcey kernel  $K_{\text{Pearcey}}^{(\tau)}(x, y)$  yields the convergence (3.20) for the gap probabilities around a cusp. The convergences (3.7) and (3.10) to the Tracy-Widom law however require a stronger mode of convergence (such as the trace-class norm convergence, or the Hilbert-Schmidt norm plus trace convergence, for the associated operators acting on  $L^2(s, \infty)$ , for every  $s \in \mathbb{R}$ ; we refer to [30, Section 4.2] for further information). This essentially amounts to obtain a local uniform convergence on  $(s, +\infty) \times (s, +\infty)$  plus tail estimates for  $K_N(x, y)$ .

From now, we shall disregard these convergence issues and provide heuristics on why the Airy kernel, the Bessel kernel and the Pearcey kernel should appear in different scaling limits.

## 4.3 Towards the Airy kernel

Here we provide an heuristic for the convergence to the Airy kernel. The gap to be filled in order to make this sketch of a proof mathematically rigorous can be found in [30]; this heuristic may actually serve as a roadmap for the quite lengthy and technical proof we provided there.

Since we are dealing with contours integrals, it is more convenient to use the following alternative representation for the Airy kernel (3.3),

$$K_{\text{Airy}}(x, y) = \frac{1}{(2i\pi)^2} \int_{\Xi} dz \int_{\Xi'} dw \frac{1}{w - z} e^{-xz + yw + z^3/3 - w^3/3}, \quad (4.3)$$

which is based on the contour integral formula for the Airy function (see e.g. the proof of [30, Lemma 4.15]). The contours  $\Xi$  and  $\Xi'$  are disjoint and unbounded contours, and  $\Xi$  goes from  $e^{i\pi/3}\infty$  to  $e^{-i\pi/3}\infty$  whereas  $\Xi'$  goes from  $e^{-2i\pi/3}\infty$  to  $e^{2i\pi/3}\infty$ , as shown on Figure 7.

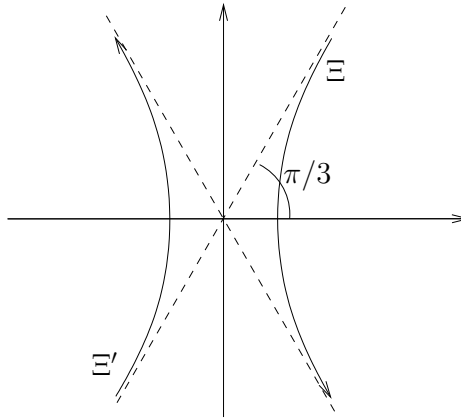


Figure 7: The paths of integration for the Airy kernel

Consider the scaling parameters associated with a soft edge provided in Theorem 1; we thus focus on the right edge setting, but the situation for a left edge is similar. More precisely,

by using the formula (4.1) where we take  $\mathbf{q} = \mathbf{c}_N$ , we investigate

$$\begin{aligned} & \frac{1}{N^{2/3}\sigma_N} K_N \left( \mathbf{a}_N + \frac{x}{N^{2/3}\sigma_N}, \mathbf{a}_N + \frac{y}{N^{2/3}\sigma_N} \right) \\ &= \frac{N^{1/3}}{(2i\pi)^2\sigma_N} \oint_{\Gamma} dz \oint_{\Theta} dw \frac{1}{w-z} e^{-N^{1/3}x(z-\mathbf{c}_N)/\sigma_N + N^{1/3}y(w-\mathbf{c}_N)/\sigma_N} e^{Nf_N(z) - Nf_N(w)}, \end{aligned} \quad (4.4)$$

where we introduced the map

$$f_N(z) = -\mathbf{a}_N(z - \mathbf{c}_N) + \log(z) - \frac{1}{N} \sum_{j=1}^n \log(1 - \lambda_j z). \quad (4.5)$$

After performing the change of variables  $z \mapsto \mathbf{c}_N + \sigma_N z / N^{1/3}$  and  $w \mapsto \mathbf{c}_N + \sigma_N w / N^{1/3}$ , the right-hand side of (4.4) becomes

$$\frac{1}{(2i\pi)^2} \oint_{\varphi_N(\Gamma)} dz \oint_{\varphi_N(\Theta)} dw \frac{1}{w-z} e^{-xz+yw+Nf_N(\mathbf{c}_N+\sigma_N \frac{z}{N^{1/3}}) - Nf_N(\mathbf{c}_N+\sigma_N \frac{w}{N^{1/3}})}, \quad (4.6)$$

where we set for convenience  $\varphi_N(z) = N^{1/3}(z - \mathbf{c}_N)/\sigma_N$ .

Next, recalling  $g_N$  was introduced in (3.2), the crucial observation that

$$f'_N(z) = g_N(z) - \mathbf{a}_N \quad (4.7)$$

allows to infer on the local behavior of  $f_N$  around  $\mathbf{c}_N$ . More precisely, since by definition of the scaling parameters we have

$$g_N(\mathbf{c}_N) = \mathbf{a}_N, \quad g'_N(\mathbf{c}_N) = 0, \quad g''_N(\mathbf{c}_N) \xrightarrow{N \rightarrow \infty} g''(\mathbf{c}) > 0,$$

a Taylor expansion for  $f_N$  around  $\mathbf{c}_N$  yields the approximation

$$N \left( f_N(\mathbf{c}_N + \sigma_N \frac{z}{N^{1/3}}) - f_N(\mathbf{c}_N) \right) \simeq \frac{1}{3} z^3, \quad (4.8)$$

where the constant  $1/3$  comes from the definition of  $\sigma_N$ , see (3.6). In conclusion, after plugging (4.8) into (4.6), we obtain the approximation

$$\begin{aligned} & \frac{1}{N^{2/3}\sigma_N} K_N \left( \mathbf{a}_N + \frac{x}{N^{2/3}\sigma_N}, \mathbf{a}_N + \frac{y}{N^{2/3}\sigma_N} \right) \\ & \simeq \frac{1}{(2i\pi)^2} \oint_{\varphi_N(\Gamma)} dz \oint_{\varphi_N(\Theta)} dw \frac{1}{w-z} e^{-xz+yw+z^3/3-w^3/3}, \end{aligned} \quad (4.9)$$

and we can almost read the Airy kernel (4.3), up to contour deformations.

To frame the previous heuristic into a rigorous mathematical setting, a few technical points should be addressed, since of course the approximation (4.8) is only valid when  $|z|$  is not too large and the contours appearing in the Airy kernel are unbounded. With this respect, the standard move is to split the contours  $\Gamma$  and  $\Theta$  into different parts and then to deform each part in an appropriate way.

In a neighborhood of  $\mathbf{c}_N$ , after simple transformations, one chooses  $\Gamma$  and  $\Theta$  to match with the contours of the Airy kernel there, and then justify rigorously the approximation (4.9) after

restriction of  $z, w$  to that neighborhood. This can be done by quantifying the approximation (4.8) and then performing tedious but rather simple computations.

Then, outside of this neighborhood, one proves that the remaining of the integrals don't contribute in the large  $N$  limit. In the present setting of a general matrix  $\Sigma_N$ , this is the hard part of the proof. To do so, one establishes the existence of admissible deformations for the contours  $\Gamma$  and  $\Theta$  so that they complete the Airy contours truncated on a neighborhood of  $\mathbf{c}_N$ , and where the contribution coming from the term  $\exp\{Nf_N(z) - Nf_N(w)\}$  brings an exponential decay on the remaining part. This can be done by looking for the so-called steepest descent/ascent contours (i.e. contours on which  $\text{Re } f_N$  is decreasing/increasing), and this was the strategy used by Baik, Ben Arous, P      [7] and El Karoui [26] when dealing with the rightmost edge. When considering any right or left soft edge, following this strategy requires to consider many sub-cases and to perform again most of the computations in several case. In [30], we instead developed a unified (abstract) method to provide the existence of appropriate contours by mean of the maximum principle for subharmonic functions.

For the reader interested in having a look at the proofs of [30], let us mention it turns out it is more convenient to work at a scale where the contours  $\Gamma, \Theta$  live in a bounded domain, and this is the reason why we did not performed there the changes of variables  $z \mapsto \mathbf{c}_N + \sigma_N z/N^{1/3}$  and  $w \mapsto \mathbf{c}_N + \sigma_N w/N^{1/3}$  as we did in the present heuristic.

#### 4.4 Towards the Pearcey kernel

Now we turn to the heuristics for Pearcey kernel and refer to [31] for a rigorous proof. The setting is essentially the same as in the Airy case, except that now  $\mathbf{c}_N$  is a simple zero for  $g_N''$  instead of  $g_N'$ , and  $g_N'(\mathbf{c}_N) \rightarrow 0$ , which entails a different behavior for the map  $f_N$  near  $\mathbf{c}_N$ .

We start with the alternative representation for the Pearcey kernel (3.16),

$$K_{\text{Pearcey}}^{(\tau)}(x, y) = \frac{1}{(2i\pi)^2} \int_{\Xi} dz \int_{-i\infty}^{i\infty} dw \frac{1}{w - z} e^{-xz - \frac{\tau z^2}{2} + \frac{z^4}{4} + yw + \frac{\tau w^2}{2} - \frac{w^4}{4}}, \quad (4.10)$$

where the contour  $\Xi$  is disjoint to the imaginary axis, and has two components. The first part goes from  $e^{i\pi/4}\infty$  to  $e^{-i\pi/4}\infty$ , whereas the other part goes from  $e^{-3i\pi/4}\infty$  to  $e^{3i\pi/4}\infty$ . See [53] for a proof (and also [16] when  $\tau = 0$ ). Notice also the symmetry  $K_{\text{Pearcey}}^{(\tau)}(x, y) = K_{\text{Pearcey}}^{(\tau)}(-x, -y)$  which follows from the change of variables  $z, w \mapsto -z, -w$ .

Consider the scaling parameters associated with a regular cusp point provided in Theorem 5. By using the formula (4.1) where we choose  $\mathbf{q} = \mathbf{c}_N$ , we now consider

$$\begin{aligned} & \frac{1}{N^{3/4}\sigma_N} K_N \left( \mathbf{a}_N + \frac{x}{N^{3/4}\sigma_N}, \mathbf{a}_N + \frac{y}{N^{3/4}\sigma_N} \right) \\ &= \frac{N^{1/4}}{(2i\pi)^2\sigma_N} \oint_{\Gamma} dz \oint_{\Theta} dw \frac{1}{w - z} e^{-N^{1/4}x \frac{(z - \mathbf{c}_N)}{\sigma_N} + N^{1/4}y \frac{(w - \mathbf{c}_N)}{\sigma_N}} e^{Nf_N(z) - Nf_N(w)}, \end{aligned} \quad (4.11)$$

where the map  $f_N$  is the same as in (4.5). After the change of variables  $z \mapsto \mathbf{c}_N + \sigma_N z/N^{1/4}$  and  $w \mapsto \mathbf{c}_N + \sigma_N w/N^{1/4}$ , the right-hand side of (4.11) reads

$$\frac{1}{(2i\pi)^2} \oint_{\varphi_N(\Gamma)} dz \oint_{\varphi_N(\Theta)} dw \frac{1}{w - z} e^{-xz + yw + Nf_N(\mathbf{c}_N + \sigma_N \frac{z}{N^{1/4}}) - Nf_N(\mathbf{c}_N + \sigma_N \frac{w}{N^{1/4}})}, \quad (4.12)$$

where we introduced for convenience  $\varphi_N(z) = N^{1/4}(z - \mathbf{c}_N)/\sigma_N$ .

In this setting, the definition of the scaling parameters yields

$$g_N(\mathbf{c}_N) = \mathbf{a}_N, \quad \sqrt{N}g'_N(\mathbf{c}_N) \xrightarrow{N \rightarrow \infty} \kappa, \quad g''_N(\mathbf{c}_N) = 0, \quad g'''_N(\mathbf{c}_N) \xrightarrow{N \rightarrow \infty} g'''(\mathbf{c}) > 0.$$

Recalling the identity (4.7) and the definition (3.19) of  $\tau$ , a Taylor expansion around  $\mathbf{c}_N$  then yields the approximation

$$N \left( f_N(\mathbf{c}_N + \sigma_N \frac{z}{N^{1/4}}) - f_N(\mathbf{c}_N) \right) \simeq -\frac{\tau}{2}z^2 + \frac{1}{4}z^4, \quad (4.13)$$

where the constant  $1/4$  comes from the definition of  $\sigma_N$ , see (3.19). Thus, by plugging (4.13) into (4.12), we obtain the approximation

$$\begin{aligned} \frac{1}{N^{3/4}\sigma_N} K_N \left( \mathbf{a}_N + \frac{x}{N^{3/4}\sigma_N}, \mathbf{a}_N + \frac{y}{N^{3/4}\sigma_N} \right) \\ \simeq \frac{1}{(2i\pi)^2} \oint_{\varphi_N(\Gamma)} dz \oint_{\varphi_N(\Theta)} dw \frac{1}{w-z} e^{-xz - \frac{\tau z^2}{2} + \frac{z^4}{4} + yw + \frac{\tau w^2}{2} - \frac{w^4}{4}}, \end{aligned} \quad (4.14)$$

and we can almost see the Pearcey kernel (4.10), up to contour deformations.

As to make this approximation rigorous, the method is the same as for the Airy kernel. Let us mention that the abstract argument we mention previously for the existence of appropriate contour deformations also applies in this setting.

## 4.5 Towards the Bessel kernel

Finally, we provide heuristics for the appearance of the Bessel kernel and refer to [31] for a rigorous proof. The main input here is, according to Section 2, the critical point  $\mathbf{c}$  associated with the hard edge is now located at infinity.

The first step is to write the Bessel kernel (3.11) as the double contour integral,

$$K_{\text{Bessel}}^{(\alpha)}(x, y) = \frac{1}{(2i\pi)^2} \left( \frac{y}{x} \right)^{\alpha/2} \oint_{|z|=r} \frac{dz}{z} \oint_{|w|=R} \frac{dw}{w} \frac{1}{z-w} \left( \frac{z}{w} \right)^{\alpha} e^{-\frac{x}{z} + \frac{z}{4} + \frac{y}{w} - \frac{w}{4}}, \quad (4.15)$$

where  $0 < r < R$ , and which is provided in [30, Lemma 6.2]. The contours of integration are circles oriented counterclockwise. Let us stress this formula is only available when  $\alpha \in \mathbb{Z}$ , since otherwise the term  $(z/w)^{\alpha}$  in the integrand would not make sense on the whole of the integration contours.

Setting  $\sigma_N$  as in Theorem 4 and using the formula (4.1) where we choose  $\mathbf{q} = 0$ , we now consider

$$\begin{aligned} \frac{1}{N^2\sigma_N} K_N \left( \frac{x}{N^2\sigma_N}, \frac{y}{N^2\sigma_N} \right) \\ = \frac{1}{(2i\pi)^2 N\sigma_N} \oint_{\Gamma} dz \oint_{\Theta} dw \frac{1}{w-z} \left( \frac{z}{w} \right)^N e^{-\frac{zx}{N\sigma_N} + \frac{wy}{N\sigma_N}} \prod_{j=1}^n \frac{w - \lambda_j^{-1}}{z - \lambda_j^{-1}}. \end{aligned} \quad (4.16)$$

Having in mind the critical point is located at infinity, we perform the change of variables  $z \mapsto N\sigma_N/z$  and  $w \mapsto N\sigma_N/w$ , so that the right-hand side of (4.16) reads

$$\frac{1}{(2i\pi)^2} \oint_{\varphi_N(\Gamma)} \frac{dz}{z} \oint_{\varphi_N(\Theta)} \frac{dw}{w} \frac{1}{z-w} \left( \frac{z}{w} \right)^{\alpha} e^{-\frac{x}{z} + \frac{y}{w} - NG_N(z) + NG_N(w)}, \quad (4.17)$$



where we introduced the maps

$$G_N(z) = \frac{1}{N} \sum_{j=1}^n \log \left( \frac{z}{N\sigma_N} - \lambda_j \right)$$

and  $\varphi_N(z) = N\sigma_N/z$ . We emphasize that, during the previous step, we used that  $n = N + \alpha$  and witnessed a cancellation leading to the term  $(z/w)^\alpha$ , which does not depend on  $N$ .

Now, a Taylor expansion of  $G_N$  around zero yields the approximation

$$N \left( G_N(z) - G_N(0) \right) \simeq -\frac{z}{4N}. \quad (4.18)$$

Thus, by plugging (4.18) into (4.17), we obtain the approximation

$$\begin{aligned} \frac{1}{N^2\sigma_N} K_N \left( \frac{x}{N^2\sigma_N}, \frac{y}{N^2\sigma_N} \right) \\ \simeq \frac{1}{(2i\pi)^2} \oint_{\varphi_N(\Gamma)} \frac{dz}{z} \oint_{\varphi_N(\Theta)} \frac{dw}{w} \frac{1}{z-w} \left( \frac{z}{w} \right)^\alpha e^{-\frac{x}{z} + \frac{y}{w} + \frac{z}{4} - \frac{w}{4}}, \end{aligned} \quad (4.19)$$

and we can almost see the Bessel kernel (4.15), up to contour deformations and the prefactor  $(y/x)^{\alpha/2}$ . Finally, in order to deal with that prefactor, one considers the operator  $E$  of multiplication by  $x^{\alpha/2}$  acting on  $L^2(0, s)$ , and then use that replacing the Bessel kernel  $K_{\text{Bessel}}^{(\alpha)}(x, y)$  by the kernel of the operator  $E K_{\text{Bessel}}^{(\alpha)} E^{-1}$ , which is (4.15) without the prefactor  $(y/x)^{\alpha/2}$ , leaves the Fredholm determinant  $\det(I - K_{\text{Bessel}}^{(\alpha)})_{L^2(0, s)}$  invariant.

To make this heuristic rigorous, the method is far less demanding than in the setting of the Airy or the Pearcey kernel. Indeed, in the present setting one can legally deform the contours  $\Gamma$  and  $\Theta$  in such a way that  $\varphi_N(\Gamma)$  and  $\varphi_N(\Theta)$  match the integration contours for the Bessel kernel (4.15). After making this deformation, a simple Taylor expansion of the map  $G_N$  around zero will be enough to establish the convergence towards the Bessel kernel and therefore to obtain (3.14).

By pushing the Taylor expansion (4.18) one step further, one can also obtain the more accurate estimate (3.15), provided with an identity involving the resolvent of the Bessel kernel established by Tracy and Widom [52]. We refer the reader to [31] for further information.

## 5 Open questions

The results presented here naturally entail a number of open questions that we list below.

1. **At the edge of the definition domain and exotic local behaviors.** The results on the eigenvalues local behavior presented in this survey only concern edges or cusp points  $\mathfrak{a}$  which read  $\mathfrak{a} = g(\mathfrak{c})$  with  $\mathfrak{c} \in D$ . If we focus on the rightmost edge for the sake of simplicity, then Proposition 2.1 states this edge may actually belong to  $g(\partial D)$  (notice that this cannot happen if the limiting spectral distribution  $\nu$  of  $\Sigma_N$  is a finite combination of Dirac measures). In this case, the square root behavior of the density around this edge is not guaranteed anymore, and the laws describing the fluctuations of the eigenvalues near such an edge seem completely unknown and a priori different from Tracy-Widom distribution. We believe the fluctuations will actually depend on  $\nu$ , and

hence lie outside of the random matrix universality class. Quite interestingly, the same phenomenon arise in the study of the additive deformation of a GUE random matrix [20] and random Gelfand-Tsetlin patterns [22, 23].

2. **Alternative regime at a cusp point I.** In the context of Theorem 5, our speed assumption (3.18) does not cover the following case

$$\sqrt{N} g'_N(\mathfrak{c}_N) \xrightarrow{N \rightarrow \infty} +\infty .$$

This condition corresponds to the situation where the density of the deterministic equivalent  $\mu_N$  is positive in a neighborhood of  $\mathfrak{c}_N$ . It essentially states that the bulk of  $\mu_N$  will degenerate into a cusp around  $g(\mathfrak{c})$  quite slowly and we do not expect to witness Pearcey-like fluctuations around  $g_N(\mathfrak{c}_N)$  anymore. We believe instead that the sine kernel will arise at the scale  $\sqrt{N/g'_N(\mathfrak{c}_N)}$ , which strictly lies in between  $N^{1/2}$  and  $N^{3/4}$ .

3. **Alternative regime at a cusp point II.** Another case that is not covered by our assumption (3.18) is when

$$\sqrt{N} g'_N(\mathfrak{c}_N) \xrightarrow{N \rightarrow \infty} -\infty .$$

In this case,  $\mathfrak{c}_N$  lies outside the support of  $\mu_N$  and  $g'_N$  has two distinct real zeroes near  $\mathfrak{c}_N$ , say  $\mathfrak{c}_{N,1}$  and  $\mathfrak{c}_{N,2}$ . Hence, for  $N$  sufficiently large,  $g_N(\mathfrak{c}_{N,1})$  and  $g_N(\mathfrak{c}_{N,2})$  both correspond to edges of the support of  $\mu_N$  which both converge towards the cusp point  $g(\mathfrak{c})$ . The previous condition entails that such a convergence will happen at a quite slow rate. In this case, we do not expect to observe the Pearcey kernel around  $g_N(\mathfrak{c}_N)$  either, because of the absence of particles, but we believe a local analysis around the edge  $g_N(\mathfrak{c}_{N,1})$  or  $g_N(\mathfrak{c}_{N,2})$  may uncover the Airy kernel at an intermediary scale.

4. **Study of the fluctuations at the hard edge in more general cases.** The hard edge fluctuations were described here when  $n = N + \alpha$  with a fixed  $\alpha \in \mathbb{Z}$ , but the hard edge is always present as soon as  $n/N \rightarrow 1$ . Thus it would be of interest to describe the hard edge fluctuations in more general situations, for example when  $\alpha = \alpha(n) \rightarrow +\infty$  so that  $n/N \rightarrow 1$ . In this case one would expect Tracy-Widom fluctuations near the leftmost edge  $\mathfrak{a}_N$  of  $\mu_N$ , the latter being positive but converging to zero as  $N \rightarrow \infty$ .

## 5. Non-Gaussian entries.

All the fluctuations results presented here rely on the fact that the entries of matrix  $\mathbf{X}_N$  are complex Gaussian. It is however of major interest, for applications and for the general theory as well, to study the universality of such results for non-Gaussian complex random variables. As explained in Remark 3.9, the Tracy-Widom fluctuations for the extremal eigenvalues associated with any regular soft edges are now established in this general setting (under some moment conditions for the entries), by combining theorems 1 and 2 with Knowles and Yin's recent preprint [37] (see also [8]). However, natural related questions remain open: Would it be possible to describe the fluctuations at

- (a) the hard edge, for general complex entries?
- (b) a (regular) cusp point, for general complex entries?

Another universality class of interest is the case where the matrix  $\mathbf{X}_N$  has real entries. In this case, the techniques based on the determinantal structure of the eigenvalues are no longer available. Lee and Schnelli [40] recently succeeded to establish GOE Tracy-Widom fluctuations of the largest eigenvalue when the entries are real Gaussian or simply real (with subexponential decay), under the assumption that the covariance matrix  $\Sigma_N$  is diagonal. The techniques developed by Knowles and Yin [37] enable to relax the diagonal assumption for the covariance matrix  $\Sigma_N$ . A number of questions remain open: Would it be possible to describe the fluctuations at

- (c) any (regular) soft edge, when the entries are real Gaussian?
- (d) the hard edge, when the entries are real (Gaussian or not)?
- (e) a (regular) cusp point, when the entries are real (Gaussian or not)?

## References

- [1] M. Adler, M. Cafasso, P. van Moerbeke, *Nonlinear PDEs for gap probabilities in random matrices and KP theory*, Physica D: Nonlinear Phenomena **241** (2012), 2265–2284.
- [2] G. Anderson, A. Guionnet, O. Zeitouni, *An Introduction to Random Matrices*, Cambridge University Press, 2009.
- [3] Z. Bai and J. Silverstein, *Spectral analysis of large dimensional random matrices*, Springer Series in Statistics. Springer, New York, second edition, 2010.
- [4] Z. Bai, Y. Chen, Y C. Liang (editors): Random matrix theory and its applications, volume 18 of Lecture Notes Series. Institute for Mathematical Sciences. National University of Singapore. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2009, ISBN 978-981-4273-11-4; 981-4273-11-2, Multivariate statistics and wireless communications, Papers from the workshop held at the National University of Singapore, Singapore, 2006.
- [5] Z. Bai, J. Silverstein, *No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices*, Ann. Probab., **26** (1998), 316–345.
- [6] Z. Bai, J. Silverstein, *Exact separation of eigenvalues of large-dimensional sample covariance matrices*, Ann. Probab. **27** (1999), 1536–1555.
- [7] J. Baik, G. Ben Arous, S. Péché, *Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices*, Ann. Probab. **33** (2005), 1643–1697.
- [8] Z. Bao, G. Pan, W. Zhou, *Universality for the largest eigenvalue of sample covariance matrices with general population*, Ann. Statist. Volume 43 (**1**) (2015), 382–421.
- [9] E. Basor, Y. Chen, L. Zhang, *PDEs satisfied by extreme eigenvalues distributions of GUE and LUE*, Random Matrices Theory Appl. **1** (2012), 1150003.
- [10] F. Benaych-Georges, R. R. Nadakuditi, *The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices*, Adv. Math. **227**(1) (2011), 494–521.
- [11] M. Bertola, M. Cafasso, *The transition between the gap probabilities from the Pearcey to the Airy Process; a Riemann–Hilbert approach*, Int. Math. Res. Not. 2012 no. 7, 1519–1568.

- [12] P. Bianchi, M. Debbah, M. Maïda and J. Najim, *Performance of statistical tests for single source detection using random matrix theory*, IEEE Trans. Inf. Theory, **57** (2011), 2400–2419.
- [13] A. Bloemendal, A. Knowles, H.-T. Yau, J. Yin, *On the principal components of sample covariance matrices*, to appear in Prob. Theor. Rel. Fields. Preprint 75p (2014), arXiv:1404.0788.
- [14] F. Bornemann, *A note on the expansion of the smallest eigenvalue distribution of the LUE at the hard edge*. Preprint 3p (2015), arXiv:1504.00235.
- [15] A. Borodin, P. Forrester, *Increasing subsequences and the hard-to-soft edge transition in matrix ensembles*, J. Phys. A **36** (2003), 2963–2981.
- [16] E. Brézin, S. Hikami, *Universal singularity at the closure of a gap in a random matrix theory*, Phys. Rev. E **57** (1998), 4140–4149.
- [17] E. Brézin, S. Hikami, *Level spacing of random matrices in an external source*, Phys. Rev. E **58** (1998), 7176–7185.
- [18] R. Couillet, M. Debbah, Random matrix methods for wireless communications, Cambridge University Press (2011), ISBN 978-1-107-01163-2.
- [19] M. Capitaine, *Additive/multiplicative free subordination property and limiting eigenvectors of spiked additive deformations of Wigner matrices and spiked sample covariance matrices*, J. Theor. Probab. **26** (2013), 595–648.
- [20] M. Capitaine, S. Péché, *Fluctuations at the edge of the spectrum of the full rank deformed GUE*, preprint 44p (2014), arXiv:1402.2262.
- [21] P. Deift, *Orthogonal polynomials and random matrices: a Riemann-Hilbert approach*, Courant Lecture Notes in Mathematics Vol. 3, AMS Providence R.I. (1999).
- [22] E. Duse, A. Metcalfe, *Asymptotic Geometry of Discrete Interlaced Patterns: Part I*, preprint 57p (2015), arXiv:1412.6653.
- [23] E. Duse, A. Metcalfe, *Asymptotic Geometry of Discrete Interlaced Patterns: Part II*, preprint 71p (2015), arXiv:1507.00467.
- [24] A. Edelman, *Eigenvalues and condition numbers of random matrices*, SIAM J. Matrix Anal. Appl. **9** (1988), 543–560.
- [25] A. Edelman, A. Guionnet, S. Péché, *Beyond universality in random matrix theory*, Preprint 42p (2014), arXiv:1405.7590.
- [26] N. El Karoui, *Tracy-Widom limit for the largest eigenvalue of a large class of complex sample covariance matrices*, Ann. Probab. **35** (2007), 663–714.
- [27] L. Erdős, *Universality of Wigner random matrices: a survey of recent results*, Russ. Math. Surv. **66** (2011), 507.
- [28] P. J. Forrester, *The spectrum edge of random matrix ensembles*, Nucl. Phys. B **402** (1993), 709–728.

- [29] H. H. Goldstine, J. von Neumann, *Numerical inverting of matrices of high order. II*, Proc. Amer. Math. Soc., **2** (1951), 188–202.
- [30] W. Hachem, A. Hardy, J. Najim, *Large complex correlated Wishart matrices: Fluctuations and asymptotic independence at the edges*, to appear in Ann. Probab. Preprint 76p (2014) arXiv:1409.7548.
- [31] W. Hachem, A. Hardy, J. Najim, *Large complex correlated Wishart matrices: The Pearcey kernel and expansion at the hard edge*. Preprint (2015) arXiv:1507.06013.
- [32] F. Hiai, D. Petz, *The Semicircle Law, Free Random Variables and Entropy*, Mathematical Surveys and Monographs Volume 77, American Mathematical Society.
- [33] J. B. Hough, M. Krishnapur, Y. Peres, B. Virág, *Determinantal processes and independence*, Probab. Surveys **3** (2006), 206–229.
- [34] K. Johansson, *Shape fluctuations and random matrices*, Comm. Math. Phys. **209** (2000), 437–476.
- [35] K. Johansson, *Random matrices and determinantal processes*, Mathematical Statistical Physics: Lecture Notes of the Les Houches Summer School 2005 (Bovier et al., eds.), Elsevier, 2006, 1–55.
- [36] I.M. Johnstone, *On the distribution of the largest eigenvalue in principal components analysis*, Ann. Statist., **29** (2001), 295–327.
- [37] A. Knowles, J. Yin. *Anisotropic local laws for random matrices*, Preprint 60p (2014) arXiv:1410.3516.
- [38] S. Kritchman and B. Nadler, *Non-parametric detection of the number of signals: hypothesis testing and random matrix theory*, IEEE Trans. Signal Processing, **57** (2009), 3930–3941.
- [39] L. Laloux, P. Cizeau, J.-P. Bouchaud, M. Potters, *Noise dressing of financial correlation matrices*, Physical Review Letters **83** 1467 (1999).
- [40] J.O. Lee, K. Schnelli. *Tracy-Widom distribution for the largest eigenvalue of real sample covariance matrices with general population*, Preprint 36p (2014) arXiv:1409.4979.
- [41] V. Lysov, F. Wielonsky, *Strong asymptotics for multiple Laguerre polynomials*, Constr. Approx. **28** (2008), 61–111.
- [42] V. A. Marchenko and L. A. Pastur, *Distribution of eigenvalues for some sets of random matrices*, Math. USSR-Sb. **1** (1967), 457–483.
- [43] M. Y. Mo, *Universality in Complex Wishart ensembles: The 2 cut case*, Preprint 41p (2008) arXiv:0809.3750.
- [44] M. Y. Mo, *Universality in complex Wishart ensembles for general covariance matrices with 2 distinct eigenvalues*, Journal of Multivariate Analysis **101** (2009), 1203–1225.

- [45] E. Moulines, P. Duhamel, J.F. Cardoso and S. Mayrargue, *Subspace methods for the blind identification of multichannel FIR filters*, IEEE Trans. Signal Processing, **43** (1995) 516–525.
- [46] M. C. Münnix, R. Schäfer, T. Guhr, *A random matrix approach to credit risk*, PLOS ONE **9** (2014).
- [47] J. von Neumann, H. H. Goldstine, *Numerical inverting of matrices of high order*, Bull. Amer. Math. Soc. **53** (1947), 1021–1099.
- [48] A. Onatski, *The Tracy-Widom limit for the largest eigenvalues of singular complex Wishart matrices*, Ann. Appl. Probab. **18** (2008), 470–490.
- [49] A. Perret, G. Schehr, *Finite  $N$  corrections to the limiting distribution of the smallest eigenvalue of Wishart complex matrices*, preprint 28p (2015), arXiv:1506.02387.
- [50] J. W. Silverstein, S.-I. Choi, *Analysis of the limiting spectral distribution of large-dimensional random matrices*, J. Multivariate Anal. **54** (1995), 295–309.
- [51] C. Tracy, H. Widom, *Level-spacing distributions and the Airy kernel*, Comm. Math. Phys. **159** (1994), 151–174.
- [52] C. Tracy, H. Widom, *Level spacing distributions and the Bessel kernel*, Comm. Math. Phys. **161** (1994), 289–309.
- [53] C. Tracy, H. Widom, *The Pearcey process*, Commun. Math. Phys. **263** (2006), 381–400.