

# Universality in the bulk of the spectrum for complex sample covariance matrices

Sandrine Péché

*Institut Fourier, 100 Rue des Maths, BP 74, F-38402 St Martin d'Heres, France. E-mail: sandrine.peche@ujf-grenoble.fr*

Received 22 December 2009; revised 23 November 2010; accepted 15 June 2011

---

**Abstract.** We consider complex sample covariance matrices  $M_N = \frac{1}{N}YY^*$  where  $Y$  is a  $N \times p$  random matrix with i.i.d. entries  $Y_{ij}$ ,  $1 \leq i \leq N$ ,  $1 \leq j \leq p$ , with distribution  $F$ . Under some regularity and decay assumptions on  $F$ , we prove universality of some local eigenvalue statistics in the bulk of the spectrum in the limit where  $N \rightarrow \infty$  and  $\lim_{N \rightarrow \infty} p/N = \gamma$  for any real number  $\gamma \in (0, \infty)$ .

**Résumé.** On considère des matrices de covariance empirique complexes  $M_N = \frac{1}{N}YY^*$  où  $Y$  est une matrice de taille  $N \times p$  dont les coefficients  $Y_{ij}$ ,  $1 \leq i \leq N$ ,  $1 \leq j \leq p$ , sont des variables aléatoires i.i.d. de loi  $F$ . Sous certaines hypothèses de régularité et de décroissance sur  $F$ , on montre l'universalité de certaines statistiques locales de valeurs propres au milieu du spectre quand  $N \rightarrow \infty$  et  $\lim_{N \rightarrow \infty} p/N = \gamma$  pour tout réel  $\gamma \in (0, \infty)$ .

MSC: 60B20; 60B10; 60B12

Keywords: Random matrix; Bulk universality; Sample covariance matrices

---

## 1. Introduction

### 1.1. Model and result

This paper is concerned with universal properties of large complex sample covariance matrices in the bulk of the spectrum. We consider  $N \times p$  random matrices  $Y = (Y_{ij})$  where the  $Y_{ij}$ 's are i.i.d. random variables with some probability distribution  $F$ . Let then  $M_N$  be the sample covariance matrix:

$$M_N = \frac{1}{N}YY^*.$$

In the whole paper we assume that  $p \geq N$  and that

$$\exists \gamma \in [1, \infty) \text{ such that } p/N \rightarrow \gamma \text{ as } N \rightarrow \infty.$$

The case where  $p < N$  can be deduced from the above setting using the fact that  $YY^*$  and  $Y^*Y$  have the same non-zero eigenvalues. In the sequel we call  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$  the ordered eigenvalues of  $M_N$  and  $\pi_N := \frac{1}{N} \sum_{i=1}^N \delta_{\lambda_i}$  its spectral measure. Assuming that  $F$  has a finite variance  $\sigma^2$ , Marčenko and Pastur ([18], see also [20]) have shown that  $\pi_N$  almost surely converges as  $N \rightarrow \infty$  to the so-called Marčenko–Pastur distribution  $\rho_{\gamma, \sigma^2}^{\text{MP}}$ . This probability

distribution depends on  $\sigma^2$  only and not on any other detail (higher moments e.g.) of  $F$ : in this sense it is *universal*. It is defined by the density with respect to the Lebesgue measure:

$$\frac{d\rho_{\gamma, \sigma^2}^{\text{MP}}(x)}{dx} = \frac{\sqrt{(u_+ - x)(x - u_-)}}{2\pi x \sigma^2} 1_{x \in [u_-, u_+]}, \quad (1.1)$$

where  $u_+ = \sigma^2(1 + \sqrt{\gamma})^2$  and  $u_- = \sigma^2(1 - \sqrt{\gamma})^2$ .

It has been conjectured (see [19] for instance) that, in the large- $N$ -limit, some finer properties of the spectrum are also universal. For instance, the spacing between nearest neighbor eigenvalues in the vicinity of a point  $u \in (u_-, u_+)$  is expected to be universal, under the sole assumption that the variance of  $F$  is finite. The spacing is actually believed to be “more universal” than the limiting Marčenko–Pastur distribution, in the sense that it is expected to be the same as for Hermitian Wigner matrices. To investigate such local properties of the spectrum, we introduce the so-called *local eigenvalue statistics* in the bulk of the spectrum. Given a function  $f \in L^\infty(\mathbb{R}^m)$  ( $m$  fixed) with compact support, a point  $u \in [u_-, u_+]$ , and a scaling factor  $\rho_N$ , we define the local eigenvalue statistic  $S_N^{(m)}(f, u, \rho_N)$  by

$$S_N^{(m)}(f, u, \rho_N) = \sum_{i_1, \dots, i_m} f(\rho_N(\lambda_{i_1} - u), \dots, \rho_N(\lambda_{i_m} - u)), \quad (1.2)$$

where the sum is over all distinct indices from  $\{1, \dots, N\}$ . When  $u$  is in the bulk of the spectrum, that is  $u \in (u_-, u_+)$ , the natural choice for the scaling factor is  $\rho_N = N\rho_{\gamma, \sigma^2}^{\text{MP}}(u)$ , which should give the scale of the spacing between nearest neighbor eigenvalues in the vicinity of  $u$ .

We here prove universality of some local linear statistics in the bulk of the spectrum for a wide class of complex sample covariance matrices. We follow the approach used by [14] (and a series of papers [11–13]) where universality in the bulk of Wigner matrices is proved. We now define the class of matrices under consideration in this paper. Let  $\mu$  be the real Gaussian distribution with mean 0 and variance  $1/2$ . Let  $F$  be a complex probability distribution whose real and imaginary parts are of the form

$$\nu(dx) = e^{-V(x)} \mu(dx), \quad (1.3)$$

for some real function  $V$  satisfying the following assumptions:

- $V \in \mathcal{C}^6$  and there exists an integer  $k \geq 1$  such that

$$\sum_{j=1}^6 |V^{(j)}(x)| \leq C_*(1 + x^2)^k, \quad (1.4)$$

- there exist  $\delta_1, C' > 0$  such that  $\forall x \in \mathbb{R}$ ,

$$\nu(x) \leq C' e^{-\delta_1 |x|^2}. \quad (1.5)$$

This assumption can actually be relaxed as explained in Remark 1.2 below to consider distributions  $\nu$  with sub-exponential decay only.

- $F$  is normalized so that

$$\int x \, d\nu(x) = 0 \quad \text{and} \quad \int |x^2| \, d\nu(x) = 1/2. \quad (1.6)$$

In the sequel we consider sample covariance matrices  $M_N = \frac{1}{N} Y Y^*$  where  $Y = (Y_{ij})$  is a  $N \times p$  random matrix such that:

$$Y_{ij}, 1 \leq i \leq N; 1 \leq j \leq p, \text{ are i.i.d. random variables with distribution } F. \quad (1.7)$$

One shall remark that the condition (1.6) can always be achieved by rescaling and does not impact on the generality of our next results.

We now give our two main results. Let  $\epsilon > 0$  small be given.

**Theorem 1.1.** *Assume that  $F$  satisfies (1.3)–(1.6). Let also  $u \in (u_- + \epsilon, u_+ - \epsilon)$  be a point in the bulk of the spectrum and  $\rho_N = N\rho_{\gamma,1}^{\text{MP}}(u)$ . Then*

$$\lim_{N \rightarrow \infty} \mathbb{E} S_N^{(2)}(f, u, \rho_N) = \int_{\mathbb{R}^2} f(x, y) \left( 1 - \left( \frac{\sin(x-y)\pi}{\pi(x-y)} \right)^2 \right) dx dy.$$

**Remark 1.1.** *It is also possible to prove universality of local eigenvalue statistics for  $m = 3$  using the approach developed hereafter. Nevertheless to consider higher integers  $m$ , one needs to increase the regularity of  $V$  (see Remark 1.1 in [14]).*

We can also prove that the spacing distribution close to a point  $u$  in the interior of the support of Marčenko–Pastur’s law is universal. Let  $s \geq 0$  and  $(t_N)$  be a sequence such that  $\lim_{N \rightarrow \infty} t_N = +\infty$  and  $\lim_{N \rightarrow \infty} \frac{t_N}{N^{1-\beta}} = 0$  for some  $\beta > 0$ . Let  $u \in (u_- + \epsilon, u_+ - \epsilon)$  be given and  $\rho_N = N\rho_{\gamma,1}^{\text{MP}}(u)$ . Define then the “spacing function” of the eigenvalues by

$$S_N(s, u) := \frac{1}{2t_N} \#\left\{ 1 \leq j \leq N-1, \lambda_{j+1} - \lambda_j \leq \frac{s}{\rho_N}, |\lambda_j - u| \leq \frac{t_N}{\rho_N} \right\}. \quad (1.8)$$

Intuitively the expectation of the spacing function is the probability, knowing that there exists an eigenvalue in the interval  $[u - t_N, u + t_N]$ , to find its nearest neighbor within a distance  $\frac{s}{\rho_N}$ . Finally, we define

$$p(s) = \frac{d^2}{ds^2} \det(I - K)_{L^2(0,s)}, \quad \text{where } K(x, y) := \frac{\sin \pi(x-y)}{\pi(x-y)}. \quad (1.9)$$

**Theorem 1.2.** *Assume that  $F$  satisfies (1.3)–(1.6). Let also  $u \in (u_- + \epsilon, u_+ - \epsilon)$  be a point in the bulk of the spectrum. Then,*

$$\lim_{N \rightarrow \infty} \mathbb{E} S_N(s, u) = \int_0^s p(w) dw. \quad (1.10)$$

**Remark 1.2.** *Using truncation and centralization techniques, it is possible to prove both Theorem 1.1 and Theorem 1.2 when assumption (1.5) is replaced by the weaker assumption*

$$\exists C_1, C_2 > 0 \text{ such that } v(x) \leq C_1 e^{-C_2|x|}.$$

*This extension is examined in full detail in Section 5 of [14] for Wigner random matrices and readily extends to sample covariance matrices.*

The first proofs of universality in the bulk of the spectrum of large random matrices have been obtained for the so-called invariant ensembles [7–9]. Their proof relies on the fact that the joint eigenvalue density of such ensembles can be computed and the asymptotic local eigenvalue statistics can then be determined. A breakthrough in the proof of the conjecture was obtained in [17] (following the idea of [5] and [6]), proving universality in the bulk for the so-called Dyson Brownian motion model [10]. This then allowed to extend universality results to a wide class of non-invariant Hermitian random matrix ensembles – the so-called Gauss divisible ensembles (see Section 1.2 for the definition). [4] have then obtained the same universality result for Gauss divisible complex sample covariance matrices. Very recently, such universality results have been greatly improved by [22] and [14] for Hermitian Wigner random matrices. Both the papers remove the Gauss divisible assumption and only assume sufficient decay of the entries of Wigner random matrices. The approach of [22] is to make a Taylor expansion of local eigenvalue statistics of Wigner matrices. The core of the proof is then to show that these statistics depend, in the large- $N$ -limit, only on the first four moments of the entries of the Wigner matrix. Proving that any Wigner matrix can be matched to a Gauss divisible matrix with the same first 4 moments allows to prove a very general universality result. On the other hand,

the approach of [14] is to show that any Wigner matrix (under suitable decay of the entries) can be sufficiently well approximated by a Gauss divisible random matrix, so that the bulk universality follows. We also refer the reader to [15] where the two approaches are combined.

While writing this paper, a proof of universality in the bulk of the spectrum of large sample covariance matrices has been obtained in [23] in the sole case where  $p - N = O(N^{43/48})$ , but with much milder assumptions on the decay of the distribution  $\nu$ . Therein  $\nu$  satisfies  $\int |x|^{C_0} d\nu(x) < \infty$  for some sufficiently large  $C_0$  in place of our assumptions (1.4) and (1.5). Their approach is based on the ideas developed in [22].

Our paper closely follows the ideas developed in [14]. We give an overview of the proof in the next subsection.

## 1.2. The idea of the proof

Following the pioneering work of [17] and [4,5] have shown universality of local eigenvalue statistics in the bulk of the spectrum for complex sample covariance matrices when the distribution of the sample is *Gauss divisible*. We recall that a complex distribution  $\mu_G$  is Gauss divisible if there exist a complex probability distribution  $P$  and a non-trivial complex Gaussian distribution  $G$  such that  $\mu_G = P \star G$ . Equivalently [4] consider random matrices of the form

$$\tilde{M}_N = \frac{1}{N}(W + aX)(W + aX)^*, \quad (1.11)$$

where  $W$  and  $X$  are independent  $N \times p$  complex random matrices both with i.i.d. entries. The  $W_{ij}$ 's are  $P$  distributed and the  $X_{ij}$ 's are complex standard Gaussian random variables. In the above context  $a$  is real number independent of  $N$ . The proof of [17] and [4] relies on three main steps:

- conditionally on  $H = \frac{1}{N}WW^*$ , the eigenvalue process induced by  $\tilde{M}_N$  defines a so-called determinantal random point field;
- the corresponding correlation kernel can be expressed as a double integral in the complex plane depending on  $H$  through its sole spectral measure  $\mu_N$ ;
- under suitable assumptions on  $W$  and thanks to concentration results for the spectral measure  $\mu_N$  established by [2,3] and [16], the asymptotic analysis of the correlation kernel (and local statistics) can be performed.

One may also point out that the correlation kernel is expressed in terms of Bessel functions with a large parameter and whose order may be bounded or not. The asymptotic analysis requires uniform asymptotic expansions of such Bessel functions. In [4], the authors consider the case where  $\gamma = 1$  and  $p - N = O(N^{43/48})$  only, for which uniform asymptotic expansions of Bessel functions are easier to handle. In this paper, we are able to consider the general case with  $\gamma \geq 1$ , using different uniform asymptotic expansions of Bessel functions.

The parameter  $a$  is to be seen as the “order of the Gaussian regularization” of  $P$ . In principle the above result would yield a full proof of the universality conjecture if one could let  $a$  approach (and be smaller than)  $1/\sqrt{N}$ . Unfortunately this idea fails whatever sharp concentration results can be established for  $\mu_N$ . The asymptotic analysis is not stable in this scale.

A breakthrough to overcome this difficulty is obtained in [14]. In [11–13] concentration results are deeply sharpened so as to be able to consider a Gaussian regularization of order  $a \gg \frac{1}{\sqrt{N}}$ . Given an arbitrary (non-Gauss divisible) distribution  $F$ , the main point is however to show that one can find a Gauss divisible distribution approximating  $F$  sufficiently well so that one can deduce universality in the bulk for a sample covariance matrix with i.i.d.  $F$  distributed entries. This is the main step achieved in [14], which we now briefly expose.

Consider the Ornstein–Uhlenbeck (OU) process

$$L = \frac{1}{4} \frac{\partial^2}{\partial x^2} - \frac{x}{2} \frac{\partial}{\partial x}, \quad \partial_t u = Lu.$$

Given a complex probability distribution  $\tilde{\mu}$ , let then  $X_t$  be the  $N \times p$  matrix process with initial distribution  $\tilde{\mu}^{\otimes Np}$  and whose entries (real and imaginary parts) evolve independently according to the OU process. Then  $X_t$  evolves as

$$t \mapsto e^{-t/2}(\hat{H} + (e^t - 1)^{1/2}X),$$

where  $\hat{H}$  has i.i.d. entries with distribution  $\tilde{\mu}$  and  $X$  has standard complex Gaussian entries. Let  $e^{t\mathcal{L}} := (e^{tL})^{\otimes Np}$  denote the dynamics of the OU process for all the matrix elements. Here, for small  $t$ , one should think that  $t \simeq a^2$ .

Erdős et al. [14] prove that there exists a Gauss divisible distribution  $\mu_G^t$  approximating  $F$  in the total variation norm in a sufficiently good way. Roughly speaking, let

$$\mu_G^t = e^{tL}(1 - tL + tL^2/2)F^c, \quad \tilde{\mu} = (1 - tL + t^2L^2/2)F^c,$$

where  $F^c$  is obtained from  $F$  after truncation and centralization. Intuitively, it is reasonable to expect that  $\mu_G^t$  is a good approximation of  $F$  in the scale  $t \ll 1$ . For sake of completeness we here recall their result (Proposition 2.1).

Let  $\theta$  be a smooth cutoff function satisfying  $\theta(x) = 1$  if  $|x| \leq 1$  and  $\theta(x) = 0$  for  $|x| \geq 2$ .

**Proposition 1.1.** *Let  $V$  satisfy (1.4), (1.5) for some  $k$  and (1.6). Let  $\lambda > 0$  be sufficiently small and  $t = N^{\lambda-1}$ . Let  $c_N, d_N$  be real numbers so that  $v d\mu$  defines a centered probability density if the function  $v$  is given by:*

$$v(x) := e^{-V^c(x)}, \quad V^c(x) := V(x)\theta((x - c_N)N^{-\lambda/4k}) + d_N.$$

Let then  $\mathcal{L} := L^{\otimes Np}$ ,  $f_{N,p} := (e^{-V})^{\otimes Np}$ , and  $f_{c,N,p} := v^{\otimes Np}$ .

1. There exist constants  $C > 0, c > 0$  depending on  $k$  and  $\lambda$  such that  $\int |f_{c,N,p} - f_{N,p}| d\mu^{\otimes Np} \leq C e^{-c(Np)^{c/2}}$ .
2.  $g_t := (1 - tL + t^2L^2/2)v$  is a probability measure with respect to  $d\mu$ . Setting  $G_t := [g_t]^{\otimes Np}$ , there exists a constant  $C$  depending on  $\lambda$  and the constants  $C_*$  and  $\delta_1$  defined in (1.4) and (1.5) such that

$$\int \frac{|e^{t\mathcal{L}}G_t - f_{c,N,p}|^2}{e^{t\mathcal{L}}G_t} d\mu^{\otimes Np} \leq CNp t^{6-\lambda} \leq CN^{-4+8\lambda} \frac{p}{N}.$$

The idea is then to prove Theorem 1.1 and Theorem 1.2 for the Gauss divisible ensembles with small parameter  $a \sim N^{(\lambda-1)/2}$  for some  $\lambda > 0$ . Then, using Proposition 1.1, and following the idea of [14], Section 4, one can extend universality of local eigenvalue statistics  $S_N^{(m)}(f, u, \rho_N)$  with  $m = 2, 3$  and that of the spacing function to sample covariance matrices satisfying (1.7).

The paper is organized as follows. In Section 2 we study eigenvalue statistics in the bulk of the spectrum for Gauss divisible sample covariance matrices. To this aim, we first recall some properties of the *Deformed Wishart ensemble*. This is the conditional distribution of  $\tilde{M}_N$  knowing  $W$ . Such an ensemble is in particular known to be determinantal, as we recall. We then establish some improved convergence rates for the spectral measure of sample covariance matrices whose entries have a sub-Gaussian tail. These concentration results then allow to compute the asymptotic correlation functions as  $N \rightarrow \infty$  in the regime where  $a \rightarrow 0, a \gg \frac{1}{\sqrt{N}}$ . We then prove Theorem 1.2 and Theorem 1.1 in Section 3.

In the whole paper, we use  $C$  and  $c$  to denote constants whose value may vary from line to line.

## 2. The Gauss divisible ensemble

In this section we establish some universality results for the following Gauss divisible ensemble: let

$$\tilde{M}_N = \frac{1}{N}(W + aX)(W + aX)^*, \tag{2.1}$$

be a Hermitian  $N \times N$  random matrix where

- $(H_0)$   $X$  and  $W$  are independent  $N \times p$  random matrices where  $p \geq N$  and there exists  $\gamma \geq 1$  so that  $\lim_{N \rightarrow \infty} p/N = \gamma$ ;
- $(H_1)$   $a = a_N = \sqrt{\frac{N^\lambda}{N}}$  where  $\lambda > 0$  is a (small) real number;
- $(H_2)$   $X$  is a  $N \times p$  random matrix with complex standard Gaussian entries:

$$\operatorname{Re}X_{ij}, \operatorname{Im}X_{ij} \sim \mathcal{N}(0, 1/2) \quad \forall 1 \leq i \leq N, 1 \leq j \leq p;$$

- $(H_3)$   $W = (W_{ij})$ ,  $1 \leq i \leq N$ ,  $1 \leq j \leq p$ , is a complex random matrix such that  $\operatorname{Re} W_{ij}$ ,  $\operatorname{Im} W_{ij}$ ,  $1 \leq i \leq N$ ,  $1 \leq j \leq p$ , are i.i.d. and satisfy:

$$(A_1) \quad \text{There exists a constant } \delta_o > 0 \text{ such that } \mathbb{E}e^{\delta_o |W_{ij}|^2} < \infty, \forall i, j.$$

Without loss of generality, we also make the assumption that

$$\mathbb{E}|W_{ij}|^2 = 1/4. \quad (2.2)$$

Note that (2.2) can always be achieved by rescaling  $\tilde{M}_N$ . From now on, we denote by  $y_1 \geq y_2 \geq \dots \geq y_N$  the ordered eigenvalues of  $H = H_N = WW^*/N$  and let  $\mu_N = \frac{1}{N} \sum_{i=1}^N \delta_{y_i}$  be its spectral measure. Under assumption (2.2), it is known that  $\mu_N$  almost surely converges to the Marcenko–Pastur distribution with parameters  $\gamma$  and  $1/4$ ,  $\rho_{\gamma, 1/4}^{\text{MP}} = \rho_{\gamma}^{\text{MP}}$ , whose density function is given by (1.1) with  $\sigma^2 = 1/4$ . When  $\gamma = 1$ , we simply denote  $\rho_1^{\text{MP}}$  by  $\rho^{\text{MP}}$  for short.

### 2.1. Correlation functions

The sample covariance Gauss divisible ensemble has a nice mathematical structure that we are going to make use of in the sequel: the conditional distribution of  $\tilde{M}_N$  with respect to  $W$  is the so-called complex deformed Wishart ensemble. Such an ensemble has been widely studied in random matrix theory as it induces a determinantal random point field. We recall some results used in [4] (see references therein) that will be needed for the sequel. Let then  $P_N^H(\lambda_1, \lambda_2, \dots, \lambda_N)$  be the joint eigenvalue distribution induced by the conditional distribution of  $\tilde{M}_N$  w.r.t.  $H$ . Then (see Section 3 in [4]),  $P_N^H$  is absolutely continuous with respect to Lebesgue measure on  $\mathbb{R}_+^N$ . Its density  $f_N^H$  is given by:

$$f_N^H(x_1, x_2, \dots, x_N) = \frac{V(x)}{V(y)} \det \left( \frac{1}{S} e^{-(y_i + x_j)/S} I_\nu \left( \frac{2\sqrt{y_i x_j}}{S} \right) \left( \frac{x_j}{y_i} \right)^{v/2} \right)_{i,j=1}^N, \quad (2.3)$$

where  $V(x) := \prod_{1 \leq i < j \leq N} (x_i - x_j)$ ,  $v = p - N$ ,  $I_\nu$  is the modified Bessel function of the first kind, and  $S = a^2/N$ . The main tools to study local eigenvalue statistics are the so-called eigenvalue correlation functions (see Section 3). They are defined for any integer  $1 \leq m \leq N$  by  $R_N^{(m)}(x_1, x_2, \dots, x_m; H) = \frac{N!}{(N-m)!} \int_{\mathbb{R}_+^{N-m}} f_N^H(x_1, x_2, \dots, x_m, \lambda_{m+1}, \dots, \lambda_N) \prod_{i=m+1}^N d\lambda_i$ . Then, for any integer  $1 \leq m \leq N$ , one has that

$$R_N^{(m)}(x_1, x_2, \dots, x_m; H) = \det(K_N(x_i, x_j; H))_{i,j=1}^m,$$

for some correlation kernel  $K_N(u, v; H)$ . This gives the determinantal random point field structure. Furthermore  $K_N(u, v; H)$  is given by

$$\begin{aligned} K_N(u, v; H) &:= \frac{1}{2i^2 \pi^2 S^2} \int_{i\mathbb{R}+A} dw \int_{\Gamma} dz e^{(-z^2 + w^2)/S} K_\nu \left( \frac{2w\sqrt{v}}{S} \right) I_\nu \left( \frac{2z\sqrt{u}}{S} \right) \\ &\quad \times \prod_{i=1}^N \frac{w^2 - y_i}{z^2 - y_i} \left( \frac{w}{z} \right)^v \left( \frac{w+z}{w-z} - \frac{w-z}{w+z} \right), \end{aligned} \quad (2.4)$$

where  $K_\nu$  is the modified Bessel function of the second kind.  $\Gamma$  is a contour oriented counterclockwise enclosing the  $\pm y_i^{1/2}$ ,  $i = 1, \dots, N$ , and  $A$  is large enough so that the two contours  $\Gamma$  and  $\mathcal{T} = i\mathbb{R} + A$  do not cross each other (see Fig. 1).

Due to the determinantal structure, correlation functions are not impacted by conjugation of the correlation kernel. In particular, for any  $b \in \mathbb{R}$  and any integer  $m \geq 1$ ,

$$\det(K_N(u_i, u_j; H))_{i,j=1}^m = \det(K_N(u_i, u_j; H) e^{2b(\sqrt{u_j} - \sqrt{u_i})/S})_{i,j=1}^m.$$

In the sequel we consider the conjugation for some  $b$  that will be defined in the asymptotic analysis. We set

$$K_N^b(u_i, u_j; H) := K_N(u_i, u_j; H) e^{2b(\sqrt{u_j} - \sqrt{u_i})/S}. \quad (2.5)$$

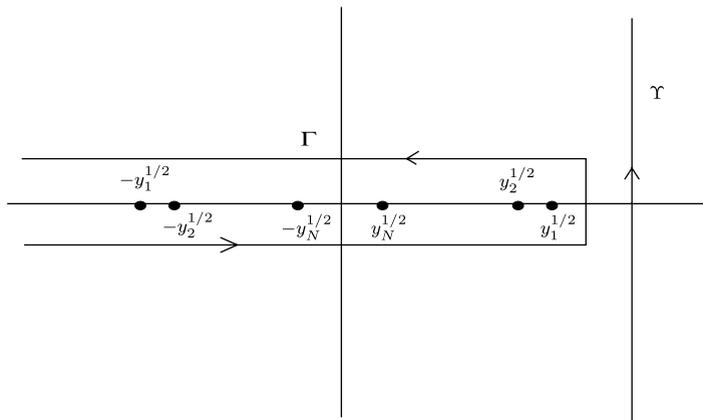


Fig. 1. The two contours defining the correlation kernel  $K_N(u, v; H)$ .

The correlation kernel  $K_N^b(u, v; H)$  only depends on  $H$  through its spectral measure  $\mu_N$ . As in the case where  $a$  is fixed (independent of  $N$ ), the idea is to use the convergence of  $\mu_N$  to the Marčenko–Pastur distribution. In particular, one would like to make the replacement (outside a suitably negligible set)  $\prod_{i=1}^N (w^2 - y_i) = \exp\{N(1 + o(1)) \int \ln(w^2 - y) d\rho_\gamma^{\text{MP}}(y)\}$ . In Section 2.2, we prove that this replacement can be made in some sense provided  $\text{Im}(w^2)$  is not too small.

Also, the kernel  $K_N^b(u, v; H)$  depends on Bessel functions with large argument and whose order can be bounded or grow unboundedly. For some technical reasons (namely uniform asymptotic expansions of Bessel functions with bounded/unbounded order), we will separate the asymptotic analysis of the kernel depending on the case where  $v$  is bounded (Section 2.3) or not (Section 2.4). This has no impact (up to technicalities) on the strategy of the proof.

## 2.2. Concentration results

This subsection is devoted to the proof of the following Proposition 2.1, which precises the rate of convergence of  $\mu_N$  to  $\rho^{\text{MP}}$ . Before exposing this proposition we need a few definitions and notations. Given a complex number  $z = u + i\eta$ ,  $u \in \mathbb{R}$ ,  $\eta > 0$ , we define the Stieltjes transform of  $\mu_N$  by

$$m_N(z) := \int \frac{1}{\lambda - z} d\mu_N(\lambda), \quad (2.6)$$

and the Stieltjes transform of the limiting Marčenko–Pastur distribution by

$$m_{\text{MP}}(z) := \int \frac{1}{\lambda - z} d\rho_\gamma^{\text{MP}}(\lambda). \quad (2.7)$$

Let  $\epsilon > 0$  be given (small enough).

**Proposition 2.1.** *Let  $z = u + i\eta$  for some  $u \in [u_- + \epsilon, u_+ - \epsilon]$  and  $\eta > 0$ . Then, there exist a constant  $c_1$  and  $C_o, C > 0, c > 0$  depending on  $\epsilon$  only such that  $\forall \delta < c_1\epsilon$ ,*

$$\mathbb{P}\left(\sup_{u \in [u_- + \epsilon, u_+ - \epsilon]} |m_N(z) - m_{\text{MP}}(z)| \geq \left(\delta + C_o \left| \frac{p}{N} - \gamma \right| \right)\right) \leq C e^{-c\delta\sqrt{N\eta}},$$

for any  $(\ln N)^4/N \leq \eta \leq 1$ . Furthermore, given  $\eta \geq (\ln N)^4/N$ , there exist constants  $c > 0, C > 0$  and  $K_o$  such that  $\forall \kappa \geq K_o$ ,

$$\mathbb{P}\left(\sup_{|x| > (\epsilon/200)^2, y \geq \eta} |m_N(x + iy)| \geq \kappa\right) \leq C e^{-c\sqrt{\kappa N\eta}}.$$

**Proof.** To ease the reading, the proof is postponed to Appendix A. The proof is actually a modification of that of Theorems 3.1 and 4.1 in [13], where the Stieltjes transform of Hermitian Wigner matrices is considered. We thus simply indicate the main changes.  $\square$

### 2.3. Asymptotics of the correlation kernel when $\nu = p - N$ is bounded

The aim of this subsection is here to prove the following proposition.

Let  $\epsilon > 0$  be given (small) independent of  $N$ .

**Proposition 2.2.** *Let  $u_* \in [u_- + \epsilon, u_+ - \epsilon] = [\epsilon, 1 - \epsilon]$  and  $m \geq 1$  be given. Consider a sequence  $u = u_N$  such that  $N^{1-\lambda}|u - u_*| = o(1)$ . Then, there exist  $b \in \mathbb{R}$ , a set  $\Omega_N$ , and positive constants  $C, c$  such that:*

- the complement of  $\Omega_N$  is negligible:  $\mathbb{P}(\Omega_N^c) \leq C e^{-cN^{\lambda/4}}$ ,
- on  $\Omega_N$  one has that

$$\begin{aligned} & \lim_{N \rightarrow \infty} \det \left( \frac{1}{N \rho^{\text{MP}}(u)} K_N^b \left( u + \frac{x_i}{N \rho^{\text{MP}}(u)}, u + \frac{x_j}{N \rho^{\text{MP}}(u)}; H \right) \right)_{i,j=1}^m \\ &= \det \left( \frac{\sin(\pi(x_i - x_j))}{\pi(x_i - x_j)} \right)_{i,j=1}^m. \end{aligned}$$

**Remark 2.1.** *In Remark 2.3 we explain how to extend Proposition 2.2 to the case where  $|u - u_*| = o(N^{-c})$  with  $c > 0$  arbitrarily small. This extension is needed for the proof of Theorem 1.2.*

This subsection is devoted to the proof of Proposition 2.2. The proof is divided into two parts: first we obtain a new expression for the correlation kernel, which then allows to derive its asymptotics by a saddle point argument.

#### 2.3.1. Rewriting the kernel

Our strategy consists into two parts: we first replace the Bessel functions with their asymptotic expansion and then remove the singularity  $1/(w - z)$  in the correlation kernel as it will be proved to prevent a direct saddle point analysis. As this part is highly technical and needs a few notations, we summarize our main result in Lemma 2.3 stated at the end of this subsection. The reader might skip this part and is simply referred to the above cited lemma.

Our first task to replace in (2.4) Bessel functions with their asymptotic expansion given in Appendix B, formula (B.1). This replacement can be made, up to a negligible error, provided the contour  $\Gamma$  is cut in a small neighborhood of the imaginary axis: see Fig. 2. In the sequel (see Fig. 2) we call  $x_o^\pm$  (resp.  $x_1^\pm$ ) the endpoints of  $\Gamma_c$  with  $\text{Re}(x_o) < 0$  (resp.  $\text{Re}(x_1) > 0$ ) with positive/negative imaginary part surrounding the imaginary axis:  $x_o^\pm$  and  $x_1^\pm$  will be chosen

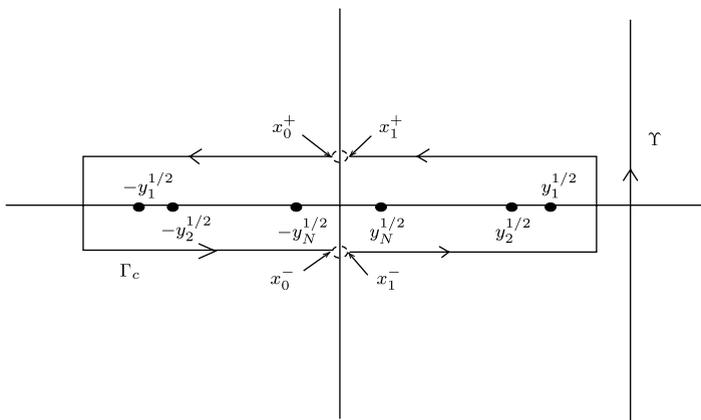


Fig. 2. The contour  $\Gamma_c$  has been cut along the imaginary axis.

in Section 2.3.2. We also call  $\Gamma_1$  (resp.  $\Gamma_2$ ) the part of  $\Gamma_c$  lying to the right (resp. left) of  $i\mathbb{R}$ . Then  $\Gamma_1$  is the image of  $\Gamma_2$  by  $z \mapsto -z$  and  $\Gamma_c = \Gamma_1 \cup \Gamma_2$ . One obtains the following lemma.

**Lemma 2.1.** *For any  $b \in \mathbb{R}$ , one has that*

$$K_N^b(u, v; H) = (1 + o(1))(K_N^{1,b}(u, v; H) + K_N^{2,b}(u, v; H)), \quad (2.8)$$

where

$$\begin{aligned} K_N^{1,b}(u, v; H) &= \frac{e^{2b(\sqrt{v}-\sqrt{u})/S}}{4i^2\pi^2S} \int_{i\mathbb{R}+A} dw \int_{\Gamma_1} dz e^{(-z^2+w^2)/S} e^{-2w\sqrt{v}/S} e^{2z\sqrt{u}/S} \\ &\quad \times \frac{1}{\sqrt{w}\sqrt{z}(uv)^{1/4}} \prod_{i=1}^N \frac{w^2 - y_i}{z^2 - y_i} \left(\frac{w}{z}\right)^v \frac{w+z}{w-z}, \\ K_N^{2,b}(u, v; H) &= \frac{e^{2b(\sqrt{v}-\sqrt{u})/S}}{4i^2\pi^2S} \int_{i\mathbb{R}+A} dw \int_{\Gamma_1} dz e^{(-z^2+w^2)/S} e^{-2w\sqrt{v}/S} e^{-2z\sqrt{u}/S} \\ &\quad \times \frac{1}{\sqrt{w}\sqrt{z}(uv)^{1/4}} \prod_{i=1}^N \frac{w^2 - y_i}{z^2 - y_i} \left(\frac{w}{z}\right)^v \frac{w+z}{w-z} e^{vi\pi+i\pi/2}. \end{aligned} \quad (2.9)$$

**Proof.** Here we will call on some arguments already used in [4], which we won't develop entirely, yet trying to be the most self-contained as possible. Lemma 2.1 essentially follows from the two claims exposed in the sequel and whose proof relies on the saddle point analysis in the next subsection.

**Claim 2.1.** *One has that*

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{e^{2b(\sqrt{v}-\sqrt{u})/S}}{2i^2\pi^2S^2} \int_{i\mathbb{R}+A} dw \int_{\Gamma \setminus \Gamma_c} dz e^{(-z^2+w^2)/S} K_\nu\left(\frac{2w\sqrt{v}}{S}\right) I_\nu\left(\frac{2z\sqrt{u}}{S}\right), \\ \prod_{i=1}^N \frac{w^2 - y_i}{z^2 - y_i} \left(\frac{w}{z}\right)^v \left(\frac{w+z}{w-z} - \frac{w-z}{w+z}\right) = 0. \end{aligned} \quad (2.10)$$

Claim 2.1 essentially follows from the fact that  $|I_\nu(z)| \leq e^z$  if  $\text{Re}(z) > 0$  (and a similar bound for  $K_\nu$ ) and the fact that  $\Gamma \setminus \Gamma_c$  lies far away from the critical points (a full justification can be derived from the saddle point analysis performed in Section 2.3.2, see Remark 2.2).

**Claim 2.2.** *One has that*

$$\begin{aligned} \frac{e^{2b(\sqrt{v}-\sqrt{u})/S}}{2i^2\pi^2S^2} \int_{i\mathbb{R}+A} dw \int_{\Gamma_c} dz e^{(-z^2+w^2)/S} K_\nu\left(\frac{2w\sqrt{v}}{S}\right) I_\nu\left(\frac{2z\sqrt{u}}{S}\right) \\ \times \prod_{i=1}^N \frac{w^2 - y_i}{z^2 - y_i} \left(\frac{w}{z}\right)^v \left(\frac{w+z}{w-z} - \frac{w-z}{w+z}\right) \\ = \frac{e^{2b(\sqrt{v}-\sqrt{u})/S}}{4i^2\pi^2S} \int_{i\mathbb{R}+A} dw \int_{\Gamma_1} dz e^{(-z^2+w^2)/S} e^{-2w\sqrt{v}/S} e^{2z\sqrt{u}/S} \frac{1}{(uv)^{1/4}\sqrt{w}\sqrt{z}} \\ \times \prod_{i=1}^N \frac{w^2 - y_i}{z^2 - y_i} \left(\frac{w}{z}\right)^v \left(\frac{w+z}{w-z} - \frac{w-z}{w+z}\right) (1 + O(N^{-\lambda/2})). \end{aligned} \quad (2.11)$$

By a straightforward change of variables, we can reduce  $\Gamma_c$  to  $\Gamma_1$ . Claim 2.2 then follows by uniform asymptotic expansions of first the  $w$ -integral, using the asymptotics of the Bessel function  $K_\nu$  given in (B.1), and then of the  $z$ -integral. Again this follows from the saddle point analysis of Section 2.3.2. This saddle point analysis will prove that the correlation kernel does not vanish in the large  $N$ -limit. Then splitting  $\Gamma_1$  into sufficiently many pieces lying to the left or to the right of the critical points and moving accordingly  $\Upsilon = i\mathbb{R} + A$  to pass close enough to the same critical points yields that the error term is of order  $N^{-\lambda/2}$ . This yields Claim 2.2 and Lemma 2.1. We skip the detail.  $\square$

A saddle point analysis of the correlation kernel  $K_N^{1,b}$  cannot be performed at that point, due to the singularity  $1/(w-z)$ . Indeed, assuming that  $\prod_{i=1}^N (w^2 - y_i) = e^{\int N \ln(w^2 - y) d\rho^{\text{MP}}(y)(1+o(1))}$ , it is not difficult to see that the  $z$ - and  $w$ -integrands have the same critical points. Thus we first remove this singularity (see [4] for a more detailed explanation).

Define

$$\begin{aligned} f_u^N(z) &:= z^2 - 2\sqrt{u}z + S \sum_{i=1}^N \ln(z^2 - y_i), & f_N(z) &= f_u^N(z) + 2b\sqrt{u}, \\ f_v^N(z) &:= z^2 - 2\sqrt{v}z + S \sum_{i=1}^N \ln(z^2 - y_i), & \tilde{f}_N(z) &= f_v^N(z) + 2b\sqrt{v}, \\ g_{N,u}(w, z) &:= (w-b) \frac{(f_u^N)'(w) - (f_u^N)'(z)}{w-z} + (f_u^N)'(z), \\ g_{N,u}^1(w, z) &= \frac{vbS}{wz} + \frac{Sb}{2wz} \frac{w-z}{w+z}, & \theta(w, b) &= -\frac{e^{2(w-b)(\sqrt{u}-\sqrt{v})/S} - 1}{2(w-b)(\sqrt{u}-\sqrt{v})}, \\ \theta^1(z, b) &= \frac{e^{-2(z-b)(\sqrt{u}+\sqrt{v})/S} - 1}{2(z-b)(\sqrt{u}+\sqrt{v})}. \end{aligned} \tag{2.12}$$

**Lemma 2.2.** *One has that*

$$\begin{aligned} &K_N^{1,b}(u, v; H) \\ &= \frac{1}{(uv)^{1/4} 4i^2 \pi^2 S} \int_{i\mathbb{R}+A} dw \int_{\Gamma_1} dz e^{(f_N(w) - f_N(z))/S} \\ &\quad \times \frac{w+z}{\sqrt{w}\sqrt{z}} (g_{N,u}(w, z) + g_{N,u}^1(w, z)) \left(\frac{w}{z}\right)^v \theta(w, b) \end{aligned} \tag{2.13}$$

$$\begin{aligned} &- \frac{(x_1^- - b)}{4i^2 \pi^2 (uv)^{1/4}} \int_{i\mathbb{R}+A} dw \frac{w+x_1^-}{\sqrt{w}\sqrt{x_1^-}} \left(\frac{w}{x_1^-}\right)^v e^{(f_N(w) - f_N(x_1^-))/S} \theta(w, b) \\ &+ \frac{(x_1^+ - b)}{4i^2 \pi^2 (uv)^{1/4}} \int_{i\mathbb{R}+A} dw \frac{w+x_1^+}{\sqrt{w}\sqrt{x_1^+}} \left(\frac{w}{x_1^+}\right)^v e^{(f_N(w) - f_N(x_1^+))/S} \theta(w, b). \end{aligned} \tag{2.14}$$

Setting  $\tilde{K}_N^{2,b}(u, v; H) := e^{-vi\pi - i\pi/2} K_N^{2,b}(u, v; H)$ , we have that

$$\begin{aligned} &\tilde{K}_N^{2,b}(u, v; H) \\ &= \frac{e^{-4b\sqrt{u}/S}}{4i^2 \pi^2 S} \int_{i\mathbb{R}+A} dw \int_{\Gamma_1} dz (g_{N,v}(w, z) + g_{N,v}^1(w, z)) e^{(\tilde{f}_N(w) - \tilde{f}_N(z))/S} \\ &\quad \times \frac{w+z}{\sqrt{w}\sqrt{z}(uv)^{1/4}} \left(\frac{w}{z}\right)^v \theta^1(z, b) \end{aligned}$$

$$\begin{aligned}
& - \frac{e^{-(4b\sqrt{v})/S}(x_1^- - b)}{4i^2\pi^2(uv)^{1/4}} \int_{i\mathbb{R}+A} dw \frac{w+x_1^-}{\sqrt{w}\sqrt{x_1^-}} \left(\frac{w}{x_1^-}\right)^v e^{(\tilde{f}_N(w)-\tilde{f}_N(x_1^-))/S} \theta^1(x_1^-, b) \\
& + \frac{e^{-4b\sqrt{u}/S}(x_1^+ - b)}{4i^2\pi^2(uv)^{1/4}} \int_{i\mathbb{R}+A} dw \frac{w+x_1^+}{\sqrt{w}\sqrt{x_1^+}} e^{(\tilde{f}_N(w)-\tilde{f}_N(x_1^+))/S} \left(\frac{w}{x_1^+}\right)^v \theta^1(x_1^+, b).
\end{aligned} \tag{2.15}$$

**Proof.** We only consider  $K_N^{1,b}$  (the arguments are similar for  $K_N^{2,b}$ ). To this aim we make the change of variables  $w = b + \beta(w' - b)$ ,  $z = b + \beta(z' - b)$  for some  $\beta$  real close to 1 and get the following: set

$$E(w) := e^{(w^2 - 2\sqrt{v}w)/S} \prod_{i=1}^N (w^2 - y_i); \quad G(z) = e^{(z^2 - 2\sqrt{u}z)/S} \prod_{i=1}^N (z^2 - y_i). \tag{2.16}$$

Then one has that

$$\begin{aligned}
K_N^{1,b}(u, v; H) &= \frac{e^{2b(\sqrt{v}-\sqrt{u})/S}}{(uv)^{1/4}4i^2\pi^2S} \int_{i\mathbb{R}+A} dw \int_{\Gamma_1} dz \frac{1}{\sqrt{wz}} \frac{E(w)}{G(z)} \left(\frac{w}{z}\right)^v \frac{w+z}{w-z} \\
&= \frac{e^{2b(\sqrt{v}-\sqrt{u})/S}}{4i^2\pi^2S(uv)^{1/4}} \int_{i\mathbb{R}+A} \int_{\Gamma_1} \frac{\beta dw dz}{\sqrt{(b+\beta(w'-b))(b+\beta(z'-b))}} \\
&\quad \times \left(\frac{b+\beta(w'-b)}{b+\beta(z'-b)}\right)^v \frac{2b+\beta(w'+z'-2b)}{w'-z'} \frac{E(b+\beta(w'-b))}{G(b+\beta(z'-b))} \\
&\quad - \frac{e^{2b(\sqrt{v}-\sqrt{u})/S}}{4i^2\pi^2S(uv)^{1/4}} \int_{i\mathbb{R}+A} \int_{x_1^+}^{b+\beta(x_1^+-b)} \frac{dw dz}{\sqrt{wz}} \left(\frac{w}{z}\right)^v \frac{w+z}{w-z} \frac{E(w)}{G(z)} \\
&\quad + \frac{e^{2b(\sqrt{v}-\sqrt{u})/S}}{4i^2\pi^2S(uv)^{1/4}} \int_{i\mathbb{R}+A} \int_{x_1^-}^{b+\beta(x_1^--b)} \frac{dw dz}{\sqrt{wz}} \left(\frac{w}{z}\right)^v \frac{w+z}{w-z} \frac{E(w)}{G(z)}.
\end{aligned} \tag{2.17}$$

Differentiating with respect to  $\beta$  (close to 1), we find that (see [17] Section 2 for the detail)

$$\begin{aligned}
& (uv)^{1/4} K_N^{1,b} + (\sqrt{v} - \sqrt{u}) \frac{d}{d\sqrt{v}} ((uv)^{1/4} K_N^{1,b}) \\
&= \frac{e^{2b(\sqrt{v}-\sqrt{u})/S}}{4i^2\pi^2S^2} \int_{i\mathbb{R}+A} dw \int_{\Gamma_1} dz \frac{1}{\sqrt{wz}} \left(\frac{w}{z}\right)^v \frac{w+z}{w-z} \frac{E(w)}{G(z)} \\
&\quad \times \left( \left[ -2(w+z-b) + 2\sqrt{u} + S \sum_{i=1}^N \frac{2y_i(w+z) - b(2wz+y_i)}{(w^2-y_i)(z^2-y_i)} \right] \right. \\
&\quad \left. - \frac{vbS}{wz} - \frac{Sb}{2wz} \frac{w-z}{w+z} \right) \\
&\quad - \frac{(x_1^+ - b)e^{2b(\sqrt{v}-\sqrt{u})/S}}{4i^2\pi^2S} \int_{i\mathbb{R}+A} dw \frac{1}{\sqrt{wx_1^+}} \left(\frac{w}{x_1^+}\right)^v \frac{w+x_1^+}{w-x_1^+} \frac{E(w)}{G(x_1^+)} \\
&\quad + \frac{(x_1^- - b)e^{2b(\sqrt{v}-\sqrt{u})/S}}{4i^2\pi^2S} \int_{i\mathbb{R}+A} dw \frac{1}{\sqrt{wx_1^-}} \left(\frac{w}{x_1^-}\right)^v \frac{w+x_1^-}{w-x_1^-} \frac{E(w)}{G(x_1^-)}.
\end{aligned} \tag{2.18}$$

$$\begin{aligned}
& - \frac{(x_1^+ - b)e^{2b(\sqrt{v}-\sqrt{u})/S}}{4i^2\pi^2S} \int_{i\mathbb{R}+A} dw \frac{1}{\sqrt{wx_1^+}} \left(\frac{w}{x_1^+}\right)^v \frac{w+x_1^+}{w-x_1^+} \frac{E(w)}{G(x_1^+)} \\
& + \frac{(x_1^- - b)e^{2b(\sqrt{v}-\sqrt{u})/S}}{4i^2\pi^2S} \int_{i\mathbb{R}+A} dw \frac{1}{\sqrt{wx_1^-}} \left(\frac{w}{x_1^-}\right)^v \frac{w+x_1^-}{w-x_1^-} \frac{E(w)}{G(x_1^-)}.
\end{aligned} \tag{2.19}$$

By the definition of  $f_u^N$ , the term in the bracket (2.18) can also be written

$$\begin{aligned}
(2.18) &= -\frac{(w-b)(f_u^N)'(w) - (z-b)(f_u^N)'(z)}{w-z} \\
&= -(w-b)\frac{(f_u^N)'(w) - (f_u^N)'(z)}{w-z} - (f_u^N)'(z).
\end{aligned} \tag{2.20}$$

Integrating by parts yields Lemma 2.2.  $\square$

To summarize this highly technical subsection, we have proved that

**Lemma 2.3.** *Assume that  $v$  is a bounded integer. Then*

$$K_N^b(u, v; H) = (1 + o(1))(K_N^{1,b}(u, v; H) + K_N^{2,b}(u, v; H)),$$

where the kernels  $K_N^{1,b}$  and  $K_N^{2,b}$  are defined in Lemma 2.2.

### 2.3.2. Saddle point analysis of the correlation kernel

We are now in position to perform the saddle point analysis of the correlation kernel  $K_N^b(u, v; H)$  and prove Proposition 2.2. Let  $\tau$  be a given real number independent of  $N$  and assume that  $v - u = \tau/(N\rho^{\text{MP}}(u))$ . We mainly focus on the asymptotics of the first term (2.13) in  $K_N^{1,b}$ . The asymptotic expansion of the two other terms (2.14) in  $K_N^{1,b}$  and of  $K_N^{2,b}$  will be an easy corollary of the arguments used hereafter.

We first examine the saddle point analysis for the approximate exponential term:

$$f(w) = f_u(w) := w^2 - 2(w-b)\sqrt{u} + a^2 \int \ln(w^2 - y) d\rho^{\text{MP}}(y). \tag{2.21}$$

Note that  $f$  is the almost sure limit of the exponential term arising in the definition of (2.13) and should thus (approximately) lead its asymptotic analysis. Then using the fact that, when  $\gamma = 1$ , a Marčenko–Pastur random variable has the same law as a squared semi-circle random variable, one gets that

$$f'(w) := 2w - 2\sqrt{u} + a^2 \int \frac{2w}{w^2 - y} d\rho^{\text{MP}}(y) = 2w - 2\sqrt{u} + 4a^2(w - \sqrt{w^2 - 1}). \tag{2.22}$$

Given a point  $u \in [\epsilon, 1 - \epsilon]$ ,  $\epsilon > 0$ , it is easy to see that  $f$  admits two conjugate critical points:

$$w_c^\pm = \frac{(1 + 2a^2)\sqrt{u} \pm 2a^2 i \sqrt{1 + 4a^2 - u}}{1 + 4a^2} = \sqrt{u} \pm a^2 i \pi \sqrt{u} \rho^{\text{MP}}(u) + O(a^2).$$

More precisely, one has that  $\text{Im}(w_c^\pm) = \pm a^2 \pi \sqrt{u} \rho^{\text{MP}}(u) + O(a^4)$ , where the  $O$  is uniform in the bulk (depending on  $\epsilon$  only). One can also check that

$$f''(w_c^\pm) = 2(1 + 2a^2) + \frac{4a^2 w_c^\pm}{\sqrt{(w_c^\pm)^2 - 1}} = 2 + O(a^2),$$

where the  $O(\cdot)$  is uniform as long as  $u \in [\epsilon, 1 - \epsilon]$ ,  $\epsilon > 0$ , and depends on  $\epsilon$  only.

We now give the relevant contours for a saddle point analysis of the approximate exponential terms. Let  $\mathcal{Y} = \text{Re}(w_c^+) + it$ ,  $t \in \mathbb{R}$ , and  $\Gamma_1 = \Gamma_1^+ \cup \overline{\Gamma_1^+}$  with  $\Gamma_1^+ = \{w_c^+(t), t \in [\frac{\epsilon}{2}, 1 - \frac{\epsilon}{2}]\} \cup \{w_c^+(1 - \frac{\epsilon}{2}) + x, x > 0\} \cup \{i \text{Im}(w_c^+(\frac{\epsilon}{2})) + x, \epsilon/180 \leq x \leq \text{Re}(w_c^+(\frac{\epsilon}{2}))\}$ . Then it is easy to check that  $\text{Re} f$  achieves its maximum (resp. minimum) at  $w_c^\pm$  on  $\mathcal{Y}$  (resp.  $\Gamma$ ).

We now turn to the saddle point analysis for the true exponential term:

$$f_N(w) := w^2 - 2(w - b)\sqrt{u} + \frac{a^2}{N} \sum_{i=1}^N \ln(w^2 - y_i). \quad (2.23)$$

One has that  $f'_N(w) = 2(w - \sqrt{u}) - 2wa^2m_N(w^2)$ . Using the concentration results of Proposition 2.1, we now show that the first derivatives of  $f_N$  and  $f$  are close (on a suitable set). Let  $\zeta > 0$  be a very small number and set

$$S_{\epsilon, \zeta} = \left\{ z = x + iy, x \in \left[ u_- + \frac{\epsilon}{180}, u_+ - \frac{\epsilon}{8} \right], \zeta a^2 \leq y \leq 1 \right\}, \quad (2.24)$$

with  $u_- = 0$  and  $u_+ = 1$  here as  $\gamma = 1$ . Define for  $H = H_N = \frac{w w^*}{N}$ ,

$$\Omega_N = \left\{ \left| m_N(z) - m(z) \right| \leq N^{-\lambda/4}, \text{ if } \frac{\epsilon}{180} \leq \operatorname{Re}(z) \leq 1 - \frac{\epsilon}{8}, \zeta a^2 \epsilon \leq \operatorname{Im}(z) \leq 1; \right. \\ \left. \sup y_i(H) \leq K; \quad \sup_{|x| > (\epsilon/200)^2, y \geq \zeta a^2} |m_N(x + iy)| \leq K \right\}. \quad (2.25)$$

Using Proposition 2.1 and Lemma 7.3 in [12], we can deduce that there exist  $K$  large enough and constants  $c, C > 0$  such that

$$\mathbb{P}(\Omega_N^c) \leq C e^{-cN^{\lambda/4}}.$$

From now on we assume that  $H_N \in \Omega_N$ . As a consequence, there exists  $C > 0$  such that for any  $w \in S_{\epsilon, \zeta}$ ,

$$|f'_N(w) - f'(w)| \leq \frac{Ca^2}{N^{\lambda/4}}. \quad (2.26)$$

By Cauchy's formula, one can then deduce that for any integer  $l \geq 1$

$$|f_N^{(l)}(w) - f^{(l)}(w)| \leq \frac{C}{a^{2(l-2)} N^{\lambda/4}}. \quad (2.27)$$

Here the value of the constant  $C$  varies from line to line.

Let us now show that  $f_N$  admits two conjugate critical points  $w_{c,N}^\pm$  which are very close to  $w_c^\pm$ . The proof directly follows the arguments of [14], Section 3. The critical points of  $f_N$  are the solutions of the fixed point equation:

$$w = F_N(w) = \sqrt{u} - \frac{a^2}{N} \sum_{i=1}^N \frac{w}{w^2 - y_i}.$$

This equation clearly admits  $2N - 1$  real solutions which are interlaced with  $-\sqrt{y_N}, \dots, -\sqrt{y_1}, \sqrt{y_1}, \dots, \sqrt{y_N}$ . It admits also two non real solutions  $w_{c,N}^\pm$ . We now show that  $w_{c,N}^\pm$  are very close to  $w_c^\pm$ . Set  $F(w) = \sqrt{u} - a^2 \int \frac{w}{w^2 - y} d\rho^{\text{MP}}(y)$ . Define  $\Theta = \{z \in \mathbb{C}, |\operatorname{Re}(z) - \sqrt{u}| \leq C_0 a^2, \zeta a^2 \leq \operatorname{Im}(z) \leq C_0 a^2\}$ , for some large constant  $C_0$ . Then, on  $\Theta$ ,  $|F_N - F| \leq Ca^2 N^{-\lambda/4}$ . As on  $\Theta$ ,  $\operatorname{Re}(F(z)) = \sqrt{u} + O(a^2)$  and  $\operatorname{Im}(F(z)) = a^2 \pi \sqrt{u} \rho^{\text{MP}}(u) + o(a^2)$  we deduce that  $F_N(\Theta) \subset \Theta$ . Now it is an easy consequence of (2.27) that  $F_N$  is a contraction: one has that  $|F'_N(w) - F'(w)| \leq CN^{-\lambda/4}$ . Thus  $F_N$  admits a unique fixed point with  $\operatorname{Im}(w_{c,N}^+) > \zeta a^2$ . Furthermore, it is an easy fact that

$$|w_{c,N}^+ - w_c^+| \leq \frac{Ca^2}{N^{\lambda/4}}.$$

We now slightly modify the contours  $\Gamma$  and  $\Upsilon$  for the saddle point analysis of  $f_N$ . Set  $\Upsilon_N = \{\operatorname{Re}(w_{c,N}^+) + it, t \in \mathbb{R}\}$ . Then  $\Upsilon_N$  lies within a  $C^1$ -distance of at most  $Ca^2N^{-\lambda/4}$  from  $\Upsilon$ . Furthermore, setting  $w = \operatorname{Re}(w_{c,N}^+) + it$ ,

$$\operatorname{Re}\left(\frac{d}{dt} f_N(w)\right) = -t \left\{ 1 - \frac{a^2}{N} \sum_{i=1}^N \left( \frac{1}{|w - \sqrt{y_i}|^2} + \frac{1}{|w + \sqrt{y_i}|^2} \right) \right\}.$$

By the monotonicity of  $t \mapsto \sum_{i=1}^N \left( \frac{1}{|w - \sqrt{y_i}|^2} + \frac{1}{|w + \sqrt{y_i}|^2} \right)$ , we conclude that  $\operatorname{Re}(f_N)$  is maximum on  $\Upsilon_N$  at  $w_{c,N}^\pm$ . More precisely, using that  $\operatorname{Re} f_N''(w) \geq 1$  for any  $w \in \Upsilon \cap \{\operatorname{Im}(w) \geq \zeta a^2\}$ , one has that, if  $t$  stands for  $\operatorname{Im}(w)$ ,

$$\begin{aligned} \operatorname{Re}(f_N(w) - f_N(w_{c,N}^+)) &= -\operatorname{Re}\left(\int_0^{t - \operatorname{Im}(w_{c,N}^+)} s \, ds \int_0^1 f_N''(w_{c,N}^+ + ius) \, du\right) \\ &\leq -\frac{(t - \operatorname{Im}(w_{c,N}^+))^2}{4}. \end{aligned} \quad (2.28)$$

This implies in particular that

$$\operatorname{Re}(f_N(\operatorname{Re}(w_{c,N}^+) + i\zeta a^2) - f_N(w_{c,N}^+)) \leq -(\zeta - a^{-2}\operatorname{Im}(w_{c,N}^+))^2 a^4/4.$$

Let then  $w \in \Upsilon \cap \{z, 0 \leq \operatorname{Im}(z) \leq \zeta a^2\}$ . Then, with a slight abuse of notation when  $\operatorname{Im}(w) = 0$ ,

$$|\exp\{f_N(w) - f_N(\operatorname{Re}(w_{c,N}^+) + i\zeta a^2)\}| \leq e^{\zeta^2 a^4}.$$

If one chooses  $\zeta$  in (2.24) small enough so that for any  $v \in [\epsilon, 1 - \epsilon]$ ,  $\zeta a^2 \leq \operatorname{Im}(w_{c,N}^+)/8$ , one can then deduce that the contribution of the contour  $\Upsilon \cap \{z, \operatorname{Im}(z) \leq \zeta a^2\}$  is negligible. Thus, using (2.28), the main contribution to the  $w$ -integral comes from a neighborhood of width  $\sqrt{S}$  of the critical point.

We now turn to the  $z$ -contour. Let

$$\begin{aligned} \Gamma_N^+ &= \{w_{c,N}^+(t), \epsilon/2 \leq t \leq 1 - \epsilon/2\} \cup \{w_{c,N}^+(1 - \epsilon/2) + x, x > 0\} \\ &\quad \cup \{i\operatorname{Im}(w_{c,N}^+(\epsilon/2)) + x, \epsilon/180 \leq x \leq \operatorname{Re}(w_{c,N}^+(\epsilon/2))\}. \end{aligned} \quad (2.29)$$

We can now define  $x_1^\pm := \epsilon/180 \pm i\operatorname{Im}(w_{c,N}^+(\epsilon/2))$ . The choice of 180 here is arbitrary: we only need a sufficiently large constant. We also set  $x_o^\pm = -x_1^\mp$ .

By construction, for any  $t$ ,  $\frac{\epsilon}{2} \leq t \leq 1 - \frac{\epsilon}{2}$ ,

$$\operatorname{Re}(-f_N(w_{c,N}^+(t)) + f_N(w_{c,N}^+)) \leq -c(\sqrt{u} - \sqrt{t})^2,$$

for some constant  $c > 0$  small enough. This follows from the fact that for any  $t$ ,  $\frac{\epsilon}{2} \leq t \leq 1 - \frac{\epsilon}{2}$ ,  $\operatorname{Im}(w_{c,N}^+(t)) \geq \zeta a^2$  and  $\operatorname{Re}\left(\frac{dw_{c,N}^+(t)}{dt}\right) = \frac{1+O(a^2)}{2\sqrt{t}}$ . Moreover  $\operatorname{Re}(f_N(w_{c,N}^+(1 - \frac{\epsilon}{2})) - f_N(w_{c,N}^+(1 - \frac{\epsilon}{2}) + x)) \leq -cx^2$ , for some constant  $c > 0$  small enough. This follows from the fact that on  $\Omega_N$   $a^2 z m_N(z^2) \rightarrow 0$  uniformly along  $\Gamma_N$  as  $m_N$  is bounded. Thus the contribution of  $\{w_{c,N}^+(1 - \frac{\epsilon}{2}) + x, x > 0\}$  is negligible as  $N \rightarrow \infty$ . For the same reason, the contribution of  $\{i\operatorname{Im}(w_{c,N}^+(\frac{\epsilon}{2})) + x, \epsilon/180 \leq x \leq \operatorname{Re}(w_{c,N}^+(\frac{\epsilon}{2}))\}$  is negligible in the large  $N$  limit. As for the  $w$ -integral, the main contribution to the  $z$ -integral comes from a neighborhood of width  $\sqrt{S}$  of the critical point.

**Remark 2.2.** This analysis justifies Claim 2.1 as  $\operatorname{Re}(f_N(w_{c,N}^+) - f_N(x_1^+)) \leq -c$  for some  $c > 0$ . One can choose  $\Gamma \setminus \Gamma_c$  as  $\{x_1^+ + ia^2 t, 0 < t < A - \operatorname{Im}(x_1^+)\} \cup \{x_o^+ + ia^2 t, 0 < t < A - \operatorname{Im}(x_1^+)\} \cup \{x + ia^2 A, \operatorname{Re}(x_o^+) \leq x \leq \operatorname{Re}(x_1^+)\}$  plus its conjugate, for some large enough constant  $A$ . There exists  $C > 0$  such that  $\operatorname{Re}(f_N(x_1^+) - f_N(z)) \leq Ca^2, \forall z \in \Gamma \setminus \Gamma_c$ , thus yielding a negligible contribution.

We can now conclude to the asymptotic expansion of the correlation kernel and prove Proposition 2.2. Let  $u_*$  be as in Proposition 2.2.

We first consider the asymptotics for the first term (2.13) in  $K_N^{1,b}$ . We now fix  $b$  as follows:

$$b = \operatorname{Re}(w_{c,N}^+(u_*)).$$

Thanks to this choice,  $|\operatorname{Re}(w_{c,N}^+) - b| = o(N^{\lambda-1})$ . Therefore the function  $\theta$  has no impact on the saddle point argument exposed in the above, neither do the functions  $g_{N,u}$  and  $g_{N,u}^1$ . At the critical points, one deduces from (2.12) that:

$$\begin{aligned} g_{N,u}(w_{c,N}^\pm, w_{c,N}^\pm) &= (w_{c,N}^\pm - b) f_N''(w_{c,N}^\pm) \simeq i \operatorname{Im}(w_{c,N}^\pm) f_N''(w_{c,N}^\pm); \\ g_{N,u}(w_{c,N}^\pm, w_{c,N}^\mp) &= 0; \\ g_{N,u}^1(w_{c,N}^\pm, w_{c,N}^\pm) &= g_{N,u}^1(w_{c,N}^\pm, w_{c,N}^\mp) = O(S) = O(\operatorname{Im}(w_{c,N}^\pm)/N); \\ \theta(w_{c,N}^\pm, b) &= \frac{e^{i2(\sqrt{u}-\sqrt{v})\operatorname{Im}(w_{c,N}^\pm)/S} - 1}{2i\operatorname{Im}(w_{c,N}^\pm)(\sqrt{u}-\sqrt{v})} (1 + o(1)). \end{aligned}$$

Let us now consider the 4 combined contributions of the different critical points. The contribution to (2.13) from  $z = w = w_{c,N}^\pm$  gives (at the leading order)

$$\frac{\pm i}{4i^2\pi^2 S} \left( \frac{\sqrt{2\pi S}}{\sqrt{f_N''(w_{c,N}^\pm)}} \right)^2 \frac{1}{(uv)^{1/4}} f_N''(w_{c,N}^\pm) \frac{e^{2\operatorname{Im}w_{c,N}^\pm(\sqrt{u}-\sqrt{v})/S} - 1}{(\sqrt{u}-\sqrt{v})}. \quad (2.30)$$

The contribution to (2.13) from  $z = \bar{w} = w_{c,N}^\pm$  is  $O(N^{1-\lambda/2})$  due to the fact that  $g_{N,u}$  anneals at that point. Combining the above yields that:

$$\lim_{N \rightarrow \infty} \frac{1}{N\rho^{\text{MP}}(u)} (2.13) = \lim_{N \rightarrow \infty} \frac{1}{N\rho^{\text{MP}}(u)} (2.30) = \frac{\sin(\pi\tau)}{\pi\tau}.$$

The contributions of the two other terms in (2.14) is exponentially small since there exists a constant  $c > 0$  depending on  $\epsilon$  such that:

$$\operatorname{Re} \left( f_N \left( w_{c,N}^\pm \left( \frac{\epsilon}{2} \right) \right) \right) - \operatorname{Re}(f_N(w_{c,N}^+)) > c.$$

This finishes the asymptotic expansion of  $K_N^{1,b}$ .

Let us now turn to the asymptotics for  $K_N^{2,b}$ . The function  $\theta^1$  is not bounded. Nevertheless there exists  $\chi > 0$  depending on  $\epsilon$  such that  $|e^{-4b\sqrt{u}/S}\theta^1(z, b)| \leq e^{-\chi S^{-1}}$ . This follows from the fact that  $\operatorname{Re}(b - z) < b$  along the contour  $\Gamma_N$ . One can now use the same saddle point arguments as for the study of  $K_N^{1,b}$  to show that

$$\frac{1}{N\rho^{\text{MP}}(u)} K_N^{2,b}(u, v; H) = O(e^{-\chi S^{-1}}).$$

This is enough to ensure Proposition 2.2.

**Remark 2.3.** In the case where  $|u - u_*| = o(N^{-c})$  for some  $c \leq 1 - \lambda$ , we choose a sequence  $\hat{u}(N)$  such that  $N^{1-\lambda}|u - \hat{u}(N)| \rightarrow 0$ ,  $|u_* - \hat{u}(N)| = o(N^{-c})$  and  $\hat{u}(N) \in [u_- + \epsilon, u_+ - \epsilon]$  for all  $N$ . We can then replace  $u_* \rightarrow \hat{u}(N)$  in all the asymptotic analysis (in particular  $b = b(N) = \operatorname{Re}(w_{c,N}^+(\hat{u}(N)))$ ). This has no impact on the validity of the arguments.

2.4. The case where  $v \rightarrow \infty$ 

We now turn to the case where  $v$  is unbounded (that is either  $\gamma = 1$  and  $v \rightarrow \infty$  or  $\gamma \neq 1$ ) and study the asymptotics of the correlation kernel (2.4). The aim of this section is to prove the following Proposition. Let  $\epsilon > 0$  be given (small).

**Proposition 2.3.** *Assume that  $\gamma \geq 1$  and  $\lambda < 1/2$ . Let  $u_* \in [u_- + \epsilon, u_+ - \epsilon]$  and  $u = u(N)$  be a sequence such that  $\lim_{N \rightarrow \infty} N^{1-\lambda}(u - u_*) = 0$ . Then there exists  $b \in \mathbb{R}$  such that outside a set of negligible probability,*

$$\lim_{N \rightarrow \infty} \frac{1}{N \rho_\gamma^{\text{MP}}(u)} K_N^b \left( u, u + \frac{\tau}{N \rho_\gamma^{\text{MP}}(u)}; H \right) = \frac{\sin(\pi \tau)}{\pi \tau}.$$

**Remark 2.4.** *Proposition 2.3 extends to the case where  $|u - u_*| = o(N^{-c})$  for some  $c \leq 1 - \lambda$  by the same arguments as in Remark 2.3.*

The whole subsection is devoted to the proof of Proposition 2.3. The proof follows the same steps as in Section 2.3.2 and we explain the main changes only.

## 2.4.1. Rewriting the kernel

Again the basic argument is to replace Bessel functions with their asymptotic expansion (large order large argument) and then derive the asymptotics of the correlation kernel by a saddle point argument.

We again cut the contour  $\Gamma$  in a small neighborhood of the imaginary axis as on Fig. 2. We also call  $\Gamma_c$  this cut contour and the endpoints  $x_1^\pm, x_o^\pm$  will be defined in the sequel. Let us denote again  $\Gamma_1$  the part of  $\Gamma_c$  lying to the right of the imaginary axis, so that  $\Gamma_c := \Gamma_1 \cup (-\Gamma_1)$ . We now consider the uniform asymptotic expansion (B.2) of Appendix B. We assume for a while that we can replace the modified Bessel functions with their asymptotic expansion in the correlation kernel (this will be proved in Lemma 2.4 below). Thus, setting  $v = u + \frac{\tau}{N \rho_\gamma^{\text{MP}}(u)}$ , we consider the kernel:

$$\begin{aligned} \widetilde{K}_N(u, v; H) &:= \frac{1}{4i^2 \pi^2 S(uv)^{1/4}} \int_{i\mathbb{R}+A} dw \int_{\Gamma_1} dz \frac{e^{(w^2 - z^2 + 2(z-b)\sqrt{u} - 2(w-b)\sqrt{v})/S}}{\sqrt{w}\sqrt{z}} \\ &\quad \times e^{((\gamma-1)^2 N^\lambda / (4u_*))(1/w - 1/z)} \prod_{i=1}^N \frac{w^2 - y_i}{z^2 - y_i} \left(\frac{w}{z}\right)^v \frac{w+z}{w-z} \end{aligned} \quad (2.31)$$

$$\begin{aligned} &- \frac{1}{4i^2 \pi^2 S(uv)^{1/4}} \int_{i\mathbb{R}+A} dw \int_{\Gamma_1} dz \frac{e^{(w^2 - z^2 + 2(z-b)\sqrt{u} - 2(w-b)\sqrt{v})/S}}{\sqrt{w}\sqrt{z}} \\ &\quad \times e^{((\gamma-1)^2 N^\lambda / (4u_*))(1/w - 1/z)} \prod_{i=1}^N \frac{w^2 - y_i}{z^2 - y_i} \left(\frac{w}{z}\right)^v \frac{w-z}{w+z}. \end{aligned} \quad (2.32)$$

We call again  $K_N^{1,b}$  (resp.  $K_N^{2,b}$ ) the kernel given by (2.31) (resp. (2.32)). Set

$$\Psi(w, z) = \exp \left\{ \frac{(\gamma-1)^2 N^\lambda}{4u_*} (1/w - 1/z) \right\}.$$

We can now slightly modify the arguments of Section 2.3.1 to deduce that:

$$\begin{aligned} K_N^{1,b}(u, v; H) &= \frac{1}{(uv)^{1/4} 4i^2 \pi^2 S} \int_{i\mathbb{R}+A} dw \int_{\Gamma_1} dz e^{(f_N(w) - f_N(z))/S} \Psi(w, z) \\ &\quad \times \frac{w+z}{\sqrt{w}\sqrt{z}} (g_{N,u}(w, z) + g_{N,u}^2(w, z)) \left(\frac{w}{z}\right)^v \theta(w, b) \end{aligned} \quad (2.33)$$

$$\begin{aligned}
& - \frac{(x_1^- - b)}{4i^2 \pi^2 (uv)^{1/4}} \int_{i\mathbb{R}+A} dw \frac{w + x_1^-}{\sqrt{w} \sqrt{x_1^-}} \left( \frac{w}{x_1^-} \right)^v e^{(f_N(w) - f_N(x_1^-))/S} \theta(w, b) \Psi(w, x_1^-) \\
& + \frac{(x_1^+ - b)}{4i^2 \pi^2 (uv)^{1/4}} \int_{i\mathbb{R}+A} dw \frac{w + x_1^+}{\sqrt{w} \sqrt{x_1^+}} \left( \frac{w}{x_1^+} \right)^v e^{(f_N(w) - f_N(x_1^+))/S} \theta(w, b) \Psi(w, x_1^+),
\end{aligned} \tag{2.34}$$

where

$$g_{N,u}^2(w, z) = g_{N,u}^1(w, z) + \frac{SN^\lambda(\gamma - 1)^2}{4u_*} \left( \frac{1}{wz} - b \frac{w+z}{z^2 w^2} \right).$$

We leave  $K_N^{2,b}$  unchanged since the singularity  $1/(w+z)$  will not prevent its direct saddle point analysis.

#### 2.4.2. Saddle point analysis

We shall now perform the saddle point analysis of the correlation kernels hereabove. The arguments follow closely those of Section 2.3. The approximate exponential term to be considered here is given by

$$h_u(z) := z^2 - 2\sqrt{u}(z-b) + a^2 \int \ln(z^2 - y) d\rho_\gamma^{\text{MP}}(y) + a^2(\gamma - 1) \ln z.$$

We are indeed going to show that the perturbative term coming from  $\Psi$  does not play a role in the asymptotics. Let  $m_\gamma(z)$  be the Stieltjes transform of the Marčenko–Pastur distribution  $\rho_\gamma^{\text{MP}}$ . One can check that

$$m_\gamma(z) = -2 + \frac{\gamma - 1}{2z} + 2 \frac{\sqrt{(z - (1 + \sqrt{\gamma})^2/4)(z - (1 - \sqrt{\gamma})^2/4)}}{z}.$$

The function  $h_u$  admits two non-real critical points which are conjugate:

$$w_c^\pm = w_c^\pm(u) = \sqrt{u}(1 - 2a^2) \pm ia^2 \pi \sqrt{u} \rho_\gamma^{\text{MP}}(u) + O(a^4). \tag{2.35}$$

We now define the contours relevant for the saddle point analysis of the approximate exponential term. Let  $\mathcal{Y} = \text{Re}(w_c^+) + it, t \in \mathbb{R}$ , be oriented positively from bottom to top. The contour  $\Gamma_1$  is defined as in Section 2.3.2: we set  $\Gamma_1 = \Gamma_1^+ \cup \overline{\Gamma_1^+}$  where  $\Gamma_1^+$  is oriented counterclockwise and

$$\begin{aligned}
\Gamma_1^+ &= \{w_c^+(t), t \in [u_- + \epsilon/90, u_+ - \epsilon/2]\} \cup \{w_c^+(u_+ - \epsilon/2) + x, x > 0\} \\
&\cup \left\{ i \text{Im} \left( w_c^+ \left( u_- + \frac{\epsilon}{90} \right) \right) + x, \epsilon/180 \leq x \leq \text{Re} \left( w_c^+ \left( u_- + \frac{\epsilon}{90} \right) \right) \right\}.
\end{aligned}$$

We first consider the  $z$ -contour. There exists  $c(\epsilon) > 0$  depending on  $\epsilon$  only such that for all points  $t \in [u_- + \epsilon/90, u_+ - \epsilon/2]$  in the bulk of the Marčenko–Pastur distribution  $\text{Re}(h_u''(w_c^+(t))) > c(\epsilon)$ . Using the fact that

$$\frac{d}{dt} \text{Re}(h_u(w_c^+(t))) = (\sqrt{t} - \sqrt{u}) \text{Re} \frac{1}{\sqrt{t} h_u''(w_c^+(t))}, \tag{2.36}$$

$-\text{Re}(h_u)$  achieves its maximum at  $z = w_c^\pm$  along the first part of  $\Gamma_1^+$ . There now remains to show that  $-\text{Re}(h_u)$  also decreases along  $\{w_c^+(u_+ - \epsilon/2) + x, x > 0\}$  and  $\{w_c^+(u_- + \frac{\epsilon}{90}) - x, 0 < x < \text{Re}(w_c^+(u_- + \frac{\epsilon}{90})) - \frac{\epsilon}{180}\}$ . This assertion follows from the fact that  $m_\gamma$  is bounded along this curve. Using the same arguments as in Section 2.3.2, it is easily seen that  $\text{Re}(h_u)$  admits its maximum on  $\mathcal{Y}$  at  $w_c^\pm$ . All the above reasoning holds unchanged if  $\gamma$  is replaced by  $p/N$ . Thus, without loss of generality, we assume in the rest of this section that  $p/N = \gamma$ .

We now consider the true exponential term. We set

$$h_{N,u}(z) = h_N(z) := z^2 - 2\sqrt{u}(z-b) + \frac{a^2}{N} \sum_{i=1}^N \ln(z^2 - y_i) + a^2(\gamma - 1) \ln z.$$

Again we can show that, provided  $H$  belongs to a suitable set, the derivatives of  $h_u$  and  $h_{N,u}$  are very close and that  $h_{N,u}$  admits two non-real critical points  $w_{c,N}^\pm$  very close to  $w_c^\pm$ . Let  $\zeta > 0$  be given (small) and  $S_{\epsilon,\zeta}$  be given by (2.24). Define for  $H = H_N = \frac{WW^*}{N}$ ,

$$\Omega_{N,\gamma} = \left\{ \begin{aligned} &\sup y_i(H) \leq K; \quad \sup_{|x| > (\epsilon/200)^2, y \geq \zeta a^2} |m_N(x + iy)| \leq K; \\ &|m_N(z) - m(z)| \leq N^{-\lambda/4}, \forall z \text{ with } \zeta a^2 \epsilon \leq \text{Im}(z) \leq 1 \text{ and } u_- + \frac{\epsilon}{180} \leq \text{Re}(z) \leq u_+ - \frac{\epsilon}{8}. \end{aligned} \right. \quad (2.37)$$

Using Proposition 2.1 and Lemma 7.3 in [11], we can deduce that there exist  $K$  large enough and constants  $c, C > 0$  such that

$$\mathbb{P}(\Omega_{N,\gamma}^c) \leq C e^{-cN^{\lambda/4}}.$$

From now on we assume that  $H_N \in \Omega_{N,\gamma}$ . Mutatis mutandis, all the arguments used to consider the true exponential term when  $\nu$  is bounded (starting with (2.26)) can be copied to consider the true exponential term when  $\nu$  is not bounded. In particular, one has that

$$|w_{c,N}^\pm - w_c^\pm| \leq \frac{Ca^2}{N^{\lambda/4}}, \quad |h_{N,u}^{(l)}(w) - h_u^{(l)}(w)| \leq \frac{C}{a^{2(l-2)}N^{\lambda/4}} \quad \forall l \geq 1, \forall w \in S_{\epsilon,\zeta}.$$

We now consider the subsequent modified contours. We set  $\Gamma_N = \Gamma_N^+ \cup \overline{\Gamma_N^+}$  with

$$\begin{aligned} \Gamma_N^+ &= \left\{ w_{c,N}^+(t), u_- + \frac{\epsilon}{90} \leq t \leq u_+ - \frac{\epsilon}{2} \right\} \cup \left\{ w_{c,N}^+ \left( u_+ - \frac{\epsilon}{2} \right) + x, x > 0 \right\} \\ &\cup \left\{ i \text{Im} \left( w_{c,N}^+ \left( u_- + \frac{\epsilon}{90} \right) \right) + x, \epsilon/180 \leq x \leq \text{Re} \left( w_{c,N}^+ \left( u_- + \frac{\epsilon}{90} \right) \right) \right\}; \\ \Upsilon_N &= \text{Re}(w_{c,N}^+) + it, \quad t \in \mathbb{R}. \end{aligned} \quad (2.38)$$

Accordingly, we now choose  $x_1^\pm = \pm i \text{Im}(w_{c,N}^+(u_- + \frac{\epsilon}{90})) + \frac{\epsilon}{180}$  and  $x_o^\pm = -x_1^\mp$ .

Then the main contribution to the exponential term of (2.33) comes from a neighborhood of width  $\sqrt{S}$  of  $w_{c,N}^\pm$  for both the  $w$ - and  $z$ -integral. Thus one is left with showing that the function  $\Psi$  does not impact on the saddle point analysis. Let  $\mathcal{O}_w = \{w' \in \mathbb{C}, |w' - w| \leq N^\alpha \sqrt{S}\}$  for some  $\alpha$  that we determine hereafter. In order to ensure that  $\Psi$  has no impact on the saddle point analysis outside  $\mathcal{O}_{w_{c,N}^\pm}$ , it is enough that  $\alpha > \lambda/2$ . Now one can check that

$$\forall w, z \in \mathcal{O}_{w_{c,N}^\pm}, \quad |\Psi(w, z) - 1| \leq CN^{3\lambda/2 + \alpha - 1} \ll 1,$$

if  $\alpha$  and  $\lambda$  are small enough (in particular  $\lambda < 1/2$ ). This finishes the proof that the main contribution to (2.33) comes from a neighborhood of width  $\sqrt{S}$  of  $w_{c,N}^\pm$  for both the  $w$ - and  $z$ -integral.

We now set  $b = \text{Re}(w_{c,N}^+(u_*))$ , so that  $|b - w_{c,N}^+| = o(N^{\lambda-1})$ , which implies in particular that the function  $\theta$  does not impact the saddle point argument. Let us now consider the contribution of the functions  $g_{N,u}(w, z) + g_{N,u}^2(w, z)$  and  $\theta(w, z)$  close to the critical points. One has that

$$\begin{aligned} g_{N,u}(w_{c,N}^\pm, w_{c,N}^\pm) + g_{N,u}^2(w_{c,N}^\pm, w_{c,N}^\pm) &= (w_{c,N}^\pm - b)h_N''(w_{c,N}^\pm) + O(S) \\ &\quad + \frac{SN^\lambda(\gamma - 1)^2}{4u_*} \frac{w_{c,N}^\pm - 2b}{(w_{c,N}^\pm)^3} \\ &= (w_{c,N}^\pm - b)h_N''(w_{c,N}^\pm) + O(SN^\lambda); \end{aligned}$$

$$g_{N,u}(w_{c,N}^\pm, w_{c,N}^\mp) + g_N^2(w_{c,N}^\pm, w_{c,N}^\mp) = O(SN^\lambda);$$

$$\theta(w_{c,N}^\pm, b) = n \frac{e^{i2(\sqrt{u}-\sqrt{v})\text{Im}(w_{c,N}^\pm)/S} - 1}{2(w_{c,N}^\pm - b)(\sqrt{u} - \sqrt{v})} (1 + o(1)).$$

We can copy the end of the proof of Proposition 2.2 to show that on  $\Omega_{N,\gamma}$

$$\lim_{N \rightarrow \infty} \frac{1}{N\rho_\gamma^{\text{MP}}(u)} K_N^{1,b}(u, v; H) = \frac{\sin(\pi\tau)}{\pi\tau}.$$

To study the asymptotics of  $K_N^{2,b}$ , we use the same saddle point analysis argument. Note that this is possible since  $\Upsilon_N \cap (-\Gamma_N) = \emptyset$ . We consider as leading exponential term for both the  $w$ - and  $z$ -integrals the function  $h_{N,u}$  as above. Note that the  $w$ -integrand has then a ‘‘perturbative term’’  $e^{2(w-b)(\sqrt{u}-\sqrt{v})/S}$  which does not play a role in the saddle point analysis. It is easy to deduce that

$$\lim_{N \rightarrow \infty} \frac{1}{N\rho_\gamma^{\text{MP}}(u)} K_N^{2,b}(u, v; H) = 0.$$

This follows from the fact that the contribution from equal critical points is negligible (due to the  $(w-z)$  factor), while that of conjugate critical points is of order  $1/N$  due to the rescaling and the fact that  $|e^{S^{-1}(h_{N,u}(z_{c,N}^\pm(u)) - h_{N,v}(w_{c,N}^\mp(v)))}| \leq C$  for some  $C > 0$  independent of  $N$ .

Thus we are now left with showing the following lemma.

**Lemma 2.4.** *One has that*

$$\lim_{N \rightarrow \infty} \frac{1}{N\rho_\gamma^{\text{MP}}(u)} |\widetilde{K}_N(u, v; H) - K_N^b(u, v; H)| = 0.$$

**Proof.** Note that it is enough to show that  $|\widetilde{K}_N(u, v; H) - K_N^b(u, v; H)| = o(N)$ . First Claim 2.1 can be translated with no modification to the case where  $v$  is unbounded. Also we can reduce  $\Gamma_c$  to  $\Gamma_1$  in  $K_N^b$  using the change of variables  $z \mapsto -z$ . There remains to prove the counterpart of Claim 2.2. Set

$$Z = \frac{2\sqrt{u}z}{(\gamma-1)a^2}, \quad Z_* = \frac{2\sqrt{u_*}z}{(\gamma-1)a^2}, \quad W = \frac{2\sqrt{v}w}{(\gamma-1)a^2}, \quad W_* = \frac{2\sqrt{u_*}w}{(\gamma-1)a^2},$$

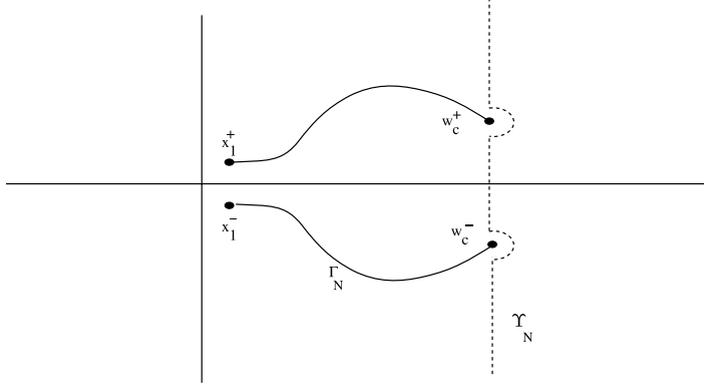
$$A'_1(Z) := \frac{e^{\nu(Z-1/(2Z))}}{\sqrt{2\pi\nu Z}}, \quad A_1(Z) = \frac{e^{\nu(Z-1/(2Z_*))}}{\sqrt{2\pi\nu Z}},$$

$$A'_2(W) := \sqrt{\pi} \frac{e^{-\nu(W-1/(2W))}}{\sqrt{2\nu W}}, \quad A_2(W) = \sqrt{\pi} \frac{e^{-\nu(W-1/(2W_*))}}{\sqrt{2\nu W}}.$$

In  $\widetilde{K}_N$  we have replaced the Bessel functions  $I_\nu$  and  $K_\nu$  with the ‘‘approximations’’  $A_1$  and  $A_2$ . It is an easy computation that, along  $\Upsilon_N$  and  $\Gamma_N$ , one has:

$$\left| \frac{I_\nu(\nu Z)}{A'_1(Z)} - 1 \right| \leq O(S), \quad \left| \frac{K_\nu(\nu W)}{A'_2(W)} - 1 \right| \leq O(S). \quad (2.39)$$

We first consider the contribution of  $\Gamma_N^l := \Gamma_N \cap \{\text{Re}(x_1) \leq \text{Re}(z) \leq \text{Re}(w_{c,N}^+)\}$  to  $\widetilde{K}_N(u, v; H) - K_N^b(u, v; H)$  where the new contour  $\Upsilon_N^{\epsilon'}$  is slightly deformed around  $w_{c,N}^\pm$  so that  $d(\Upsilon_N, \Gamma_N^l) = \epsilon'\sqrt{S}$  for some small  $\epsilon'$  (see Fig. 3). It is

Fig. 3. The contour  $\Gamma_N$  has been deformed around  $w_{c,N}^\pm$ .

then easy to deduce from (2.39) and the previous saddle point analysis that

$$\begin{aligned} & \left| \frac{e^{2b(\sqrt{v}-\sqrt{u})/S}}{2i^2\pi^2S^2} \int_{\Gamma_N^{\epsilon'}} dw \int_{\Gamma_N^l} dz e^{(-z^2+w^2)/S} \prod_{i=1}^N \frac{w^2 - y_i}{z^2 - y_i} \left(\frac{w}{z}\right)^v \frac{w+z}{w-z} \right. \\ & \quad \left. \times \left( A_2'(W) I_\nu \left( \frac{2z\sqrt{u}}{S} \right) - A_2'(W) A_1'(Z) \right) \right| \\ & \leq O(\sqrt{S}). \end{aligned} \quad (2.40)$$

Indeed replacing  $A_1$  and/or  $A_2$  with  $A_1'$  and/or  $A_2'$  does not impact on the saddle point argument. Similarly

$$\begin{aligned} & \left| \frac{e^{2b(\sqrt{v}-\sqrt{u})/S}}{2i^2\pi^2S^2} \int_{\Gamma_N^{\epsilon'}} dw \int_{\Gamma_N^l} dz e^{(-z^2+w^2)/S} \prod_{i=1}^N \frac{w^2 - y_i}{z^2 - y_i} \left(\frac{w}{z}\right)^v \frac{w+z}{w-z} \right. \\ & \quad \left. \times \left( A_2'(W) I_\nu \left( \frac{2z\sqrt{u}}{S} \right) - K_\nu \left( \frac{2w\sqrt{v}}{S} \right) I_\nu \left( \frac{2z\sqrt{u}}{S} \right) \right) \right| \\ & \leq O(\sqrt{S}), \end{aligned} \quad (2.41)$$

by using (2.39), the fact that  $I_\nu(\nu Z)/A_1'(Z)$  is uniformly bounded for  $N$  large enough and the previous saddle point analysis. Lastly

$$\begin{aligned} & \left| \frac{e^{2b(\sqrt{v}-\sqrt{u})/S}}{2i^2\pi^2S^2} \int_{\Gamma_N^{\epsilon'}} dw \int_{\Gamma_N^l} dz e^{(-z^2+w^2)/S} (A_2(W)A_1(Z) - A_2'(W)A_1'(Z)) \right. \\ & \quad \left. \times \prod_{i=1}^N \frac{w^2 - y_i}{z^2 - y_i} \left(\frac{w}{z}\right)^v \frac{w+z}{w-z} \right| \\ & \leq O\left( \frac{N^{\lambda+\alpha}\sqrt{S}|u-u_*|}{\epsilon'\sqrt{S}} + \frac{N^\lambda|u-u_*|}{N^\alpha\sqrt{S}} + N^{\lambda/2} \right) \end{aligned} \quad (2.42)$$

$$\leq O(N^{\alpha+2\lambda-1} + N^{3\lambda/2-\alpha} + N^{\lambda/2}) = o(N). \quad (2.43)$$

In (2.42) we have separated the cases where  $d(w, z) \geq N^\alpha\sqrt{S}$  or not, where  $z \in \Gamma_N^l$  and  $w \in \Gamma_N^{\epsilon'}$  and used the fact that  $|u-v| \simeq N^{-1}$  and  $|u-u_*| = o(N^{\lambda-1})$ . Replacing  $\frac{w+z}{w-z}$  with  $\frac{w-z}{w+z}$  in (2.40), (2.41) and (2.43) yields similar estimates.

We now turn to the contribution of  $\Gamma_N^r := \Gamma_N \cap \{\operatorname{Re}(z) \geq \operatorname{Re}(w_{c,N}^+)\}$  to  $\widetilde{K}_N(u, v; H) - K_N^b(u, v; H)$ . Note that  $\Upsilon_N$  can be moved to the left of  $\operatorname{Re}(w_{c,N}^+)$  up to adding the residue at  $w = z$ . Define  $\Upsilon_N^{-\epsilon'}$  as the contour obtained by reflecting  $\Upsilon_N^{\epsilon'}$  with respect to the line  $\operatorname{Re}(z) = \operatorname{Re}(w_{c,N}^+)$ . Then

$$\begin{aligned} & \int_{A+i\mathbb{R}} dw \int_{\Gamma_N^r} dz e^{(-z^2+w^2)/S} \prod_{i=1}^N \frac{w^2 - y_i}{z^2 - y_i} \left(\frac{w}{z}\right)^v \frac{w+z}{w-z} \\ & \quad \times \left( K_\nu \left( \frac{2w\sqrt{v}}{S} \right) I_\nu \left( \frac{2z\sqrt{u}}{S} \right) - A'_1(Z) A'_2(W) \right) \\ & = \int_{\Upsilon_N^{-\epsilon'}} dw \int_{\Gamma_N^r} dz e^{(-z^2+w^2)/S} \prod_{i=1}^N \frac{w^2 - y_i}{z^2 - y_i} \left(\frac{w}{z}\right)^v \frac{w+z}{w-z} \\ & \quad \times \left( K_\nu \left( \frac{2w\sqrt{v}}{S} \right) I_\nu \left( \frac{2z\sqrt{u}}{S} \right) - A'_1(Z) A'_2(W) \right) \end{aligned} \quad (2.44)$$

$$+ 2i\pi \int_{\Gamma_N^r} dz 2z \left( K_\nu \left( \frac{2z\sqrt{v}}{S} \right) I_\nu \left( \frac{2z\sqrt{u}}{S} \right) - A'_1(Z) A'_2 \left( \frac{2z\sqrt{v}}{a^2(\gamma-1)} \right) \right). \quad (2.45)$$

The analysis of (2.44) is similar to that of  $\Gamma_N^l \cap \Upsilon_N^{\epsilon'}$ . We thus have to consider (2.45). One has that

$$|(2.45)| = O \left( S \int_{\Gamma_N^r} dz 2z A'_1(Z) A'_2 \left( \frac{2z\sqrt{v}}{a^2(\gamma-1)} \right) \right). \quad (2.46)$$

As the integrand in the r.h.s. of (2.46) has no singularity, one can thus move  $\Gamma_N^r$  to the line joining the two critical points  $w_{c,N}^\pm$ . Using now that  $|b - \operatorname{Re}(w_{c,N}^+)| = o(N^{\lambda-1})$ , one can easily see that there exists a constant  $C > 0$  such that

$$e^{2b(\sqrt{v}-\sqrt{u})/S} (2.46) S^{-2} \leq C.$$

One gets similar estimates replacing  $\frac{w+z}{w-z}$  with  $\frac{w-z}{w+z}$ . This yields Lemma 2.4.  $\square$

### 3. Proof of Theorem 1.1 and Theorem 1.2

We only give the proof of Theorem 1.1 when  $\gamma = 1$  and  $\nu$  is bounded. The extension to other parameters  $\gamma$  is straightforward. The proof of universality for the spacing distribution can easily be deduced from [17] and [14] and is based on the extension of Proposition 2.2 given in Remark 2.1. Let then  $f \in L^\infty(\mathbb{R}^2)$  with compact support and  $S_N^{(2)}(f, u, \widetilde{\rho}_N)$  be defined by (1.2) with  $u \in [\epsilon, 1 - \epsilon]$  and  $\widetilde{\rho}_N = N\rho^{\text{MP}}(u)$ . Then if  $\mathbb{E}$  denotes the expectation w.r.t. the distribution of  $\widetilde{M}_N$  in (2.1),

$$\mathbb{E} S_N^{(2)}(f, u, \widetilde{\rho}_N) = \int dP_N(H) \int_{\mathbb{R}_+^N} f_N^H(x_1, \dots, x_N) S_N^{(2)}(f, u, \widetilde{\rho}_N)[x] dx^N,$$

where  $f_N^H(\cdot)$  (resp.  $S_N^{(2)}(f, u, \widetilde{\rho}_N)[\cdot]$ ) is the density function (2.3) (resp. local eigenvalue statistic) of the deformed Wishart ensemble. Then, using (2.25) and setting

$$d\overline{P}_N(H) = \frac{1}{Z_N} dP_N(H) 1_{H \in \Omega_N},$$

where  $Z_N$  is the normalizing constant, we get that  $\exists C > 0$  such that

$$\begin{aligned} & \left| \mathbb{E} S_N^{(2)}(f, u, \widetilde{\rho}_N) - \int d\overline{P}_N(H) \int_{\mathbb{R}_+^N} f_N^H(x) S_N^{(2)}(f, u, \widetilde{\rho}_N)[x] dx^N \right| \\ & \leq CN^2 |f|_\infty e^{-cN^{\lambda/4}} = o(1). \end{aligned} \quad (3.1)$$

Now,

$$\begin{aligned}
& \lim_{N \rightarrow \infty} \int d\bar{P}_N(H) \int_{\mathbb{R}_+^N} f_N^H(x) S_N^{(2)}(f, u, \widetilde{\rho}_N)[x] dx^N \\
&= \lim_{N \rightarrow \infty} \int d\bar{P}_N(H) \int_{\mathbb{R}^2} \frac{f(t_1, t_2)}{\widetilde{\rho}_N^2} R_N^{(2)}\left(u + \frac{t_1}{\widetilde{\rho}_N}, u + \frac{t_2}{\widetilde{\rho}_N}; H\right) dt_1 dt_2 \\
&= \int_{\mathbb{R}^2} f(t_1, t_2) \left(1 - \left(\frac{\sin \pi(t_1 - t_2)}{\pi(t_1 - t_2)}\right)^2\right) dt_1 dt_2, \tag{3.2}
\end{aligned}$$

where we used the definition of correlation functions, Proposition 2.2 and the fact that  $f$  has compact support.

There now remains to extend the result to non-Gauss divisible ensembles in order to prove the full Theorem 1.1. The argument exactly follows the arguments of [14]. We recall them for seek of completeness. (3.1) and (3.2) show that the sine kernel holds for the complex measure  $e^{t\mathcal{L}} G_t d\mu^{\otimes Np} \otimes e^{t\mathcal{L}} G_t d\mu^{\otimes Np}$  if  $t = N^{-1+\lambda}$ . More precisely, let  $p_{N,t}(x)$  denote the density function of the eigenvalues  $x = (x_1, \dots, x_N)$  w.r.t. this measure and let  $R_{N,t}^{(2)}$  be its two point correlation function. Similarly, we define  $p_{N,c}(x)$  (resp.  $p_N(x)$ ) and  $R_{N,c}^{(2)}$  (resp.  $R_{N,F}^{(2)}$ ) for the eigenvalue density and two point correlation function w.r.t. truncated complex measure  $F_{c,N,p} = f_{c,N,p} d\mu^{\otimes Np} \otimes f_{c,N,p} d\mu^{\otimes Np}$  (resp. w.r.t. measure  $F_{N,p} = F^{\otimes Np}$ ). We also set in the following

$$u(t_1) = u + \frac{t_1}{\rho_N} \quad \text{and} \quad u(t_2) = u + \frac{t_2}{\rho_N} \quad \text{with} \quad \rho_N = N\rho_{1,1}^{\text{MP}}(u).$$

Then

$$\left| \int [R_{N,F}^{(2)}(u(t_1), u(t_2)) - R_{N,t}^{(2)}(u(t_1), u(t_2))] \frac{f(t_1, t_2)}{\rho_N^2} dt_1 dt_2 \right| \leq (I) + (II),$$

where

$$(I) := \left| \int [R_{N,F}^{(2)}(u(t_1), u(t_2)) - R_{N,c}^{(2)}(u(t_1), u(t_2))] \frac{f(t_1, t_2)}{\rho_N^2} dt_1 dt_2 \right|,$$

$$(II) := \int |R_{N,c}^{(2)}(u(t_1), u(t_2)) - R_{N,t}^{(2)}(u(t_1), u(t_2))| \left| \frac{f(t_1, t_2)}{\rho_N^2} \right| dt_1 dt_2.$$

It is easy to see from Proposition 1.1 that

$$(I) \leq N^2 \|f\|_\infty 2C e^{-cN^c} \leq C' e^{-c'N^c}$$

with some  $C', c' > 0$  as  $N \rightarrow \infty$ . To estimate (II), we use the fact that

$$(II)^2 \leq \int \left[ \frac{R_{N,c}^{(2)}(u(t_1), u(t_2))}{R_{N,t}^{(2)}(u(t_1), u(t_2))} - 1 \right]^2 R_{N,t}^{(2)}(u(t_1), u(t_2)) \left| \frac{f(t_1, t_2)}{\rho_N^2} \right| dt_1 dt_2 \tag{3.3}$$

$$\times \int R_{N,t}^{(2)}(u(t_1), u(t_2)) \left| \frac{f(t_1, t_2)}{\rho_N^2} \right| dt_1 dt_2. \tag{3.4}$$

The factor (3.4) is bounded using (3.1) and (3.2). We now use the beautiful idea of [14] to bound (3.3). As  $f$  is bounded, one has that

$$\begin{aligned}
((3.3))^{1/2} &\leq C \left[ \int \left[ \frac{R_{N,c}^{(2)}(z, y)}{R_{N,t}^{(2)}(z, y)} - 1 \right]^2 R_{N,t}^{(2)}(z, y) dz dy \right]^{1/2} \\
&\leq C \left[ N^2 \int \left( \frac{p_{N,c}(x)}{p_{N,t}(x)} - 1 \right)^2 p_{N,t}(x) dx \right]^{1/2}
\end{aligned}$$

$$\begin{aligned}
&\leq C \left[ N^2 \int \frac{|e^{t\mathcal{L}}G_t - f_{c,N,p}|^2}{e^{t\mathcal{L}}G_t} d\mu^{\otimes Np} \right]^{1/2} \\
&\leq CN^{-1+4\lambda},
\end{aligned} \tag{3.5}$$

where  $C > 0$  is a constant that varies from line to line. Here the basic argument is that the distance  $D(f, g) := \int |f/g - 1|^2 g$  between two probability measures  $f$  and  $g$  decreases when taking marginals as well as when passing from the matrix ensemble to the induced joint eigenvalue density. Finally, we used the estimate of Proposition 1.1. This completes the proof of Theorem 1.1.

## Appendix A: Proof of Proposition 2.1

To ease the reading, we here recall Proposition 2.1.

**Proposition A.1.** *Let  $z = u + i\eta$  for some  $u \in [u_- + \epsilon, u_+ - \epsilon]$  and  $\eta > 0$ . Then, there exist a constant  $c_1$  and  $C_o, C > 0, c > 0$  depending on  $\epsilon$  only such that  $\forall \delta < c_1\epsilon$ ,*

$$\mathbb{P} \left( \sup_{u \in [u_- + \epsilon, u_+ - \epsilon]} |m_N(z) - m_{\text{MP}}(z)| \geq \left( \delta + C_o \left| \frac{p}{N} - \gamma \right| \right) \right) \leq C e^{-c\delta\sqrt{N\eta}},$$

for any  $(\ln N)^4/N \leq \eta \leq 1$ . Furthermore, given  $\eta \geq (\ln N)^4/N$ , there exist constants  $c > 0, C > 0$  and  $K_o$  such that  $\forall \kappa \geq K_o$ ,

$$\mathbb{P} \left( \sup_{|x| > (\epsilon/200)^2, y \geq \eta} |m_N(x + iy)| \geq \kappa \right) \leq C e^{-c\sqrt{\kappa N\eta}}.$$

**Proof.** The proof follows closely that of Theorems 3.1, 4.1 and 4.6 in [13]. We only prove here the first statement. The second one is simple adaptation of the latter statement and ideas given in the proof of Theorem 4.6 in [13] (see also the proof of Theorem 2.1 in [12]).

Using a discretization scheme with step  $\delta/4\eta^2$  (see [13]), it is possible to show that it is actually enough to prove that

$$\mathbb{P} \left( |m_N(z) - m_{\text{MP}}(z)| \geq \left( \delta + C_o \left| \frac{p}{N} - \gamma \right| \right) \right) \leq C e^{-c\delta\sqrt{N\eta}},$$

for any  $(\ln N)^4/N \leq \eta \leq 1$  and for a given  $z = u + i\eta$  where  $u \in [u_- + \epsilon, u_+ - \epsilon]$ . Denote by  $C_k$  the  $k$ th column of  $W/\sqrt{N}$ ,  $C_k := \frac{1}{\sqrt{N}}(W)_k$ , then one has that:

$$H_N = \sum_{k=1}^p C_k C_k^*. \tag{A.1}$$

We define also  $\mathbf{R}_N(z) = (H_N - zI)^{-1}$  and for any integer  $k = 1, \dots, p$ ,  $\mathbf{R}_N^{(k)}(z) := (H_N - C_k C_k^* - zI)^{-1}$ . Then  $m_N(z) = \frac{1}{N} \text{Tr} \mathbf{R}_N(z)$  and one has that

$$1 + z m_N(z) = \frac{p}{N} - \frac{1}{N} \sum_{k=1}^p \frac{1}{1 + C_k^* \mathbf{R}_N^{(k)}(z) C_k}. \tag{A.2}$$

Equation (A.2) simply follows from (A.1) and the identity  $\mathbf{R}_N(z)(H_N - zI) = I$ . In addition we denote by  $y_1^{(k)} \geq y_2^{(k)} \geq \dots \geq y_N^{(k)}$  the ordered eigenvalues of  $H_N - C_k C_k^*$  and set  $\mu_N^{(k)} := \frac{1}{N} \sum_{i=1}^N \delta_{y_i^{(k)}}$ . By the well-known interlacing property of eigenvalues, for any  $k = 1, \dots, p$ ,  $y_1 \geq y_1^{(k)} \geq y_2 \geq y_2^{(k)} \geq \dots \geq y_N \geq y_N^{(k)}$ . Denote by  $F_N$  (resp.  $F_N^{(k)}$ ) the p.d.f. of the spectral measure  $\mu_N$  (resp.  $\mu_N^{(k)}$ ). Then one has that

$$|NF_N(x) - NF_N^{(k)}(x)| \leq 1 \quad \forall x \in \mathbb{R}. \tag{A.3}$$

Last call  $v_i^{(k)}$ ,  $i = 1, \dots, N$ , a set of orthonormal eigenvectors associated to the ordered eigenvalues  $y_i^{(k)}$  and define

$$\xi_i^{(k)} := |\langle v_i^{(k)}, \sqrt{N}C_k \rangle|^2.$$

From (A.2), one gets that

$$1 + zm_N(z) = \frac{p}{N} - \frac{1}{N} \sum_{k=1}^p \frac{1}{1 + (1/N) \sum_{i=1}^N \xi_i^{(k)} / (y_i^{(k)} - z)}. \quad (\text{A.4})$$

The latter formula (A.4) is the counterpart of formula (2.6) in [13]. The proof of [13] can be summarized into 2 basic ideas.

First the random vectors  $C_k$  are centered with i.i.d. entries and they are independent of  $\mathbf{R}_N^{(k)}$ . It is known that the random variables  $C_k^* \mathbf{R}_N^{(k)} C_k$  concentrate around their means which is given by  $\mathbb{E}|W_{11}|^2 \text{Tr} \mathbf{R}_N^{(k)} / N$ . The speed of concentration is explicit under assumption  $(\mathbf{A}_1)$  in Lemma 4.2 in [13] for Wigner matrices. The extension to Wishart matrices states as follows.

**Lemma A.1.** *Set  $X^{(k)} = X = \frac{1}{N} \sum_{k=1}^p \frac{\sigma^{-2} \xi_i^{(k)} - 1}{y_i^{(k)} - z}$  where  $z = u + i\eta$ ,  $u \in [u_- + \epsilon, u_+ - \epsilon]$ . Then there exists a positive constant  $c$  (depending on  $\epsilon$ ) so that for every  $\delta > 0$  we have*

$$\mathbb{P}[|X| \geq \delta] \leq 5e^{-c \min\{\delta \sqrt{N\eta/\gamma}, \delta^2 N\eta/\gamma\}}$$

if  $N\eta \geq (\ln N)^2$  and  $N$  is sufficiently large (independently of  $\delta$ ).

The proof of Lemma A.1 is postponed to the end of Appendix A.

**Remark A.1.** *Lemma A.1 is established under the assumption that the real and imaginary parts of the components of  $C_k$  are i.i.d. (in addition to the gaussian decay assumption  $(\mathbf{A}_1)$ ). This is the reason for our assumption  $(H_3)$ .*

One can then use (A.3) to show that for any  $k = 1, \dots, p$ ,  $|\text{Tr} \mathbf{R}_N^{(k)} - \text{Tr} \mathbf{R}_N| \leq \eta^{-1}$ . By Lemma A.1,  $|X^{(k)}| \leq \delta$ ,  $\forall k = 1, \dots, p$  in (A.4) with high probability. A bootstrap argument exposed in [11], Section 2, yields that with high probability, the Stieltjes transform  $m_N(z)$  satisfies

$$1 + zm_N(z) = \frac{p}{N} - \frac{p}{N} \frac{1}{1 + \sigma^2 m_N(z)} + \Delta, \quad (\text{A.5})$$

where  $|\Delta| \leq C'\delta$  for some  $C' > 0$  is a small error term and  $\sigma^2 = \mathbb{E}|W_{11}|^2 = 1/4$ .

The second basic idea is the stability of the equation (A.5). The equation

$$1 + zm(z) = \frac{p}{N} - \frac{p}{N} \frac{1}{1 + \sigma^2 m(z)} \quad (\text{A.6})$$

admits a unique solution satisfying  $\text{Im}(m(z)) > 0$  whenever  $\text{Im}(z) > 0$ . This solution is  $m_{MP,p/N}(z)$ , that is the Stieltjes transform of the Marčenko–Pastur distribution with parameter  $p/N$ . Now, the stability of equation (A.5) implies that there exists a constant  $C$  such that for any  $z \in \{z = u + i\eta, u \in [u_- + \epsilon, u_+ - \epsilon], (\ln N)^4/N \leq \eta \leq 1\}$ ,

$$|m_N(z) - m_{MP,p/N}(z)| \leq C\Delta.$$

The constant  $C$  here depends on  $\epsilon$  only. Now, as  $|z - u_{\pm}| \geq \epsilon$  and  $1/|z| = O(1)$ , there exists  $C_o > 0$  depending on  $\epsilon$  only such that

$$|m_{MP,p/N}(z) - m^{\text{MP}}(z)| \leq C|p/N - \gamma|.$$

This finishes the proof of Proposition 2.1 provided we show Lemma A.1. □

**Proof of Lemma A.1.** Define for  $n \geq 1$  the intervals  $I_n = [u - 2^{n-1}\eta, u + 2^{n-1}\eta]$  and let  $M$  and  $K_0$  be sufficiently large numbers. We have  $[-K_0, K_0] \subset I_{n_0}$  with  $n_0 = C_1 \ln(K_0/\eta) \leq C_2 \ln(NK_0)$  for some constants  $C_1, C_2 > 0$ . Denote by  $A$  the event:

$$A = \left\{ \max_{n \leq n_0} \frac{\mathcal{N}(I_n)}{\sqrt{Np}|I_n|} \geq M \right\} \cup \{\max y_i \geq K_0\},$$

where for a given interval  $I$ , we denote by  $|I|$  its length and set  $\mathcal{N}(I) = \#\{i = 1, \dots, N | y_i \in I\}$ . Let then  $\mathbb{P}_k$  denote the probability w.r.t.  $C_k$ . One has that

$$P[|X^{(k)}| \geq \delta] \leq \mathbb{E}(1_{A^c} \mathbb{P}_k[|X| \geq \delta]) + P(A).$$

The first term can be handled as in [13]. It is proved in [13], Proposition 4.5, under assumption  $(\mathbf{A}_1)$ , and if  $\eta \geq 1/N$ , and for sufficiently large  $M$  and  $K_0$ , that there is a positive constant  $c = c(M, K_0)$  such that for any  $\delta > 0$

$$\mathbb{E}(1_{A^c} \mathbb{P}_k(|X^{(k)}| \geq \delta)) \leq 4e^{-c \min\{\delta\sqrt{N\eta}/\sqrt{p/N}, \delta^2 N^2 \eta/p\}}.$$

The above estimate requires again the full assumption  $(H_3)$ .

We now turn to the estimate of  $\mathbb{P}(A)$ . This is the only part which has to be modified to consider sample covariance matrices. For this result, we need that  $(\ln N)^2/N \leq \eta \leq 1$ . Then for sufficiently large  $M$  and  $K_0$  there are positive constants  $c, C$  such that for all  $N \geq 2$ ,

$$\mathbb{P}(A) \leq Ce^{-c\sqrt{M\eta\sqrt{Np}}}. \quad (\text{A.7})$$

To prove (A.7), we first recall Lemma 7.3 of [12]. Let  $Y$  be a  $N \times p$ ,  $p \geq N$ , random matrix with i.i.d. centered entries with variance 1. If the entries of  $Y$  also satisfy assumption  $(\mathbf{A}_1)$ , there exists a positive constant  $c$  such that for  $C > 0$  large enough

$$\mathbb{P}(y_{\max}(YY^*/p) \geq C) \leq e^{-cCp}.$$

Thus we only have to consider

$$\mathbb{P}\left(\left\{ \max_{n \leq n_0} \frac{\mathcal{N}(I_n)}{\sqrt{Np}|I_n|} \geq M \right\}\right).$$

To this aim we consider an interval  $I$  of length  $|I| = \alpha$  for some  $\alpha \geq \eta$  and call  $u$  its midpoint. In the following we set  $z = u + i\alpha$  and assume that  $|z| \geq \varepsilon$  for some  $\varepsilon > 0$  small. Let  $0 < \vartheta_o < 1/8$  be given. For each  $k = 1, 2, \dots, N$ , we define the events

$$B^{(k)} := \sum_{m, y_m \in I} \xi_m^{(k)} \leq \vartheta_o(\mathcal{N}(I) - 1),$$

and set  $B = \bigcup_k B^{(k)}$ . As the eigenvalues of  $H_N$  and  $H_N^{(k)}$  are interlaced, at least  $\mathcal{N}(I) - 1$  eigenvalues of  $H_N^{(k)}$  lie in  $I$ . It is proved in Lemma 4.7 of [13] that

$$\exists c > 0, \quad \mathbb{P}_k(B^{(k)}) \leq e^{-c\sqrt{\mathcal{N}(I)-1}}.$$

From this, we deduce that there exists  $C > 0$  such that if  $M$  is large enough,

$$\mathbb{P}(B \cap \{\mathcal{N}(I) \geq M\alpha\sqrt{Np}\}) \leq Ce^{-c\sqrt{M\alpha\sqrt{Np}}}.$$

Then on  $B^c$  we have for some constant  $C_2 > 0$

$$\begin{aligned} \mathcal{N}(I) &\leq C_2 N \alpha \operatorname{Im}(m_N(u + i\alpha)) \\ &= C_2 N \alpha \operatorname{Im}\left(\frac{p/N - 1}{z} - \frac{1}{N} \sum_{k=1}^p \frac{1}{z + zC_K^* \mathbf{R}_N^{(k)} C_k}\right) \end{aligned}$$

$$\leq C_2 N \alpha \operatorname{Im} \left( -\frac{1}{N} \sum_{k=1}^p \frac{1}{z + z C_K^* \mathbf{R}_N^{(k)} C_k} \right) \quad (\text{A.8})$$

$$\leq \frac{4C_2 N p \alpha^2}{\varepsilon \vartheta_o \mathcal{N}(I)}. \quad (\text{A.9})$$

In (A.8) we used that  $\operatorname{Im}(1/z) < 0$  and  $p/N \geq 1$ . To derive (A.9), we used the fact that

$$\operatorname{Im} \left( -\frac{1}{z + z C_K^* \mathbf{R}_N^{(k)} C_k} \right) \leq \frac{1}{|z| |1 + C_K^* \mathbf{R}_N^{(k)} C_k|} \leq \frac{1}{|z| \operatorname{Im}(C_K^* \mathbf{R}_N^{(k)} C_k)},$$

with  $\frac{1}{|z|} \leq \varepsilon^{-1}$  and we estimated from below

$$\sum_{m=1}^N \frac{\xi_m^k}{|y_m^{(k)} - z|^2} \geq \vartheta_o \mathcal{N}(I) / (4\alpha^2).$$

This now implies that  $\mathcal{N}(I) \leq \alpha \sqrt{4C_2 N p / \vartheta_o \varepsilon}$  on  $B^c$ . This yields the desired result: one simply chooses  $M$  large enough.  $\square$

## Appendix B: Asymptotics of Bessel functions

We use two types of asymptotics for Bessel functions.

The first one deals with is well-known asymptotics of Bessel functions with bounded order and large argument (see [21] e.g.). They are used in Section 2.3.

Assume that  $\nu$  is bounded. Then, for large  $|z|$ ,  $z \in \mathbb{C}$

$$\begin{aligned} I_\nu(z) &= \frac{1}{\sqrt{2\pi z}} (e^z + e^{-z+(v+1/2)i\pi}) (1 + O(1/z)), \quad -\frac{\pi}{2} < \operatorname{Arg} z < \frac{3\pi}{2}; \\ K_\nu(z) &= \frac{\sqrt{\pi}}{\sqrt{2z}} e^{-z} (1 + O(1/z)), \quad |\operatorname{Arg} z| \leq \frac{3\pi}{2}. \end{aligned} \quad (\text{B.1})$$

We also make use in Section 2.4 of asymptotics of Bessel functions with large order and large argument. Abramowitz and Stegun [1] (p. 378) give the following uniform asymptotic expansion of modified Bessel functions of large order:

$$\begin{aligned} I_\nu(\nu z) &= \frac{e^{\nu\varphi}}{\sqrt{2\pi\nu(1+z^2)^{1/4}} \left( 1 + \sum_{k=1}^{\infty} \frac{u_k(t)}{\nu^k} \right)}, \\ K_\nu(\nu z) &= \frac{\sqrt{\pi} e^{-\nu\varphi}}{\sqrt{2\nu(1+z^2)^{1/4}} \left( 1 + \sum_{k=1}^{\infty} \frac{(-1)^k u_k(t)}{\nu^k} \right)}, \\ t &= \frac{1}{\sqrt{1+z^2}}, \quad \varphi = \sqrt{1+z^2} + \ln \left( \frac{z}{1 + \sqrt{1+z^2}} \right), \quad |\arg(z)| \leq \frac{\pi}{2} - \epsilon, \end{aligned} \quad (\text{B.2})$$

where  $u_k(t) = t^k v_k(t)$  for some polynomial  $v_k$ .

## References

- [1] M. Abramowitz and I. Stegun. *Handbook of Mathematical Functions With Formulas, Graphs and Mathematical Tables*. National Bureau of Standards Applied Mathematics Series 55. For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C., 1964. MR0167642

- [2] Z. D. Bai. Convergence rate of expected spectral distributions of large random matrices. II Sample covariance matrices. *Ann. Probab.* **21** (1993) 649–672. [MR1217560](#)
- [3] Z. D. Bai, B. Miao and J. Tsay. Remarks on the convergence rate of the spectral distributions of Wigner matrices. *J. Theoret. Probab.* **12** (1999) 301–311. [MR1684746](#)
- [4] G. Ben Arous and S. Péché. Universality of local eigenvalue statistics for some sample covariance matrices. *Comm. Pure Appl. Math.* **LVIII** (2005) 1–42. [MR2162782](#)
- [5] E. Brézin and S. Hikami. Spectral form factor in random matrix theory. *Phys. Rev. E* **55** (1997) 4067–4083. [MR1449379](#)
- [6] E. Brézin and S. Hikami. Correlations of nearby levels induced by a random potential. *Nucl. Phys. B* **479** (1996) 697–706. [MR1418841](#)
- [7] P. Deift. *Orthogonal Polynomials and Random Matrices: A Riemann–Hilbert Approach*. *Courant Lecture Notes in Mathematics* **3**. American Mathematical Society, Providence, RI, 1999. [MR1677884](#)
- [8] P. Deift, T. Kriecherbauer, K. T.-R. McLaughlin, S. Venakides and X. Zhou. Uniform asymptotics for polynomials orthogonal with respect to varying exponential weights and applications to universality questions in random matrix theory. *Comm. Pure Appl. Math.* **52** (1999) 1335–1425. [MR1702716](#)
- [9] P. Deift, T. Kriecherbauer, K. T.-R. McLaughlin, S. Venakides and X. Zhou. Strong asymptotics of orthogonal polynomials with respect to exponential weights. *Comm. Pure Appl. Math.* **52** (1999) 1491–1552. [MR1711036](#)
- [10] F. J. Dyson. A Brownian-motion model for the eigenvalues of a random matrix. *J. Math. Phys.* **3** (1962) 1191–1198. [MR0148397](#)
- [11] L. Erdős, B. Schlein and H.-T. Yau. Local semicircle law and complete delocalization for Wigner random matrices. *Commun. Math. Phys.* **287** (2009) 641–655. [MR2481753](#)
- [12] L. Erdős, B. Schlein and H.-T. Yau. Semicircle law on short scales and delocalization of eigenvectors for Wigner random matrices. *Ann. Probab.* **37** (2009) 815–852. [MR2537522](#)
- [13] L. Erdős, B. Schlein and H.-T. Yau. Wegner estimate and level repulsion for Wigner random matrices. *Int. Math. Res. Notices* **2010** (2010) 436–479. [MR2587574](#)
- [14] L. Erdős, S. Péché, J. Ramirez, B. Schlein and H.-T. Yau. Bulk universality for Wigner matrices. *Comm. Pure Appl. Math.* **63** (2010) 895–925. [MR2662426](#)
- [15] L. Erdős, J. Ramirez, B. Schlein, T. Tao, V. Vu and H.-T. Yau. Bulk universality for Wigner Hermitian matrices with subexponential decay. *Math. Research Letters* **17** (2010) 667–674. [MR2661171](#)
- [16] A. Guionnet and O. Zeitouni. Concentration of the spectral measure for large random matrices. *Electron. Comm. Probab.* **5** (2000) 119–136. [MR1781846](#)
- [17] K. Johansson. Universality of the local spacing distribution in certain ensembles of Hermitian Wigner matrices. *Commun. Math. Phys.* **215** (2001) 683–705. [MR1810949](#)
- [18] V. A. Marčenko and L. Pastur. The distribution of eigenvalues for some sets of random matrices. *Math. Sb.* **72** (1967) 507–536. [MR208649](#)
- [19] M. L. Mehta. *Random Matrices*. Academic Press, New York, 1991. [MR1083764](#)
- [20] J. W. Silverstein. Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices. *J. Multivariate Anal.* **55** (1995) 331–339. [MR1370408](#)
- [21] F. Olver. *Asymptotics and Special Functions*. *Computer Science and Applied Mathematics*. Academic Press, New York–London, 1974. [MR0435697](#)
- [22] T. Tao and V. Vu. Random matrices: Universality of local eigenvalue statistics. Preprint. Available at [arxiv:0906.0510](#).
- [23] T. Tao and V. Vu. Random covariance matrices: Universality of local statistics of eigenvalues. Preprint. Available at [arXiv:0912.0966](#).