Probab. Theory Relat. Fields (2016) 165:117–161
DOI 10.1007/s00440-015-0628-6

Fluctuations at the edges of the spectrum of the full rank deformed GUE

Mireille Capitaine[1] · Sandrine Péché[2]

Received: 12 February 2014 / Revised: 2 February 2015 / Published online: 10 April 2015

Abstract We consider a full rank deformation of the GUE $\frac{W_N}{\sqrt{N}} + A_N$ where A_N is a full rank Hermitian matrix of size N and W_N is a GUE. The empirical eigenvalue distribution μ_{A_N} of A_N converges to a probability distribution ν. We identify all the possible limiting eigenvalue statistics at the edges of the spectrum, including outliers, edges and merging points of connected components of the limiting spectrum. The results are stated in terms of a deterministic equivalent of the empirical eigenvalue distribution of $W_N + A_N$, namely the free convolution of the semi-circle distribution and the empirical eigenvalues distribution of A_N.

Keywords Random matrices · Free probability · Deformed GUE · Asymptotic spectrum · Fluctuations

Mathematics Subject Classification 15B52 · 60B20 · 46L54 · 15A18 · 60F05

✉ Mireille Capitaine
mireille.capitaine@math.univ-toulouse.fr

Sandrine Péché
sandrine.peche@math.univ-paris-diderot.fr

1 Institut de Mathématiques de Toulouse, CNRS, Université Paul Sabatier, 118, Route de Narbonne, 31062 Toulouse, Cedex 9, France

2 LPMA, Université Paris Diderot, 5 Rue Thomas Mann, 75013 Paris, France

1 Introduction and results

1.1 Motivations

Enormous progress has been accomplished in the very recent years in the study of asymptotic spectral properties of large random matrices. A Hermitian Wigner random matrix is a $N \times N$ matrix $W_N = \frac{1}{\sqrt{N}}(W_{ij})_{i,j=1}^N$,with i.i.d. entries off the diagonal $W_{ij}, i < j$ (modulo the symmetry assumption) and independent diagonal real entries. The entries are standardized to be centered and of variance σ^2. The asymptotic local properties of the spectrum of Wigner random matrices are now quite well understood thanks to the fantastic work of Erdös–Schlein–Yau (see [13,14] and references therein) and Tao–Vu [27]. In particular, it is known (assuming that the matrix elements admit enough moments) that the fluctuations of eigenvalues in the bulk or at the edges of the spectrum are universal. In particular, they coincide with those identified for a Gaussian (GUE) matrix with variance σ^2. In other words, the limiting asymptotic spectral properties of a Wigner matrix in the large N limit do not depend on the detail of the distribution of the matrix elements W_{ij}, $1 \leq i, j \leq N$.
In this article, we are interested in deformed random matrix ensembles. A *deformation* of a standard random matrix can be more or less understood as the modification of the distribution of some of the entries of a Wigner matrix. The set of possible deformations is non exhaustive (one can force some of the entries to be zero such as for sparse matrices) but we here restrict to some additive deformations. More precisely, we consider a matrix A_N of size N, which is deterministic. Our study could be extended to the case where it is random but we do not wish to pursue this direction here. We consider the deformed matrices

$$\frac{W_N}{\sqrt{N}} + A_N,$$

where W_N is a standard Wigner matrix. The question is to understand the asymptotic properties of the eigenvalues and eigenvectors of the deformed matrix, knowing that of A_N and $\frac{W_N}{\sqrt{N}}$. Such ensembles have first been introduced by [10], and [16] when W_N is a GUE.

In the case where A_N is a fixed rank (independent of the size N) matrix, the asymptotic properties of the spectrum are quite clear. Finite rank perturbed ensembles have first been considered in [4] (see also [6] and [22]). First, the global properties of the spectrum are not impacted by A_N. Indeed, denoting by $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$ the ordered eigenvalues of $\frac{W_N}{\sqrt{N}} + A_N$, the empirical eigenvalue distribution $\mu_N := \frac{1}{N}\sum_{i=1}^N \delta_{\lambda_i}$ still converges (as in the case where $A_N = 0$) to the semi-circle distribution with density $\sigma_{sc}(x) = \frac{1}{2\pi\sigma^2}\sqrt{4\sigma^2 - x^2}\mathbb{1}_{|x|\leq 2\sigma}$. The asymptotic local eigenvalue statistics of eigenvalues in the bulk of the spectrum are also unchanged by the deformation matrix A_N. Only the local behavior of the spectrum at the edges may be impacted by the deformation A_N, as we now explain. The deformation A_N may cause some eigenvalues to separate from the bulk of the spectrum. Each eigenvalue of A_N greater than σ is called a *spike*. To each spike θ_i of A_N such that $|\theta_i| > \sigma$ (if it exists) there corresponds an eigenvalue λ_i satisfying

$$\lambda_i \to \left(\theta_i + \frac{\sigma^2}{\theta_i}\right)$$

a.s. Such eigenvalues λ_i outside the support of the semi-circle distribution are called *outliers*. Interestingly, [11] and then [23,24] have proved that the fluctuations of spikes are not universal in general. More precisely

$$\sqrt{N}\left(\lambda_i - \left(\theta_i + \frac{\sigma^2}{\theta_i}\right)\right) \xrightarrow{d} \mu,$$

where the distribution μ may depend explicitly on the distribution of the matrix elements W_{ij}. It can be shown that eigenvectors of the matrix A_N play a fundamental role in the universality/non universality of the deformation matrix A_N. On the contrary, when there is no spike, the limiting distribution of extreme eigenvalues is the same as in the non deformed case. In particular, extreme eigenvalues stick to the bulk of the spectrum. The scale of their fluctuations is $N^{-2/3}$ and the limiting distribution of the largest (and smallest) eigenvalues is the Tracy–Widom distribution, provided the matrix elements W_{ij} admit enough moments. A complete study of such deformed ensembles has been achieved in [17] and [18] and we refer the reader to these articles for a complete state of the art in finite rank deformations of Wigner matrices.

The study of deformed ensembles extends to the case where the matrix A_N has low rank $r_N << N, r_N \to \infty$ (see [22] e.g.) or full rank i.e. when $r_N = O(N)$. In this case, it is natural to assume that the empirical eigenvalue distribution of A_N has a weak limit as $N \to \infty$, which is possibly δ_0. Denote by $y_1 \geq y_2 \geq \cdots \geq y_N$ the ordered eigenvalues of A_N. Let $\mu_{A_N} = \frac{1}{N}\sum_{i=1}^N \delta_{y_i}$. We assume the norms of $(A_N)_N$ are uniformly bounded and that there exists a probability distribution ν on $\mathbb{R}$ such that

$$\mu_{A_N} \underset{N\to\infty}{\overset{w}{\to}} \nu.$$

Let us diagonalize A_N through $A_N = V\mathrm{diag}(y_1, \ldots, y_N)V^*$. Roughly speaking the deformed model is now understood in the sense that $\frac{W_N}{\sqrt{N}} + A_N$ is a "small" perturbation of the matrix $\frac{W_N}{\sqrt{N}} + VA_0V^*$ where A_0 would be a diagonal matrix made up with quantiles of the probability ν. The asymptotic global behavior of the spectrum is well-known in this case. Indeed, let μ_N be the empirical eigenvalue distribution of $\frac{W_N}{\sqrt{N}} + A_N$. Its Stieltjes transform is

$$m_N(z) := \int \frac{1}{z-y} d\mu_N(y), \quad \mathrm{Im} z \neq 0.$$

According to [3,30], m_N converges as $N \to \infty$ to the Stieltjes transform m_τ of a probability distribution τ, called the free convolution of ν and the semi-circle distribution. This probability distribution τ is uniquely characterized by a fixed point equation satisfied by m_τ, as we review in Sect. 2; it has a density p. We emphasize that the support of the probability distribution τ may have distinct connected components, depending on ν.

The question of the asymptotic behavior of extreme eigenvalues naturally arises in this setting also. This question has been much less investigated actually. So far, only the case where W_N is a GUE has been investigated.
In [26], the author considers the case where μ_{A_N} concentrates quite fast to the measure ν. In particular, there are no spikes. When W_N is a GUE, she investigates the local edge regime which deals with the behavior of the eigenvalues near any extremity point u_0 of a connected component of $\text{supp}(\tau)$. More precisely let some $\epsilon > 0$ be given and assume that either

$$\begin{aligned} & p(u) > 0, \quad \forall u \in]u_0; u_0 + \epsilon[, \text{ and } p(u) = 0, \quad \forall u \in]u_0 - \epsilon; u_0], && (1) \\ \text{or } & p(u) > 0, \quad \forall u \in]u_0 - \epsilon; u_0[, \text{ and } p(u) = 0, \quad \forall u \in [u_0; u_0 + \epsilon[. && (2) \end{aligned}$$

Shcherbina [26] makes a technical assumption on the uniform convergence of the Stieltjes transform of μ_{A_N} to m_ν:

$$\sup_{z \in K} |m_{\mu_{A_N}}(z) - m_\nu(z)| \leq N^{-2/3-\epsilon}, \tag{3}$$

where K is some compact subset of the complex plane at a positive distance of the support of ν. This is a rather strong assumption on the rate of convergence of μ_{A_N} to ν. Shcherbina [26] proves that the joint distribution of the largest (or smallest) eigenvalues converging to u_0 have universal asymptotic behavior, characterized by the famous Tracy–Widom distribution. We note that Shcherbina [25] also investigates the asymptotic spacing distribution of eigenvalues in the bulk of the spectrum. The same behavior as for non deformed ensemble is obtained (and described by the sine kernel). The extension to a non Gaussian matrix W_N has recently been obtained by O'Rourke and Vu [21] in the case where A_N is diagonal.
In [2] and [1], the authors consider the case where $\mu_{A_N} = \nu$ is a finite combination of Dirac delta masses. They identify different possible limiting statistics at the edges of the support of τ, after suitable normalization of the eigenvalues. If u_0 is a point such that $p(u) = 0,\ u_0 - \epsilon \leq u \leq u_0,\ p(u) > 0, u_0 < u \leq u_0 + \epsilon$ for some $\epsilon > 0$, the asymptotic distribution of eigenvalues close to u_0 is the Tracy–Widom distribution. The authors also consider the case where u_0 is a point where two connected components of $\text{supp}(\tau)$ merge so that $p(u) > 0, \forall u \in (u_0 - \epsilon, u_0 + \epsilon) \setminus \{u_0\}$ and $p(u_0) = 0$. In this case, the limiting eigenvalue statistics are described by the so-called Pearcey kernel (whose definition is reviewed hereafter).

In both cases, a strong assumption is made on the rate of convergence of μ_{A_N} to ν. We here remove this assumption. We identify all the possible limiting eigenvalue statistics at the edges of the spectrum of the deformed GUE, namely at a spike, at the edge of a connected component of the support or at a point where two connected components merge. We emphasize that we do not make any assumptions on the rate of convergence of μ_{A_N} to ν. To state our results, we use a deterministic equivalent of the empirical eigenvalue distribution of the deformed GUE. This equivalent is the free convolution of the semi-circle distribution and μ_{A_N}.
The choice of the deformed GUE is motivated by the fact that all eigenvalues statistics can be explicitly computed for this ensemble of deformed random matrices.

We expect that one can extend these results to full rank deformations of an arbitrary Wigner matrix, as in the fixed rank case (with universal or non universal results). We intend to consider this general case in a forthcoming paper. The techniques needed are completely different.

Our results shall be compared to [5]. Therein the authors consider a (random) additive perturbation of a complex compound Wishart matrix and establish universal results by considering mobile edges. Their model is quite comparable to the one considered in the present article. We use the same strategy for the proof as that developed in [5], except that we provide a free probabilistic interpretation of the mobile edges. Our main contribution with respect to [5] is thus to provide a free probabilistic setting that allows to fully describe the asymptotic distribution of extreme eigenvalues, without hypotheses on rate of convergence of μ_{A_N} to ν.

1.2 Model and results

We consider the following deformed GUE ensemble

$$M_N = X_N + A_N,$$

where

(H_1) $X_N = \frac{1}{\sqrt{N}} W_N$ where W_N is a $N \times N$ GUE matrix: the random variables $(W_N)_{ii}, \sqrt{2}(\Re(W_N)_{ij})_{i<j}, \sqrt{2}(\mathrm{Im}(W_N)_{ij})_{i<j}$ are i.i.d., with gaussian distribution of variance 1 and mean 0.

(H_2) A_N is a deterministic Hermitian matrix whose eigenvalues $y_i = y_i(N)$, $1 \leq i \leq N$, are such that the spectral measure $\mu_{A_N} := \frac{1}{N}\sum_{i=1}^N \delta_{y_i}$ converges weakly to some probability measure ν with compact support. We assume that

$$\forall t \in \text{ supp } (\nu), \ \lim_{\epsilon \to 0} \int \frac{d\nu(x)}{(t-x)^2 + \epsilon^2} > 1, \tag{4}$$

where supp(ν) denotes the support of ν.

(H_3) We also assume that there exists a fixed integer $r \geq 0$ (independent from N) and an integer $0 \leq J \leq r$ such that the following holds. There are J fixed real numbers $\theta_1 > \cdots > \theta_J$ independent of N which are outside the support of ν and such that each θ_j is an eigenvalue of A_N with a fixed multiplicity k_j (with $\sum_{j=1}^J k_j = r$). The θ_j's are called the spikes or the spiked eigenvalues of A_N and we set

$$\Theta = \{\theta_j, \ 1 \leq j \leq J\}.$$

The remaining $N - r$ eigenvalues of A_N, denoted by $\beta_j(N)$, $j = 1, \ldots, N - r$, satisfy

$$\max_{1 \leq j \leq N-r} \mathrm{dist}(\beta_j(N), \mathrm{supp}(\nu)) \underset{N\to\infty}{\longrightarrow} 0.$$

Remark 1.1 The inequality (4) may not hold typically for a measure ν having a density which vanishes quite fast at some point of the support. This may be the case for instance, for the following measures

$$d\nu(x) := \frac{\alpha+1}{(1+A)^{\alpha+1}}(1-x)^{\alpha} 1_{[-A,1]}(x)dx;\ \ \alpha \geq 3,$$

by suitably choosing A.

Remark 1.2 Note that we do not make any assumptions on the rate of convergence of μ_{A_N} to ν.

Denote by μ_{sc} the semicircle distribution whose density is given by

$$\frac{d\mu_{sc}}{dx}(x) = \frac{1}{2\pi}\sqrt{4-x^2}\,1\quad 1_{[-2,2]}(x). \tag{5}$$

According to [3], the spectral distribution of M_N weakly converges almost surely to the so-called *free convolution* $\mu_{sc} \boxplus \nu$ which has a continuous density p (see [8]). We recall some important facts about the free convolution with a semi-circular distribution in Sect. 2.

We are now in position to state our results. Let first consider a real number d which is a right edge of supp$(\mu_{sc} \boxplus \nu)$ that is which satisfies (2). Assume moreover that for any θ_j such that $\int \frac{d\nu(s)}{(\theta_i - s)^2} = 1$, we have $d \neq \theta_j + m_\nu(\theta_j)$. We show in Proposition 3.1 that for η small enough, for all large N, there exists a unique right edge d_N of supp$(\mu_{sc} \boxplus \mu_{A_N})$ in $]d-\eta; d+\eta[$. We derive the asymptotic distribution of eigenvalues in the vicinity of d_N. Before exposing our results, we need a few notations. Let $Ai(u)$ be the Airy function defined by

$$Ai(u) = \frac{1}{2\pi}\int e^{iua + i\frac{1}{3}a^3} da \tag{6}$$

where the contour is from $\infty e^{5i\pi/6}$ to $\infty e^{i\pi/6}$. The *Airy kernel* (see e.g. [28]) is then given by

$$\mathbf{A}(u,v) = \frac{Ai(u)Ai'(v) - Ai'(u)Ai(v)}{u-v} = \int_0^\infty Ai(u+z)Ai(z+v)dz. \tag{7}$$

Let $\mathbf{A}_x$ be the operator acting on $L^2((x,\infty))$ with kernel $\mathbf{A}(u,v)$. The GUE Tracy–Widom distribution for the largest eigenvalue is [28]

$$F_0(x) = \det(1 - \mathbf{A}_x) = F_{GUE}(x). \tag{8}$$

We refer to [28] for the more complicated definition of the GUE distribution for the k largest eigenvalues ($k > 1$).

We first prove the following result. Let k be a given fixed integer. Let $\lambda_{max} \geq \lambda_{max-1} \geq \cdots \lambda_{max-k+1}$ denote the k largest of those eigenvalues of M_N converging to d.

Theorem 1.1 *There exists $\alpha > 0$ depending on d_N only such that the vector*

$$\frac{N^{2/3}}{\alpha}\left(\lambda_{max} - d_N, \lambda_{max-1} - d_N, \ldots, \lambda_{max-k+1} - d_N\right)$$

converges in distribution as $N \to \infty$ to the so-called Tracy–Widom GUE distribution for the k largest eigenvalues.

Remark 1.3 Condition (4) is necessary to obtain Tracy–Widom asymptotics at the edges of the spectrum. If condition (4) fails e.g. at the top edge of the spectrum, meaning that the density of ν vanishes too fast at the edge, the limiting eigenvalue statistics at the edge can be proved to be Gaussian.

We now turn to the behavior of outliers. Let θ_i be a spiked eigenvalue with multiplicity k_i, such that $\int \frac{1}{(\theta_i - x)^2} d\nu(x) < 1$. In [12], the authors prove that the spectrum of M_N exhibits k_i eigenvalues in a neighborhood of

$$\rho_{\theta_i} = \theta_i + \int \frac{d\nu(x)}{\theta_i - x}. \tag{9}$$

Note that such a result is obtained when the support of ν has a finite number of connected components. However this assumption can be easily relaxed (see Remark 2.2). In Proposition 3.4, we prove that for $\epsilon > 0$ small enough, for all large N, $\mathrm{supp}(\mu_{sc} \boxplus \mu_{A_N})$ has a unique connected component $[L_i(N); D_i(N)]$ inside $]\rho_{\theta_i} - \epsilon; \rho_{\theta_i} + \epsilon[$.

Define

$$\rho_N(\theta_i) = z + \frac{1}{N}\sum_{y_j \neq \theta_i} \frac{1}{\theta_i - y_j}. \tag{10}$$

It can be shown that for all large N, $\rho_N(\theta_i) \in [L_i(N); D_i(N)]$ and $\rho_N(\theta_i) = \frac{L_i(N)+D_i(N)}{2} + o\left(\frac{1}{\sqrt{N}}\right)$.

To define the limiting correlation function at an outlier, we consider for $k = 1, 2, \ldots,$ the distribution $G_k(\cdot)$ given by

$$G_k(x) = \frac{1}{Z_k}\int_{-\infty}^{x} \cdots \int_{-\infty}^{x} \prod_{1 \le i < j \le k} |\xi_i - \xi_j|^2 \cdot \prod_{i=1}^{k} e^{-\frac{1}{2}\xi_i^2} d\xi_1 \cdots d\xi_k. \tag{11}$$

In other words, G_k is the distribution of the *largest eigenvalue* of $k \times k$ GUE. It has been shown (see [19] or [3] e.g.) that

$$G_k(x) = \det(1 - \mathbf{H}_x^{(k)}), \tag{12}$$

where $\mathbf{H}_x^{(k)}$ is the operator acting on $L^2((x, \infty))$ defined by the Christoffel Darboux kernel of some rescaled Hermite polynomials satisfying the orthogonality relationship

$\int_{-\infty}^{\infty} p_m(x)p_n(x)e^{-\frac{1}{2}x^2}dx = \delta_{mn}$. We refer the reader to [4, Section 1.2.2] for a more complete statement of this fact.

Let us denote by λ_{max} the largest of the k_i outliers around $\rho_N(\theta_i)$.

Theorem 1.2 *There exists $c > 0$ depending on θ_i and ν only such that*

$$\lim_{N\to\infty} \mathbb{P}\left(\sqrt{N}c(\lambda_{max} - \rho_N(\theta_i) \leq x\right) = G_{k_i}(x).$$

We actually prove that the k_i outliers around $\rho_N(\theta_i)$ fluctuate as the eigenvalues of a $k_i \times k_i$ GUE.

Finally, we turn to the fluctuations in a neighborhood of an isolated point of vanishing density. Let $u_0 \in \mathbb{R}$ be such that $p(u_0) = 0$ and that there exists $\epsilon > 0$ such that, $\forall u \in]u_0 - \epsilon; u_0 + \epsilon[\setminus\{u_0\}, p(u) > 0$. Assume that for any θ_i such that $\int \frac{d\nu(s)}{(\theta_i - s)^2} = 1$, we have $\theta_i + m_\nu(\theta_i) \neq u_0$. Set $t_0 = \Psi_\nu^{-1}(u_0)$ where Ψ_ν is defined in Theorem 2.1 below. We make the following assumption:

(H_4) The equation $\int \frac{d\mu_{A_N}(x)}{(t-x)^2} - 1 = 0$ admits a unique solution $t \in \mathbb{C}$ in a neighborhood of t_0.

We prove in Proposition 3.3, that for η small enough, for all large N, there exists u_N in $]u_0 - \eta; u_0 + \eta[$ such that $p_N(u_N) = 0$ and $\forall u \in]u_0 - \eta; u_0 + \eta[\setminus\{u_N\}, p_N(u) > 0$, where p_N denotes the density of $\mu_{sc} \boxplus \mu_{A_N}$.
Last we derive the asymptotic behavior of eigenvalues at the vicinity of u_N. Consider the Pearcey kernel defined by

$$K_P(x, y) := \frac{1}{2i\pi} \int_{\Gamma_0} dt \int_{-i\infty}^{i\infty} ds e^{-t^4+xt+s^4-sy} \frac{1}{s-t}. \tag{13}$$

The contour Γ_0 is formed by two curves lying respectively to the right and left of 0: one goes from $\infty e^{i\frac{\pi}{4}}$ to $\infty e^{-i\frac{\pi}{4}}$ and the other from $-\infty e^{i\pm\frac{\pi}{4}}$ to $-\infty e^{-i\frac{\pi}{4}}$. See Fig. 1 below. The Pearcey distribution has been defined in [1,2,29]. Let k be a fixed integer and $f : \mathbb{R}^k \to R$ be a symmetric bounded function with compact support.

Theorem 1.3 *There exists $\kappa > 0$ such that*

$$\mathbb{E} \sum_{1\leq i_1<i_2<\cdots<i_k\leq N} f(\kappa N^{\frac{3}{4}}(\lambda_{i_1} - u_N), \kappa N^{\frac{3}{4}}(\lambda_{i_2} - u_N), \ldots, \kappa N^{\frac{3}{4}}(\lambda_{i_k} - u_N))$$
$$\to_{N\to\infty} \int_{\mathbb{R}^k} \frac{1}{k!} f(x_1, \ldots, x_k) \det(K_P(x_i, x_j))_{i,j=1}^k \prod_{i=1}^k dx_i.$$

The article is organized as follows. In Sect. 2, we review the fundamental properties of the free convolution that we later need in the proof. Section 3 gives fine estimates on the comparison of the support of the spectral distribution of M_N on the one hand and that of $\mu_{sc} \boxplus \nu$ on the other hand. These are the fundamental tools for the asymptotic analysis of eigenvalue statistics in Sect. 4. Therein the basic tool is a saddle point analysis of the correlation functions of the deformed GUE.

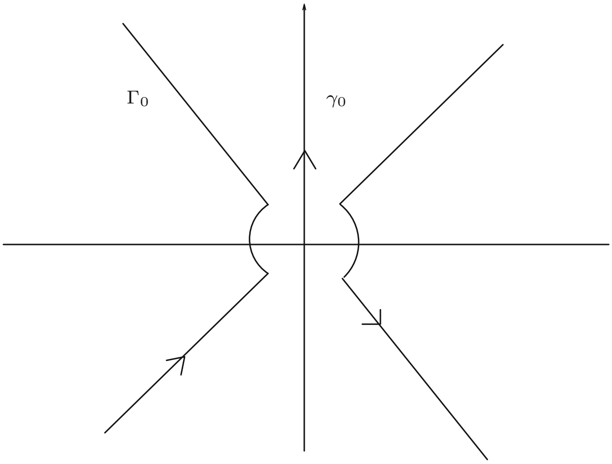

Fig. 1 Contours defining the Pearcey kernel

2 Free convolution by a semicircular distribution

2.1 The free convolution

We recall here an analytic definition of the free convolution of two probability measures. Let τ be a probability measure on $\mathbb{R}$. Its Stieltjes transform m_τ is defined by

$$m_\tau(z) := \int \frac{1}{z-y} d\tau(y).$$

m_τ is analytic on the complex upper half-plane $\mathbb{C}^+$. There exists a domain

$$D_{\alpha,\beta} = \{u + iv \in \mathbb{C}, |u| < \alpha v, \ \ v > \beta\}$$

on which m_τ is univalent. Let K_τ be its inverse function, defined on $m_\tau(D_{\alpha,\beta})$, and

$$R_\tau(z) = K_\tau(z) - \frac{1}{z}.$$

Definition 2.1 Given two probability measures τ and ν, there exists a unique probability measure λ such that

$$R_\lambda = R_\tau + R_\nu$$

on a domain where these functions are defined. The probability measure λ is called the free convolution of τ and ν and denoted by $\tau \boxplus \nu$.

We refer the reader to [15,20,32,33] for an introduction to free probability theory. The free convolution of probability measures has an important property, called subordination, which can be stated as follows: let τ and ν be two probability measures on $\mathbb{R}$;

there exists an analytic map $\omega : \mathbb{C}^+ \to \mathbb{C}^+$ such that $\omega(z)/z \to 1$ as $z \to \infty$ with $z \in D_{\alpha,\beta}$, for every such domain, and such that

$$\forall z \in \mathbb{C}^+, \quad m_{\tau \boxplus \nu}(z) = m_\nu(\omega(z)).$$

This phenomenon was first observed by Voiculescu under a genericity assumption in [31], and then proved in generality in [9, Theorem 3.1]. Later, a new proof of this result was given in [7], using a fixed point theorem for analytic self-maps of the upper half-plane.
In [8], Biane provides a deep study of the free convolution by a semicircular distribution, based on this subordination property.

2.2 The free convolution $\mu_{sc} \boxplus \nu$

We first recall here some of Biane's results that will be useful in this paper.
Let ν be a probability measure on $\mathbb{R}$. Biane [8] introduces the set

$$\Omega_\nu := \{u + iv \in \mathbb{C}^+, v > v_\nu(u)\},$$

where the function $v_\nu : \mathbb{R} \to \mathbb{R}^+$ is defined by

$$v_\nu(u) = \inf \left\{ v \geq 0, \int_{\mathbb{R}} \frac{d\nu(x)}{(u-x)^2 + v^2} \leq 1 \right\},$$

and proves the following

Proposition 2.1 [8] *The map*

$$H_\nu : z \longmapsto z + m_\nu(z)$$

is a homeomorphism from $\overline{\Omega_\nu}$ to $\mathbb{C}^+ \cup \mathbb{R}$ which is conformal from Ω_ν onto $\mathbb{C}^+$. Let $\omega_\nu : \mathbb{C}^+ \cup \mathbb{R} \to \overline{\Omega_\nu}$ be the inverse function of H_ν. One has,

$$\forall z \in \mathbb{C}^+, \quad m_{\mu_{sc} \boxplus \nu}(z) = m_\nu(\omega_\nu(z))$$

and then

$$\omega_\nu(z) = z - m_{\mu_{sc} \boxplus \nu}(z). \tag{14}$$

The previous results of [8] allows to conclude that $\mu_{sc} \boxplus \nu$ is absolutely continuous with respect to the Lebesgue measure and to obtain the following description of the support.

Theorem 2.1 [8] *Define $\Psi_\nu : \mathbb{R} \to \mathbb{R}$ by:*

$$\Psi_\nu(t) = H_\nu(t + i v_\nu(t)) = t + \int_{\mathbb{R}} \frac{(t-x)d\nu(x)}{(t-x)^2 + v_\nu(t)^2}.$$

Ψ_ν is a homeomorphism and, at the point $\Psi_\nu(t)$, the measure $\mu_{sc} \boxplus \nu$ has a density given by

$$p_\nu(\Psi_\nu(t)) = \frac{v_\nu(t)}{\pi}. \tag{15}$$

Define the set

$$U_\nu := \{t \in \mathbb{R}, v_\nu(t) > 0\}. \tag{16}$$

The support of the measure $\mu_{sc} \boxplus \nu$ is the image of the closure of the open set U_ν by the homeomorphism Ψ_ν. Ψ_ν is strictly increasing on U_ν.

Hence,

$$\mathbb{R}\backslash \mathrm{supp}(\mu_{sc} \boxplus \nu) = \Psi_\nu(\mathbb{R}\backslash \overline{U_\nu}).$$

One has $\Psi_\nu = H_\nu$ on $\mathbb{R}\backslash\overline{U_\nu}$ and $\Psi_\nu^{-1} = \omega_\nu$ on $\mathbb{R}\backslash \mathrm{supp}(\mu_{sc} \boxplus \nu)$. In particular, we have the following description of the complement of the support:

$$\mathbb{R}\backslash \mathrm{supp}(\mu_{sc} \boxplus \nu) = H_\nu(\mathbb{R}\backslash \overline{U_\nu}). \tag{17}$$

The following result will be useful later on.

Lemma 2.1 [8] *If t_0 is a point in the complement of the support of ν where two components of the set U_ν merge into one, then*

$$\int \frac{d\nu(x)}{(t_0 - x)^2} = 1,$$
$$\int \frac{d\nu(x)}{(t_0 - x)^3} = 0.$$

In [12], when ν is a compactly supported probability measure, the authors establish the following results.

Proposition 2.2 [12]

$$\overline{U_\nu} = \mathrm{supp}(\nu) \cup \left\{t \in \mathbb{R}\backslash \mathrm{supp}(\nu), \int_{\mathbb{R}} \frac{d\nu(x)}{(t-x)^2} \geq 1\right\}. \tag{18}$$

Each connected component of $\overline{U_\nu}$ contains at least one connected component of $\mathrm{supp}(\nu)$.

We also need the following additional basic results.

Lemma 2.2 *Let $]a; b[\subset \mathbb{R}\backslash\{U_\nu \cup supp(\nu)\}$. Then, Ψ_ν is strictly increasing on $]a; b[$.*

Proof Since $\forall t \in \mathbb{R}\backslash U_\nu,\ \ v_\nu(t) = 0$, we have $\Psi_\nu = H_\nu$ on $]a; b[$. Moreover $\forall t \in]a; b[,\ \ H_\nu^{'}(t) = 1 - \int \frac{d\nu(x)}{(t-x)^2} \geq 0$. The result readily follows since moreover Ψ_ν is one to one. □

Lemma 2.3 *If* $t \notin \text{supp}(\nu)$ *is such that there exists* $\delta > 0$ *such that*

$$]t-\delta; t[\subset U_\nu \text{ and } [t; t+\delta[\subset \mathbb{R}\backslash U_\nu. \tag{19}$$

Then, one has that

$$\text{(i)}: \int \frac{d\nu(x)}{(t-x)^2} = 1, \quad \text{and} \quad \text{(ii)}: \int \frac{d\nu(x)}{(t-x)^3} > 0.$$

If $t' \notin \text{supp}(\nu)$ *is such that there exists* $\delta > 0$ *such that*

$$]t'-\delta; t'] \subset \mathbb{R}\backslash U_\nu \quad \text{and} \quad]t'; t'+\delta[\subset U_\nu. \tag{20}$$

Then, one has that

$$\text{(iii)}: \int \frac{d\nu(x)}{(t'-x)^2} = 1 \quad \text{and} \quad \text{(iv)}: \int \frac{d\nu(x)}{(t'-x)^3} < 0.$$

Proof Since t and t' are in $\overline{U_\nu}\backslash U_\nu$, (18) readily implies (i) and (iii). Let us establish (ii). Let $\epsilon > 0$ be such that $]t-\epsilon; t+\epsilon[\subset \mathbb{R}\backslash\text{supp}(\nu)$. Set

$$f : \begin{array}{l}]t-\epsilon; t+\epsilon[\to \mathbb{R} \\ s \mapsto \displaystyle\int \frac{d\nu(x)}{(s-x)^2}. \end{array}$$

Note that $f''(s) = 6\int \frac{d\nu(x)}{(s-x)^4} > 0$ so that f' is strictly increasing on $]t-\epsilon; t+\epsilon[$. Therefore if $-f'(t) = 2\int \frac{d\nu(x)}{(t-x)^3} \leq 0$ then $f' > 0$ on $]t; t+\epsilon[$ and $\int \frac{d\nu(x)}{(s-x)^2} > 1$ for $s \in]t; t+\epsilon[$ which leads to a contradiction with (19). Similarly, one can prove (iv). □

Remark 2.1 In the rest of the article, since we deal with a measure ν satisfying (4), we have $\text{supp}(\nu) \subset U_\nu$.

2.3 The free convolution $\mu_{sc} \boxplus \mu_{A_N}$ and the localization of the spectrum of M_N

In [12], the authors prove that a precise localization of the spectrum of M_N can be described thanks to the support of the free convolution $\mu_{sc} \boxplus \mu_{A_N}$. In this section, we recall some of their results that we need afterwards.

Theorem 2.2 [12] *One has that* $\forall \epsilon > 0$,

$$\mathbb{P}\,(\textit{For all large } N, \text{Spect}(M_N) \subset \{x, \text{dist}(x, \text{supp}(\mu_{sc} \boxplus \mu_{A_N})) \leq \epsilon\}) = 1.$$

An outlier in the spectrum of M_N is an eigenvalue of M_N lying outside the support of $\mu_{sc} \boxplus \nu$. As we now explain, it is possible to describe outliers thanks to the support of $\mu_{sc} \boxplus \mu_{A_N}$.

Notations and definitions Throughout the rest of the article, we denote U_ν, H_ν, Ψ_ν, v_ν and p_ν by U, H, Ψ, v and p respectively. We also denote $U_{\mu_{A_N}}$, $H_{\mu_{A_N}}$, $\Psi_{\mu_{A_N}}$,

$v_{\mu_{A_N}}$, $p_{\mu_{A_N}}$ by U_N, H_N, Ψ_N, v_N and p_N respectively. Last, we define the probability measure $\hat{\nu}_N$ by

$$\hat{\nu}_N = \frac{1}{N-r} \sum_{i=1}^{N-r} \delta_{\beta_i(N)}.$$

It is easy to see that $\hat{\nu}_N$ weakly converges to ν. We define

$$\Theta_\nu = \Theta \cap (\mathbb{R} \backslash \overline{U}).$$

Furthermore, for any $\theta_j \in \Theta_\nu$, we set

$$\rho_{\theta_j} := H(\theta_j) = \theta_j + m_\nu(\theta_j). \tag{21}$$

Note that ρ_{θ_j} lies outside of the support of $\mu_{sc} \boxplus \nu$ according to (17). Define also

$$K_\nu(\theta_1, \ldots, \theta_J) := \text{supp}(\mu_{sc} \boxplus \nu) \bigcup \{\rho_{\theta_j},\ \theta_j \in \Theta_\nu\}. \tag{22}$$

In [12], the authors obtain moreover the following inclusion of the support of $\mu_{sc} \boxplus \mu_{A_N}$.

Theorem 2.3 *For any* $\epsilon > 0$,

$$\text{supp}(\mu_{sc} \boxplus \mu_{A_N}) \subset K_\nu(\theta_1, \ldots, \theta_J) + (-\epsilon, \epsilon),$$

when N *is large enough.*

In [12], the authors proved this theorem when the supp(ν) has a finite number of connected components. Nevertheless, it is still true in our more general setting as we prove in the following lines. We will use the following lemma in [12] which proof does not care about the number of connected components of the supports.

Lemma 2.4 [12] *For any* $\epsilon > 0$,

$$U_N \subset \{x, dist(x, \overline{U}) < \epsilon\} \cup \{x,\ \text{dist}(x, \Theta_\nu) < \epsilon\}, \tag{23}$$

for all large N.

Proof of Theorem 2.3 First, one can readily observe that if x satisfies dist$(x, \text{supp}(\nu)) \geq 1$ then $-m'_\nu(x) \leq 1$. This implies that the open set U is included in the compact set $\{x,\ \text{dist}(x, \text{supp}(\nu)) \leq 1\}$. Then we can choose K large enough such that $\{x, \text{dist}(x, \overline{U} \cup \Theta_\nu) \leq 1\} \subset [-K; K]$ and, since $\lim_{y \to \pm\infty} \Psi(y) = \pm\infty$ and $(\text{supp}(\mu_{A_N} \boxplus \mu_{sc}))_N$ are uniformly bounded,

$$\text{supp}(\mu_{A_N} \boxplus \mu_{sc}) \subset [\Psi(-(K-1); \Psi(K-1)]. \tag{24}$$

Let $\epsilon > 0$. Since Ψ is uniformly continuous on $[-K; K]$, there exists $0 < \alpha < 1$ such that

$$\Psi(\{x, \operatorname{dist}(x, \overline{U} \cup \Theta_\nu) < \alpha\}) \subset \{y, \operatorname{dist}(y, K_\nu(\theta_1, \ldots, \theta_J)) < \epsilon\}. \tag{25}$$

Since according to Lemma 2.4, $\overline{U_N} \subset \{x, \operatorname{dist}(x, \overline{U} \cup \Theta_\nu) < \alpha/2\}$ for all large N, we have

$$\Psi_N\left([-K; K] \cap \left\{x, \operatorname{dist}(x, \overline{U} \cup \Theta_\nu) \geq \frac{\alpha}{2}\right\}\right) \subset \Psi_N(\mathbb{R}\backslash\overline{U_N}) = \mathbb{R}\backslash \operatorname{supp}(\mu_{A_N} \boxplus \mu_{sc}). \tag{26}$$

Denote by $\mathcal{A}_{\alpha,K}$ the set $[-K; K] \cap \{x, \operatorname{dist}(x, \overline{U} \cup \Theta_\nu) \geq \alpha\}$. Note that for all large N, $\mathcal{A}_{\alpha/2,K} \subset \mathbb{R}\backslash\overline{U_N}$. Moreover,

$$\bigcup_{x \in \mathcal{A}_{\alpha,K-1}} \left]x - \frac{\alpha}{2}; x + \frac{\alpha}{2}\right[\subset \mathcal{A}_{\alpha/2,K}.$$

Thus, using Lemma 2.2 for Ψ_N, we get that

$$\bigcup_{x \in \mathcal{A}_{\alpha,K-1}} \left]\Psi_N\left(x - \frac{\alpha}{2}\right); \Psi_N\left(x + \frac{\alpha}{2}\right)\right[\subset \Psi_N(\mathcal{A}_{\alpha/2,K}).$$

Now, using the assumptions (H_3) on the spectrum of A_N, it is easy to see that Ψ_N converges uniformly towards Ψ on the compact set $\mathcal{A}_{\alpha/2,K}$. Moreover, since Ψ is continuous on the compact set $\mathcal{A}_{\alpha,K-1}$, we have

$$\inf_{x \in \mathcal{A}_{\alpha,K-1}} \min(|\Psi(x - \alpha/2) - \Psi(x)|; |\Psi(x + \alpha/2) - \Psi(x)|) = m > 0.$$

Therefore since for all large N, $\sup_{\mathcal{A}_{\alpha/2,K}} |\Psi_N(x) - \Psi(x)| < m$ and using also Lemma 2.2 for Ψ, we get that for all large N, for all $x \in \mathcal{A}_{\alpha,K-1}$, $\Psi_N\left(x - \frac{\alpha}{2}\right) < \Psi(x) < \Psi_N\left(x + \frac{\alpha}{2}\right)$ and therefore

$$\Psi(\mathcal{A}_{\alpha,K-1}) \subset \Psi_N(\mathcal{A}_{\alpha/2,K}). \tag{27}$$

(26) and (27) yield that

$$\operatorname{supp}(\mu_{A_N} \boxplus \mu_{sc}) \subset \mathbb{R}\backslash\Psi(\mathcal{A}_{\alpha,K-1})$$

with

$$\mathbb{R}\backslash\Psi(\mathcal{A}_{\alpha,K-1}) =]-\infty; \Psi(-K+1)[\cup]\Psi(K-1); +\infty[\cup\Psi(\{x, \operatorname{dist}(x, \overline{U} \cup \Theta_\nu) < \alpha\}.$$

Then, the result readily follows from (24) and (25). □

Moreover, in [12], the authors proved that the spikes of the perturbation which belong to $\mathbb{R}\backslash\overline{U}$, generate outliers in the spectrum of the deformed model.

Theorem 2.4 [12] *Let $\theta_j \in \mathbb{R}\backslash\overline{U}$ (i.e. $\in \Theta_\nu$). Denote by $n_{j-1}+1, \ldots, n_{j-1}+k_j$ the descending ranks of θ_j among the eigenvalues of A_N. Then, almost surely,*

$$\lim_{N\to\infty} \lambda_{n_{j-1}+i}(M_N) = \rho_{\theta_j} = H(\theta_j), \quad \forall\, 1 \le i \le k_j.$$

Remark 2.2 In [12], the authors proved this theorem when the support of ν has a finite number of connected components; nevertheless it is still true in our more general setting since it follows from Theorems 2.3 and 2.2 and an exact separation phenomenon (see Theorem 7.1 in [12]) which proof does not care about the number of connected components of the support of ν.

3 Comparison of the supports of $\mu_{sc} \boxplus \nu$ and $\mu_{sc} \boxplus \mu_{A_N}$

As we show in the next Sect. 4, the support of $\mu_{sc} \boxplus \mu_{A_N}$ plays a fundamental role in the study of the fluctuations of eigenvalues at the edges of the spectrum. Due to assumptions (H_2), (H_3) and (H_4), we are able to show that the supports of $\mu_{sc} \boxplus \nu$ and $\mu_{sc} \boxplus \mu_{A_N}$ exhibit very similar features at edges which are distant from outliers as we explain in Sect. 3.1 below. In Sect. 3.2, we prove that $\mu_{sc} \boxplus \mu_{A_N}$ has a connected component in the vicinity of each outlier. Sections 3.3 and 3.4 are devoted to the proof of the propositions stated in Sects. 3.1 and 3.2.

3.1 Fundamental preliminary results

The two following results will be fundamental for considering asymptotics of the correlation kernel at the edges of the support of $\mu_{sc} \boxplus \nu$.

Proposition 3.1 *Assume that for a sufficiently small $\epsilon > 0$,*

$$p(u) > 0, \quad \forall u \in]u_0 - \epsilon; u_0[, \text{ and } p(u) = 0, \quad \forall u \in [u_0; u_0 + \epsilon[.$$

Set $t_0 = \Psi^{-1}(u_0)$. Then there exists $\tau > 0$ such that $]t_0 - \tau; t_0[\subset U$, $[t_0; t_0 + \tau[\subset \mathbb{R}\backslash U$ and we have $\int \frac{d\nu(x)}{(t_0-x)^2} = 1$ and $\int \frac{d\nu(x)}{(t_0-x)^3} > 0$. Assume that for all $j \in \{1, \ldots, J\}, \theta_j \neq t_0$. Then for $\tau > 0$ small enough, for all large N, there exists one and only one $t_0(N)$ in $]t_0 - \tau; t_0 + \tau[$, such that $\int \frac{1}{(t_0(N)-x)^2} d\mu_{A_N}(x) = 1$ and $]t_0 - \tau; t_0 + \tau[\cap U_N =]t_0 - \tau; t_0(N)[$. Moreover, for $\eta > 0$ small enough, for all large N, $u_0(N) = \Psi_N(t_0(N)) \in]u_0 - \eta; u_0 + \eta[$,

$$\text{and } \forall u \in]u_0 - \eta; u_0(N)[, \ p_N(u) > 0 \text{ and } \forall u \in [u_0(N); u_0 + \eta[, \ p_N(u) = 0.$$

Moreover, we have

$$u_0(N) = u_0 + \epsilon_N(t_0(N)) + \frac{1}{4}(\epsilon_N{}'(t_0(N)))^2(1 + o(1)) + O\left(\frac{1}{N}\right), \tag{28}$$

where for t in a small neighborhood of t_0,

$$\epsilon_N(t) = \frac{N-r}{N} \int \frac{d\hat{\nu}_N(x)}{(t-x)} - \int \frac{d\nu(x)}{(t-x)}.$$

Similarly we have the following result involving the left edges of the support of $\mu_{sc} \boxplus \nu$.

Proposition 3.2 *Assume that for a sufficiently small* $\epsilon > 0$,

$$p(u) > 0, \quad \forall u \in]u_0; u_0+\epsilon[, \text{ and } p(u) = 0, \quad \forall u \in [u_0-\epsilon; u_0[.$$

Set $t_0 = \Psi^{-1}(u_0)$. *Then there exists* $\tau > 0$ *such that* $]t_0-\tau; t_0] \subset \mathbb{R}\backslash U$, $]t_0; t_0+\tau[\subset U$ *and we have* $\int \frac{d\nu(x)}{(t_0-x)^2} = 1$ *and* $\int \frac{d\nu(x)}{(t_0-x)^3} < 0$. *Assume that for all* $j \in \{1, \ldots, J\}, \theta_j \neq t_0$. *Then for* $\tau > 0$ *small enough, for all large N, there exists one and only one* $t_0(N)$ *in* $]t_0-\tau; t_0+\tau[$, *such that* $\int \frac{1}{(t_0(N)-x)^2} d\mu_{A_N}(x) = 1$ *and* $]t_0-\tau; t_0+\tau[\cap U_N =]t_0(N); t_0+\tau[$. *Moreover, for* $\eta > 0$ *small enough, for all large N*, $u_0(N) = \Psi_N(t_0(N)) \in]u_0-\eta; u_0+\eta[$

$$\text{and} \quad \forall u \in]u_0(N); u_0+\eta[, \quad p_N(u) > 0 \text{ and } \quad \forall u \in]u_0-\eta; u_0(N)], \quad p_N(u) = 0.$$

Moreover we have

$$u_0(N) = u_0 + \epsilon_N(t_0(N)) + \frac{1}{4}(\epsilon_N{}'(t_0(N)))^2(1+o(1)) + O\left(\frac{1}{N}\right),$$

where for t in a small neighborhood of t_0,

$$\epsilon_N(t) = \frac{N-r}{N} \int \frac{d\hat{\nu}_N(x)}{(t-x)} - \int \frac{d\nu(x)}{(t-x)}.$$

Remark 3.1 It is clear that, under the assumption (3) of Shcherbina ([26]), Theorem 1.1 and (28) imply her result.

The following proposition will be fundamental to study the asymptotics of the correlation kernel in a neighborhood of any point of the support of $\mu_{sc} \boxplus \nu$ where the density vanishes.

Proposition 3.3 *Let* $u_0 \in \mathbb{R}$ *be such that* $p(u_0) = 0$ *and there exists* $\epsilon > 0$ *such that*, $\forall u \in]u_0-\epsilon; u_0+\epsilon[\backslash\{u_0\}$, $p(u) > 0$. *Set* $t_0 = \Psi^{-1}(u_0) \in \mathbb{R}$. *Then* t_0 *is a point in* $\mathbb{R}\backslash supp(\nu)$ *where two components of U merge and satisfies* $\int \frac{d\nu(s)}{(t_0-s)^2} = 1$, $\int \frac{d\nu(s)}{(t_0-s)^3} = 0$. *We have* $u_0 = H(t_0)$. *Assume that assumption* (H_4) *holds true and that for any* $i = 1, \ldots, J$, $\theta_i \neq t_0$. *Then, for* η *small enough, for all large N, there exists* $u_0(N)$ *in* $]u_0-\eta; u_0+\eta[$ *such that* $p_N(u_0(N)) = 0$ *and* $\forall u \in]u_0-\eta; u_0+\eta[\backslash\{u_0(N)\}$, $p_N(u) > 0$. $t_0(N) = \Psi_N^{-1}(u_0(N))$ *is a point of* $\mathbb{R}\backslash Spect(A_N)$ *where two components of* U_N *merge and satisfies* $\int \frac{d\mu_{A_N}(s)}{(t_0(N)-s)^2} = 1$, $\int \frac{d\mu_{A_N}(s)}{(t_0(N)-s)^3} = 0$. *We have* $u_0(N) = H_N(t_0(N))$ *and* $\lim_{N\to+\infty} t_0(N) = t_0$.

3.2 In the vicinity of outliers

It turns out that the support of $\mu_{sc} \boxplus \mu_{A_N}$ exhibits a small connected component in the vicinity of each outlier.

Proposition 3.4 *Let θ_i be such that $\int \frac{d\nu(x)}{(\theta_i - x)^2} < 1$ and $\rho_{\theta_i} = H(\theta_i)$. Then, for $\epsilon > 0$ small enough, for all large N, $supp(\mu_{sc} \boxplus \mu_{A_N})$ has a unique connected component $[L_i(N); D_i(N)]$ inside $]\rho_{\theta_i} - \epsilon; \rho_{\theta_i} + \epsilon[$. Moreover, setting $\rho_N(\theta_i) = \frac{1}{N}\sum_{y_j \neq \theta_i} \frac{1}{\theta_i - y_j} + \theta_i$, we have*

$$L_i(N) = \rho_N(\theta_i) - 2\sqrt{k_i}\sqrt{1 - \int \frac{1}{(\theta_i - x)^2} d\nu(x)} \frac{1}{\sqrt{N}} + o\left(\frac{1}{\sqrt{N}}\right),$$

$$D_i(N) = \rho_N(\theta_i) + 2\sqrt{k_i}\sqrt{1 - \int \frac{1}{(\theta_i - x)^2} d\nu(x)} \frac{1}{\sqrt{N}} + o\left(\frac{1}{\sqrt{N}}\right).$$

Thus, $\rho_N(\theta_i) = \frac{L_i(N) + D_i(N)}{2} + o\left(\frac{1}{\sqrt{N}}\right)$.

3.3 Some technical lemmas

In the proof of the previous propositions, we will use the following lemmas.

Lemma 3.1 *Let $[a; b] \subset \mathbb{R} \backslash supp(\nu) \cup \Theta$. Then $m_N : z \mapsto \int \frac{d\mu_{A_N}(s)}{(z-s)}$ (resp. $-m_N^{'} : z \mapsto \int \frac{d\mu_{A_N}(s)}{(z-s)^2}$) converges uniformly towards $m : z \mapsto \int \frac{d\nu(s)}{(z-s)}$ (resp. $-m^{'} : z \mapsto \int \frac{d\nu(s)}{(z-s)^2}$) on every compact set included in $\{z \in \mathbb{C}; a < \Re z < b\}$.*

Proof Let $\gamma > 0$ be such that $[a - 3\gamma; b + 3\gamma] \subset \mathbb{R} \backslash \mathrm{supp}(\nu) \cup \Theta$. Since A_N has $N - r$ eigenvalues $\beta_j(N)$ satisfying $\max_{1 \le j \le N-r} \mathrm{dist}(\beta_j(N), \mathrm{supp}(\nu)) \longrightarrow_{N \to \infty} 0$, and the other eigenvalues of A_N are the spikes $\theta_j \in \Theta$, we can readily deduce that for all large N,

$$[a - 2\gamma; b + 2\gamma] \subset \mathbb{R} \backslash \mathrm{Spect}(A_N).$$

It is clear that the functions $m_N, g, -m_N^{'}$ and $-m^{'}$ are holomorphic on $\{z \in \mathbb{C}; a - \gamma < \Re z < b + \gamma\}$. Since for large N, $\{z \in \mathbb{C}; a - \gamma < \Re z < b + \gamma\}$ is included in $\{z \in \mathbb{C}; \mathrm{dist}(z; \mathrm{supp}(\nu)) > \gamma; \mathrm{dist}(z; \mathrm{Spect}(A_N)) > \gamma\}$, it readily follows that for large N, m_N and m (respectively $m_N^{'}$ and $m^{'}$) are uniformly bounded by $1/\gamma$ (respectively $1/\gamma^2$). Since the sequence of measures μ_{A_N} weakly converges to ν, it is easy to see that $m_N(z)$ (respectively $m_N^{'}(z)$) converges towards $m(z)$ (respectively $m^{'}(z)$) for all $z \in]a; b[$. Therefore, by Montel's theorem, the convergence is uniform on every compact set of $\{z \in \mathbb{C}; a - \gamma < \Re z < b + \gamma\}$ □

Lemma 3.2 (1) *For any t in U, $v_N(t)$ converges towards $v(t)$ when N goes to infinity.*

(2) *For any t in U such that $t \in \mathbb{R}\backslash\{supp(\nu) \cup \Theta\}$, $\Psi_N(t)$ converges towards $\Psi(t)$ when N goes to infinity.*
(3) *For any t in $\mathbb{R}\backslash\{\overline{U} \cup \Theta\}$, $\Psi_N(t)$ converges towards $\Psi(t)$ when N goes to infinity.*

Proof Let t be in U. Therefore we have $v(t) > 0$. Let $0 < \epsilon < v(t)$. We have $\int \frac{d\nu(s)}{(t-s)^2+(v(t)-\epsilon)^2} > 1$ and $\int \frac{d\nu(s)}{(t-s)^2+(v(t)+\epsilon)^2} < 1$ which implies that for all large N, $\int \frac{d\mu_{A_N}(s)}{(t-s)^2+(v(t)-\epsilon)^2} > 1$ and $\int \frac{d\mu_{A_N}(s)}{(t-s)^2+(v(t)+\epsilon)^2} < 1$. It follows that for all large N, $v(t) - \epsilon < v_N(t) < v(t) + \epsilon$.
Now, let t be in U and such that $t \in \mathbb{R}\backslash\{\mathrm{supp}(\nu) \cup \Theta\}$. Let $\delta > 0$ such that $[t - \delta; t + \delta] \subset \mathbb{R}\backslash\{\mathrm{supp}(\nu) \cup \Theta\}$. According to Lemma 3.1, $z \mapsto \int \frac{d\mu_{A_N}(x)}{z-x}$ converges towards $z \mapsto \int \frac{d\nu(x)}{z-x}$ uniformly on every compact set of $\{t - \delta < \Re z < t + \delta\}$. Since $v_N(t)$ converges towards $v(t)$, for all large N, $0 \le v_N(t) \le v(t) + 1$. The convergence of $\Psi_N(t)$ towards $\Psi(t)$ when N goes to infinity readily follows from the uniform convergence of $z \mapsto \int \frac{d\mu_{A_N}(x)}{z-x}$ towards $z \mapsto \int \frac{d\nu(x)}{z-x}$ on the compact set $\{z = t + ib, 0 \le b \le v(t) + 1\}$.
Let t be in $\mathbb{R}\backslash\{\overline{U} \cup \Theta\}$. Since $v(t) = 0$, we have $\Psi(t) = H(t)$. Since we assume that $\mathrm{supp}(\nu) \subset U$, we have $t \in \mathbb{R}\backslash\mathrm{supp}(\nu)$. According to the assumption (H_3) on the spectrum of A_N, for $\delta > 0$ small enough, for all large N, we have $[t - \delta; t + \delta] \subset \mathbb{R}\backslash\{\mathrm{supp}(\nu) \cup \mathrm{supp}(\mu_{A_N})\}$. Therefore it is easy to see that $\int \frac{d\mu_{A_N}(x)}{(t-x)}$ converges towards $\int \frac{d\nu(x)}{(t-x)}$ and $\int \frac{d\mu_{A_N}(x)}{(t-x)^2}$ converges towards $\int \frac{d\nu(x)}{(t-x)^2} < 1$ and thus for all large N, $t \notin U_N$. It follows that for all large N, $v_N(t) = 0$ and $\Psi_N(t) = H_N(t) = t + \int \frac{d\mu_{A_N}(x)}{(t-x)}$ converges towards $t + \int \frac{d\nu(x)}{(t-x)} = H(t) = \Psi(t)$. □

Lemma 3.3 *Let $t_0 \notin \mathrm{supp}(\nu) \cup \Theta$ be such that $\int \frac{1}{(t_0-x)^2} d\nu(x) = 1$ and $\int \frac{1}{(t_0-x)^3} d\nu(x) \neq 0$. Then for small enough $\epsilon > 0$, for all large N, there exists one and only one $t_0(N) \in]t_0 - \epsilon; t_0 + \epsilon[$ such that $\int \frac{1}{(t_0(N)-x)^2} d\mu_{A_N}(x) = 1$. $t_0(N)$ satisfies*

$$t_0(N) = t_0 + f_N(t_0(N))$$

where $f_N(t)$

$$= h(t) \left[\left\{ \frac{N-r}{N} \int \frac{1}{(t-x)^2} d\hat{\nu}_N(x) - \int \frac{1}{(t-x)^2} d\nu(x) \right\} + \frac{1}{N} \sum_{j=1}^{J} \frac{k_j}{(t-\theta_j)^2} \right]$$

with $h(t) = \frac{1}{\int \frac{(t-x+t_0-x)}{(t-x)^2(t_0-x)^2} d\nu(x)}$ and $0 < K_1(\epsilon) < |h(t)| < K_2(\epsilon), \forall t \in]t_0 - \epsilon; t_0 + \epsilon[$.
Moreover, if $]t_0 - \epsilon; t_0[\subset U$ and $]t_0; t_0 + \epsilon[\subset \mathbb{R}\backslash U$ (respectively $]t_0; t_0 + \epsilon[\subset U$ and $]t_0 - \epsilon; t_0[\subset \mathbb{R}\backslash U$), then for all large N, $]t_0 - \epsilon; t_0 + \epsilon[\cap U_N =]t_0 - \epsilon; t_0(N)[$ (respectively $]t_0 - \epsilon; t_0 + \epsilon[\cap U_N =]t_0(N); t_0 + \epsilon[$.)

Proof One can readily see that $t \notin \{\theta_i, i = 1, \ldots, J, \beta_j, j = 1, \ldots, N - r\}$ is in U_N if and only if $P_N(t) > 0$ where $P_N(t)$ is the polynomial defined by

$$P_N(t) = \prod_{i=1}^{N-r}(t-\beta_i)^2 \prod_{j=1}^{J}(t-\theta_j)^2 \left(\int \frac{d\mu_{A_N}}{(t-x)^2} - 1\right) \tag{29}$$

$$\begin{aligned} &= \frac{1}{N}\sum_{i=1}^{N-r}\prod_{l\neq i}(t-\beta_l)^2 \prod_{j=1}^{J}(t-\theta_j)^2 \\ &\quad + \frac{1}{N}\prod_{i=1}^{N-r}(t-\beta_i)^2 \sum_{j=1}^{J} k_j \prod_{l\neq j}(t-\theta_l)^2 \\ &\quad - \prod_{j=1}^{J}(t-\theta_j)^2 \prod_{i=1}^{N-r}(u-\beta_i)^2. \end{aligned} \tag{30}$$

Condition (H_3) on the spectrum of A_N allows us to choose $\epsilon > 0$ small enough such that for N large enough $[t_0 - 2\epsilon; t_0 + 2\epsilon]$ is in the complement of the support of ν and the support of μ_{A_N}.

$P_N(t) = 0$ for $t \in]t_0 - \epsilon; t_0 + \epsilon[$ if and only if

$$1 - \frac{N-r}{N}\int \frac{1}{(t-x)^2} d\hat{\nu}_N - \frac{1}{N}\sum_{j=1}^{J}\frac{k_j}{(t-\theta_j)^2} = 0. \tag{31}$$

Using that

$$\int \frac{1}{(t_0-x)^2} d\nu(x) = 1,$$

(31) can be rewritten as follows:

$$\begin{aligned} &\int \frac{1}{(t_0-x)^2} d\nu(x) - \int \frac{1}{(t-x)^2} d\nu(x) \\ &= \left\{\frac{N-r}{N}\int \frac{1}{(t-x)^2} d\hat{\nu}_N(x) - \int \frac{1}{(t-x)^2} d\nu(x)\right\} + \frac{1}{N}\sum_{j=1}^{J}\frac{k_j}{(t-\theta_j)^2}, \end{aligned}$$

or equivalently

$$\begin{aligned} &(t-t_0)\int \frac{(t-x+t_0-x)}{(u-x)^2(t_0-x)^2} d\nu(x) \\ &= \left\{\frac{N-r}{N}\int \frac{1}{(t-x)^2} d\hat{\nu}_N(x) - \int \frac{1}{(t-x)^2} d\nu(x)\right\} + \frac{1}{N}\sum_{j=1}^{J}\frac{k_j}{(t-\theta_j)^2}. \end{aligned}$$

Since we have $\int \frac{1}{(t_0-x)^3} d\nu(x) \neq 0$, it readily follows that for $\epsilon > 0$ small enough and for all z such that $|z - t_0| \leq \epsilon$, $\int \frac{(z-x+t_0-x)}{(z-x)^2(t_0-x)^2} d\nu(x) \neq 0$. Therefore, there exists $C_1(\epsilon) > 0$ and $C_2(\epsilon) > 0$ such that for any z such that $|z - t_0| \leq \epsilon$, $0 < C_1(\epsilon) < |\int \frac{(z-x+t_0-x)}{(z-x)^2(t_0-x)^2} d\nu(x)| < C_2(\epsilon)$. Define on $\{z; |z - t_0| \leq \epsilon\}$,

$$h(z) = \frac{1}{\int \frac{(z-x+t_0-x)}{(z-x)^2(t_0-x)^2} d\nu(x)}.$$

Using Lemma 3.1, by Rouché theorem, for large N, the function

$$z - t_0 - h(z) \left[\left\{ \frac{N-r}{N} \int \frac{d\hat{\nu}_N(x)}{(z-x)^2} - \int \frac{d\nu(x)}{(z-x)^2} \right\} + \frac{1}{N} \sum_{j=1}^{J} \frac{k_j}{|z-\theta_j|^2} \right]$$

has exactly one zero z_0 in $\{z; |z - t_0| < \epsilon\}$. Since $\bar{z_0}$ is obviously a zero too, we can conclude that z_0 is real. Hence, for ϵ small enough, for all large N, P_N has exactly one zero $t_0(N)$ in $]t_0 - \epsilon; t_0 + \epsilon[$ and

$$t_0(N) = t_0 + h(t_0(N)) \left[\left\{ \frac{N-r}{N} \int \frac{d\hat{\nu}_N(x)}{(t_0(N)-x)^2} - \int \frac{d\nu(x)}{(t_0(N)-x)^2} \right\} + \frac{1}{N} \sum_{j=1}^{J} \frac{k_j}{(t_0(N)-\theta_j)^2} \right]$$

where $0 < K_1(\epsilon) < |h(t_0(N))| < K_2(\epsilon)$.

Now, if $]t_0 - \epsilon; t_0[\subset U$ and $]t_0; t_0 + \epsilon[\subset \mathbb{R} \backslash U$ (respectively $]t_0; t_0 + \epsilon[\subset U$ and $]t_0 - \epsilon; t_0[\subset \mathbb{R} \backslash U$), then since for all large N, $P_N(t_0 - \epsilon/2) > 0$ and $P_N(t_0 + \epsilon/2) < 0$ (respectively $P_N(t_0 - \epsilon/2) < 0$ and $P_N(t_0 + \epsilon/2) > 0$), it is clear that for all large N, $]t_0 - \epsilon; t_0 + \epsilon[\cap U_N =]t_0 - \epsilon; t_0(N)[$ (respectively $]t_0 - \epsilon; t_0 + \epsilon[\cap U_N =]t_0(N); t_0 + \epsilon[$.) The proof of Lemma 3.3 is complete. □

Lemma 3.4 *Let t_0 be such that $\int \frac{d\nu(s)}{(t_0-s)^2} = 1$, $t_0 \neq \theta_j, \forall 1 \leq j \leq J$, and there exists $\tau > 0$ such that, $\forall t \in]t_0 - \tau; t_0 + \tau[\backslash \{t_0\}$, $\int \frac{d\nu(s)}{(t-s)^2} > 1$. Then, $t_0 \notin supp(\nu) \cup \Theta$. Set $d_1 = \sup\{s \in supp(\nu) \cup \Theta; s < t_0\}$ and $d_2 = \inf\{s \in supp(\nu) \cup \Theta; s > t_0\}$. Let $[a; b]$ be such that $t_0 \in]a; b[$, $[a; b] \subset]d_1; d_2[$. Then, $\forall t \in [a; b] \backslash \{t_0\}$, $\int \frac{d\nu(s)}{(t-s)^2} > 1$. Assume that (H_4) holds true. Then moreover, for all large N, $[a; b] \subset \mathbb{R} \backslash Spect(A_N)$ and there exists $t_0(N)$ in $[a; b]$ such that $\int \frac{d\mu_{A_N}(s)}{(t_0(N)-s)^2} = 1$, $\int \frac{d\mu_{A_N}(s)}{(t_0(N)-s)^3} = 0$ and $\forall t \in [a; b] \backslash \{t_0(N)\}$, $\int \frac{d\mu_{A_N}(s)}{(t-s)^2} > 1$. We have also $\lim_{N\to+\infty} t_0(N) = t_0$.*

Proof Since we assume that for any t in $\text{supp}(\nu)$, $\int \frac{d\nu(s)}{(t-s)^2} > 1$ (i.e $\text{supp}(\nu) \subset U$) and that $t_0 \neq \theta_j, \forall 1 \leq j \leq J$, it readily follows that $t_0 \notin \text{supp}(\nu) \cup \Theta$. Let $[a; b]$ be such that $t_0 \in]a; b[$, $[a; b] \subset]d_1; d_2[$.

Since $\int \frac{d\nu(s)}{(t_0-s)^2} = 1$ and there exists $\tau > 0$ such that, $\forall t \in]t_0 - \tau; t_0 + \tau[\backslash\{t_0\}$, $\int \frac{d\nu(s)}{(t-s)^2} > 1$, the strict convexity of $z \mapsto \int \frac{d\nu(s)}{(z-s)^2}$ on $[a; b]$ implies that

$$\forall t \in [a; b]\backslash\{t_0\}, \int \frac{d\nu(s)}{(t-s)^2} > 1. \tag{32}$$

By Lemma 3.1, $\phi_N : z \mapsto \int \frac{d\mu_{A_N}(s)}{(z-s)^2} - 1$ converges uniformly towards $\phi : z \mapsto \int \frac{d\nu(s)}{(z-s)^2} - 1$ on every compact set of $\{z \in \mathbb{C}; a < \Re z < b\}$.
By the principle of isolated zeroes, there exist δ_0 such that $[t_0-\delta_0; t_0+\delta_0] \subset]a; b[$ and ϕ has no other zero in $\{z \in \mathbb{C}; |z - t_0| \leq \delta_0\}$ than t_0. Thus, using Hurwitz's theorem and assumption (H_4), we can claim that for any $0 < \delta < \delta_0$, for all large N, ϕ_N has a unique real zero $t_0(N)$ in $\{z \in \mathbb{C}; |z - t_0| < \delta\}$ and that $\phi_N'(t_0(N)) = 0$. Moreover since ϕ_N is strictly convex on $[a; b]$, we have $\forall t \in [a; b]\backslash\{t_0(N)\}$, $\phi_N(t) > 0$. □

Lemma 3.5 *For each i such that* $\int \frac{1}{(\theta_i-x)^2} d\nu(x) < 1$, *for* $\epsilon > 0$ *small enough, for all large* N, $U_N \bigcap]\theta_i - \epsilon; \theta_i + \epsilon[=]t_1^i(N), t_2^i(N)[$ *where* $t_1^i(N)$ *and* $t_2^i(N)$ *satisfy*

$$t_1^i(N) = \theta_i - \sqrt{\frac{k_i}{N}\phi_N(t_1^i(N))}$$
$$t_2^i(N) = \theta_i + \sqrt{\frac{k_i}{N}\phi_N(t_2^i(N))}$$

with $\phi_N(t) = \frac{1}{1-\frac{N-r}{N}\int \frac{1}{(t-x)^2} d\hat{\nu}_N(x) - \frac{1}{N}\sum_{j\neq i} \frac{k_j}{(t-\theta_j)^2}}$ *and* $1 \leq \phi_N(t) \leq K(\epsilon)$ *for any* $t \in]\theta_i - \epsilon; \theta_i + \epsilon[$.

Proof Let θ_i be such that $\int \frac{d\nu(x)}{(\theta_i-x)^2} < 1$. Let $\epsilon > 0$ be such that $]\theta_i - 4\epsilon; \theta_i + 4\epsilon[\subset \mathbb{R}\backslash\{\text{supp}(\nu) \cup \{\theta_j, j \neq i\}\}$ and $\inf_{z\in\mathbb{C}, |z-\theta_i|\leq 2\epsilon} |\int \frac{d\nu(x)}{(z-x)^2} - 1| = m \neq 0$. In particular, we have that for any t in $[\theta_i-\epsilon; \theta_i+\epsilon]$, $\int \frac{d\nu(x)}{(t-x)^2} < 1$. According to the assumption (H_3) on the spectrum of A_N, for all large N, $[\theta_i - 3\epsilon; \theta_i[\cup]\theta_i; \theta_i + 3\epsilon] \subset \mathbb{R}\backslash\text{Spect}(A_N)$. Note that since $\int \frac{d\mu_{A_N}(x)}{(\theta_i\pm\epsilon-x)^2}$ converges towards $\int \frac{d\nu(x)}{(\theta_i\pm\epsilon-x)^2}$, we have moreover for all large N, $\int \frac{d\mu_{A_N}(x)}{(\theta_i\pm\epsilon-x)^2} < 1$, whereas $\int \frac{d\mu_{A_N}(x)}{(\theta_i-x)^2} = +\infty$. Therefore, for all large N, there exists at least one $s_N \in]\theta_i - \epsilon; \theta_i[$ and at least one $t_N \in]\theta_i; \theta_i + \epsilon[$ such that $\int \frac{d\mu_{A_N}(x)}{(s_N-x)^2} = 1$ and $\int \frac{d\mu_{A_N}(x)}{(t_N-x)^2} = 1$. Let us study the zeroes of the polynomial P_N defined by (30) in $\{z; |z - \theta_i| < \epsilon\}$. We know that there are at least two real zeroes s_N and t_N. Let us rewrite

$$P_N(t) = \frac{1}{N}\sum_{i=1}^{N-r}\prod_{l\neq i}(t-\beta_l)^2\prod_{j=1}^{J}(t-\theta_j)^2 + \frac{1}{N}\prod_{l=1}^{N-r}(t-\beta_l)^2\sum_{j\neq i}k_j\prod_{p\neq j}(t-\theta_p)^2$$
$$+\frac{1}{N}\prod_{j=1}^{N-r}(t-\beta_j)^2 k_i\prod_{l\neq i}(t-\theta_l)^2 - \prod_{j=1}^{J}(t-\theta_j)^2\prod_{l=1}^{N-r}(t-\beta_l)^2.$$

$P_N(t) = 0$ for t such that $|t - \theta_i| < 2\epsilon$ if and only if

$$(t-\theta_i)^2 \left\{ 1 - \frac{N-r}{N} \int \frac{1}{(t-x)^2} d\hat{\nu}_N(x) - \frac{1}{N} \sum_{j \neq i} \frac{k_j}{(t-\theta_j)^2} \right\} = \frac{k_i}{N},$$

Since for all large N, $[\theta_i - 3\epsilon; \theta_i + 3\epsilon] \subset \mathbb{R} \backslash \{\text{supp}(\hat{\nu}_N) \cup \text{supp}(\nu)\}$, using the same arguments as in the proof of Lemma 3.1, we get easily the uniform convergence on any compact set included in $\{z \in \mathbb{C}, \theta_i - 3\epsilon < \Re z < \theta_i + 3\epsilon\}$ of $z \mapsto 1 - \frac{N-r}{N} \int \frac{1}{(z-x)^2} d\hat{\nu}_N(x) - \frac{1}{N} \sum_{j \neq i} \frac{k_j}{(z-\theta_j)^2}$ towards $z \mapsto 1 - \int \frac{1}{(z-x)^2} d\nu(x)$. Hence, we have for all large N,

$$\inf_{z \in \mathbb{C}, |z-\theta_i| \leq 2\epsilon} \left| 1 - \frac{N-r}{N} \int \frac{1}{(z-x)^2} d\hat{\nu}_N(x) - \frac{1}{N} \sum_{j \neq i} \frac{k_j}{(z-\theta_j)^2} \right| \geq m/2$$

and

$$\inf_{t \in [\theta_i - 2\epsilon; \theta_i + 2\epsilon]} \left\{ 1 - \frac{N-r}{N} \int \frac{1}{(t-x)^2} d\hat{\nu}_N(x) - \frac{1}{N} \sum_{j \neq i} \frac{k_j}{(t-\theta_j)^2} \right\} \geq m/2$$

and the zeros of P_N in $\{z \in \mathbb{C}, |z - \theta_i| < 2\epsilon\}$ are the solutions of the equation $(z - \theta_i)^2 = \frac{k_i}{N} \phi_N(z)$ where

$$\phi_N(z) = \frac{1}{1 - \frac{N-r}{N} \int \frac{1}{(z-x)^2} d\hat{\nu}_N(x) - \frac{1}{N} \sum_{j \neq i} \frac{k_j}{(z-\theta_j)^2}}$$

and $0 < |\phi_N(z)| \leq \frac{2}{m}$. Therefore, by Hurwitz theorem, for all large N, P_N has exactly two zeroes in $\{z \in \mathbb{C}, |z - \theta_i| < \epsilon\}$. Since we have already seen that P_N has at least one zero in $]\theta_i - \epsilon; \theta_i[$ and at least one zero in $]\theta_i; \theta_i + \epsilon[$, we can conclude that for all large N, P_N has exactly one zero $t_1^i(N)$ in $]\theta_i - \epsilon; \theta_i[$ and one zero $t_2^i(N)$ in $]\theta_i; \theta_i + \epsilon[$. Moreover since $\phi_N(t) > 0$ on $[\theta_i - \epsilon; \theta_i + \epsilon]$, we have

$$t_1^i(N) = \theta_i - \sqrt{\frac{k_i}{N} \phi_N(t)} \text{ and } t_2^i(N) = \theta_i + \sqrt{\frac{k_i}{N} \phi_N(t)}.$$

Now, since $P_N(\theta_i) > 0$, it is clear that $U_N \bigcap]\theta_i - \epsilon; \theta_i + \epsilon[=]t_1^i(N), t_2^i(N)[$. The proof of Lemma 3.5 is complete. □

3.4 Proof of Propositions 3.1, 3.2, 3.3 and 3.4

Proof of Proposition 3.1 Using (15) and (16), it is clear that $\Psi^{-1}(]u_0 - \epsilon; u_0[) \subset U$ and $\Psi^{-1}([u_0; u_0 + \epsilon[) \subset \mathbb{R} \backslash U$. Note that since we assume that $\text{supp}(\nu) \subset U$, this implies that $t_0 = \Psi^{-1}(u_0) \in \mathbb{R}\backslash \mathrm{supp}(\nu)$. Let $0 < \delta < \epsilon$ be such that $\Psi^{-1}(]u_0 - \delta; u_0 + \delta[) \subset \mathbb{R}\backslash\{\mathrm{supp}(\nu) \cup \Theta\}$. Since according to Theorem 2.1, the homeomorphism Ψ is strictly increasing on U, we have $\Psi^{-1}(]u_0 - \delta; u_0[) =]\Psi^{-1}(u_0 - \delta); t_0[\subset U$. Moreover according to Lemma 2.2, $\Psi^{-1}([u_0; u_0 + \delta[) = [t_0; \Psi^{-1}(u_0 + \delta)[\subset \mathbb{R}\backslash U$. Thus, according to Lemma 2.3 (i) and (ii), we have $\int \frac{d\nu(x)}{(t_0 - x)^2} = 1$, $\int \frac{d\nu(x)}{(t_0-x)^3} > 0$. Then, using Lemma 3.3, for τ small enough, for all large N there exists one and only one $t_0(N) \in]t_0 - \tau; t_0 + \tau[$ such that $\int \frac{1}{(t_0(N)-x)^2} d\mu_{A_N}(x) = 1$. $t_0(N)$ satisfies

$$t_0(N) = t_0 + f_N(t_0(N))$$

where

$$f_N(t) = h(t) \left[\left\{ \frac{N-r}{N} \int \frac{d\hat{\nu}_N(x)}{(t-x)^2} - \int \frac{d\nu(x)}{(t-x)^2} \right\} + \frac{1}{N} \sum_{j=1}^{J} \frac{k_j}{(t-\theta_j)^2} \right]$$

with $h(t) = \frac{1}{\int \frac{(t-x+t_0-x)}{(t-x)^2(t_0-x)^2} d\nu(x)}$ and $0 < K_1(\tau) < |h(t)| < K_2(\tau)$, $\forall t \in]t_0 - \tau; t_0 + \tau[$. Moreover, for all large N,

$$]t_0 - \tau; t_0 + \tau[\cap U_N =]t_0 - \tau; t_0(N)[. \tag{33}$$

Since according to Theorem 2.1, Ψ_N is strictly increasing on U_N, we have

$$\Psi_N(]t_0 - \tau; t_0(N)[) =]\Psi_N(t_0 - \tau); \Psi_N(t_0(N))[. \tag{34}$$

Moreover according to Lemma 2.2,

$$\Psi_N([t_0(N); t_0 + \tau[) = [\Psi_N(t_0(N)); \Psi_N(t_0 + \tau)[. \tag{35}$$

Note that $\Psi_N(t_0(N)) = H_N(t_0(N)) = t_0(N) + \int \frac{d\mu_{A_N}(x)}{t_0(N) - x}$ with for τ small enough and N large enough $t_0(N) \in [t_0 - \tau; t_0 + \tau)] \subset \mathbb{R}\backslash\{\mathrm{supp}(\nu) \cup \Theta\}$. Lemma 3.1 readily yields that $u_0(N) = \Psi_N(t_0(N))$ converges towards $H(t_0) = \Psi(\Psi^{-1}(u_0)) = u_0$. Now, for τ small enough, $t_0 + \tau \in \mathbb{R}\backslash\{\overline{U} \cup \Theta\}$ and $t_0 - \tau \in U$, $t_0 - \tau \in \mathbb{R}\backslash\{\mathrm{supp}(\nu) \cup \Theta\}$, so that using Lemma 3.2, for any $\eta > 0$ small enough, for all large N,

$$\Psi_N(t_0 + \tau) > u_0 + \eta \quad \text{and} \quad \Psi_N(t_0 - \tau) < u_0 - \eta. \tag{36}$$

It readily follows from (33), (34), (35), (36), (15) and (16) that for any $\eta > 0$ small enough, for all large N,

$$\forall u \in [u_0(N); u_0 + \eta[, \;\; p_N(u) = 0 \text{ and } \forall u \in]u_0 - \eta; u_0(N)[, \;\; p_N(u) > 0.$$

Now, for t in a small neighborhood of t_0 and N large enough let us define

$$\epsilon_N(t) = \frac{N-r}{N} \int \frac{d\hat{\nu}_N(x)}{(t-x)} - \int \frac{d\nu(x)}{(t-x)}.$$

We have

$$
\begin{aligned}
\Psi_N(t_0(N)) &= H_N(t_0(N)) \\
&= t_0(N) + \int \frac{d\mu_{A_N}(x)}{t_0(N)-x} \\
&= H(t_0) + f_N(t_0(N)) + \int \frac{d\mu_{A_N}(x)}{t_0(N)-x} - \int \frac{d\nu(x)}{t_0 - x} \\
&= H(t_0) + f_N(t_0(N)) + \epsilon_N(t_0(N)) + \int \frac{d\nu(x)}{t_0(N)-x} - \int \frac{d\nu(x)}{t_0-x} \\
&\quad + O\left(\frac{1}{N}\right) \\
&= H(t_0) + f_N(t_0(N)) + \epsilon_N(t_0(N)) \\
&\quad - f_N(t_0(N))\left[\int \frac{d\nu(x)}{(t_0-x)^2} - f_N(t_0(N)) \int \frac{d\nu(x)}{(t_0(N)-x)(t_0-x)^2}\right] \\
&\quad + O\left(\frac{1}{N}\right) \\
&= H(t_0) + \epsilon_N(t_0(N)) + f_N(t_0(N))^2 \int \frac{d\nu(x)}{(t_0(N)-x)(t_0-x)^2} \\
&\quad + O\left(\frac{1}{N}\right) \\
&= H(t_0) + \epsilon_N(t_0(N)) + \frac{1}{4}(\epsilon_N{}'(t_0(N)))^2(1+o(1)) + O\left(\frac{1}{N}\right).
\end{aligned}
$$

The proof of Proposition 3.1 is complete. □

The proof of Proposition 3.2 is similar and left to the reader.

Proof of Proposition 3.3 According to Theorem 2.1,

$$\Psi^{-1}(]u_0-\epsilon; u_0+\epsilon[) \subset \overline{U}$$

and more precisely, since

$$p(\Psi(\Psi^{-1}(u))) = \frac{v(\Psi^{-1}(u))}{\pi},$$

we have $\Psi^{-1}(]u_0-\epsilon; u_0+\epsilon[\setminus\{u_0\}) \subset U$ and $x_0 = \Psi^{-1}(u_0) \notin U$. Since we assume that $\mathrm{supp}(\nu) \subset U$, $x_0 \notin \mathrm{supp}(\nu)$. Note that $u_0 = \Psi(x_0) = H(x_0)$ since $v(x_0) = 0$. Moreover, since the homeomorphism Ψ is strictly increasing on U, it is easy to see that Ψ^{-1} is strictly increasing on $]u_0-\epsilon; u_0+\epsilon[$ and $\Psi^{-1}(]u_0 - \epsilon; u_0+\epsilon[\setminus\{u_0\}) =]\Psi^{-1}(u_0-\epsilon); x_0[\cup]x_0; \Psi^{-1}(u_0+\epsilon)[$. Therefore x_0 is a point in the complement of $\mathrm{supp}(\nu)$ where two components of the set U merge into one. Therefore, Lemma 2.1 implies that $\int \frac{d\nu(s)}{(x_0-s)^2} = 1$ and $\int \frac{d\nu(s)}{(x_0-s)^3} = 0$. Since we assume that for any $\theta_i \in \Theta$, $\theta_i \neq x_0$, we have $x_0 \notin \Theta$. Therefore x_0 satisfies the assumptions of Lemma 3.4. Let η be such that $0 < 2\eta < \epsilon$ and $[\Psi^{-1}(u_0 - 2\eta); \Psi^{-1}(u_0 + 2\eta)] \subset \mathbb{R}\backslash\{\text{supp}(\nu) \cup \Theta\}$. According to Lemma 3.4, for all large N, there exists $x_0(N)$ in $[\Psi^{-1}(u_0 - 2\eta); \Psi^{-1}(u_0 + 2\eta)]$ such that $\int \frac{d\mu_{A_N}(s)}{(x_0(N)-s)^2} = 1$, $\int \frac{d\mu_{A_N}(s)}{(x_0(N)-s)^3} = 0$ and $[\Psi^{-1}(u_0-2\eta); \Psi^{-1}(u_0+2\eta)]\backslash\{x_0(N)\} \subset U_N$. We have also $\lim_{N\to+\infty} x_0(N) = x_0$. Note that since

$$\text{for any } x \in \mathbb{R},\ \ p_N(\Psi_N(x)) = \frac{v_N(x)}{\pi},$$

we have

$$p_N(\Psi_N(x_0(N))) = 0$$

and

$$\forall x \in [\Psi^{-1}(u_0 - 2\eta); \Psi^{-1}(u_0 + 2\eta)]\backslash\{x_0(N)\},\ \ p_N(\Psi_N(x)) > 0.$$

Using Lemma 3.2, we can deduce that for all large N,

$$\Psi_N(\Psi^{-1}(u_0 - 2\eta)) < u_0 - \eta \text{ and } \Psi_N(\Psi^{-1}(u_0 + 2\eta)) > u_0 + \eta.$$

Moreover since $\lim_{N\to+\infty} x_0(N) = x_0$, we have for all large N, $x_0(N) \in]\Psi^{-1}(u_0 - \eta/2); \Psi^{-1}(u_0 + \eta/2)[$ so that $u_0(N) = \Psi_N(x_0(N)) \in]u_0 - \eta; u_0 + \eta[$ for all large N by using oncemore Lemma 3.2. The proof is complete. □

Proof of Proposition 3.4 According to Lemma 3.5, for $\epsilon > 0$ small enough, for all large N, $U_N \bigcap]\theta_i - \epsilon; \theta_i + \epsilon[=]t_1^i(N), t_2^i(N)[$ where $t_1^i(N)$ and $t_2^i(N)$ satisfy

$$t_1^i(N) = \theta_i - \sqrt{\frac{k_i}{N}\phi_N(t_1^i(N))}$$
$$t_2^i(N) = \theta_i + \sqrt{\frac{k_i}{N}\phi_N(t_2^i(N))}$$

with $\phi_N(t) = \frac{1}{1-\frac{N-r}{N}\int \frac{1}{(t-x)^2} d\hat{\nu}_N(x) - \frac{1}{N}\sum_{j\neq i}\frac{k_j}{(t-\theta_j)^2}}$ and $1 \leq \phi_N(t) \leq K(\epsilon)$ for any $t \in]\theta_i - \epsilon; \theta_i + \epsilon[$. For N large enough $t_1^i(N) > \theta_i - \epsilon/2$ and $t_2^i(N) < \theta_i + \epsilon/2$ and $[\theta_i - \epsilon/2; \theta_i + \epsilon/2] \cap \overline{U_N} = [t_1^i(N), t_2^i(N)]$. Therefore, according to Theorem 2.1, $[\Psi_N(t_1^i(N)), \Psi_N(t_2^i(N))]$ is a connected component of $\text{supp}(\mu_{sc} \boxplus \mu_{A_N})$ and $[\Psi_N(\theta_i - \epsilon/2); \Psi_N(t_1^i(N))[\cup]\Psi_N(t_2^i(N)); \Psi_N(\theta_i + \epsilon/2]) \subset \mathbb{R}\backslash\text{supp}(\mu_{sc} \boxplus \mu_{A_N})$.

Now, we have

$$\Psi_N(t_1^i(N)) = \theta_i - \sqrt{\frac{k_i}{N}\phi_N(t_1^i(N))} + \frac{N-r}{N}\int \frac{d\hat{\nu}_N(x)}{\left(\theta_i - \sqrt{\frac{k_i}{N}\phi_N(t_1^i(N))} - x\right)}$$

$$+ \sum_{j\neq i} \frac{k_j}{N\left(\theta_i - \theta_j - \sqrt{\frac{k_i}{N}\phi_N(t_1^i(N))}\right)} - \frac{k_i}{N\sqrt{\frac{k_i}{N}\phi_N(t_1^i(N))}}$$

$$= \theta_i - \sqrt{\frac{k_i}{N}\phi_N(t_1^i(N))} - \frac{\sqrt{k_i}}{\sqrt{N}}\frac{1}{\sqrt{\phi_N(t_1^i(N))}}$$

$$+ \frac{N-r}{N}\int \frac{1}{\left(\theta_i - \sqrt{\frac{k_i}{N}\phi_N(t_1^i(N))} - x\right)} d\hat{\nu}_N(x) + O\left(\frac{1}{N}\right)$$

$$= \theta_i - \sqrt{\frac{k_i}{N}\phi_N(t_1^i(N))}\left\{1 - \int \frac{1}{(\theta_i - x)^2} d\nu(x)\right\}$$

$$+ \frac{N-r}{N}\int \frac{1}{\theta_i - x} d\hat{\nu}_N(x) - \frac{\sqrt{k_i}}{\sqrt{N}}\frac{1}{\sqrt{\phi_N(t_1^i(N))}}$$

$$+ \sqrt{\frac{k_i}{N}\phi_N(t_1^i(N))}\left\{\frac{N-r}{N}\int \frac{d\hat{\nu}_N(x)}{(\theta_i - x)^2} - \int \frac{d\nu(x)}{(\theta_i - x)^2}\right\} + O\left(\frac{1}{N}\right)$$

$$= \rho_N(\theta_i) - \frac{\tau_i}{\sqrt{N}} + o\left(\frac{1}{\sqrt{N}}\right)$$

with $\rho_N(\theta_i) := \frac{1}{N}\sum_{y_j \neq \theta_i} \frac{1}{\theta_i - y_j} + \theta_i$ and $\tau_i = 2\sqrt{k_i}\sqrt{1 - \int \frac{1}{(\theta_i - x)^2} d\nu(x)}$. In the same way

$$\Psi_N(t_2^i(N)) = \rho_N(\theta_i) + \frac{\tau_i}{\sqrt{N}} + o\left(\frac{1}{\sqrt{N}}\right).$$

Note that $\Psi_N(t_1^i(N))$ and $\Psi_N(t_2^i(N))$ converges towards $\rho_{\theta_i} = \Psi(\theta_i)$. Since for ϵ small enough, $[\theta_i - \epsilon; \theta_i + \epsilon] \subset \mathbb{R}\backslash(\overline{U} \cup \Theta)$ (see (18)), according to Lemma 3.2(3), $\Psi_N(\theta_i - \epsilon/2)$ and $\Psi_N(\theta_i + \epsilon/2)$ converge respectively towards $\Psi(\theta_i - \epsilon/2)$ and $\Psi(\theta_i + \epsilon/2)$ and, according to Lemma 2.2, $\Psi(\theta_i - \epsilon/2) < \Psi(\theta_i - \epsilon/4) < \Psi(\theta_i) < \Psi(\theta_i + \epsilon/4) < \Psi(\theta_i + \epsilon/2)$. Now, for all large N, $\Psi_N(\theta_i - \epsilon/2) < \Psi(\theta_i - \epsilon/4)$ and $\Psi(\theta_i + \epsilon/4) < \Psi_N(\theta_i + \epsilon/2)$. Then, for any $0 < \eta < \min\{\Psi(\theta_i + \epsilon/4) - \Psi(\theta_i); \Psi(\theta_i) - \Psi(\theta_i - \epsilon/4)\}$, for all large N, we have $\Psi_N(t_1^i(N)) > \Psi(\theta_i) - \eta$ and $\Psi_N(t_2^i(N)) < \Psi(\theta_i) + \eta$ whereas $\Psi_N(\theta_i - \epsilon/2) < \Psi(\theta_i) - \eta$ and $\Psi_N(\theta_i + \epsilon/2) > \Psi(\theta_i) + \eta$. Thus $[\Psi_N(t_1^i(N)), \Psi_N(t_2^i(N))]$ is the unique connected component of $\mathrm{supp}(\mu_{sc} \boxplus \mu_{A_N})$ inside $]\Psi(\theta_i) - \eta; \Psi(\theta_i) + \eta[$. The proof of Proposition 3.4 is complete. ◻

4 Proofs of Theorem 1.1, Theorem 1.2 and Theorem 1.3

4.1 Correlation functions of the deformed GUE

It is known from Johansson [16] (see also [10]) that the joint eigenvalue density induced by the deformed GUE M_N can be explicitly computed. Furthermore it induces a so-called "determinantal random point field". In other words, if one considers a symmetric function $f : \mathbb{R}^m \to \mathbb{R}$, one has that

$$\mathbb{E} \sum_{1 \le i_1 < i_2 < \cdots < i_m \le N} f(\lambda_{i_1}, \ldots, \lambda_{i_m}) \\ = \int f(x_1, \ldots, x_m) \frac{1}{m!} \det(K_N(x_i, x_j))_{i,j=1}^m \prod_{i=1}^m dx_i,$$

where K_N is the so-called correlation kernel of the deformed GUE, which has been explicited by [16]. We here state his result.

Proposition 4.1 [16] *The correlation of the deformed GUE M_N is given by the double complex integral:*

$$K_N(u, v) = \frac{N}{(2i\pi)^2} \int_\Gamma \int_\gamma e^{N\frac{(w-v)^2}{2} - N\frac{(z-u)^2}{2}} \frac{1}{w - z} \prod_{i=1}^N \frac{w - y_i}{z - y_i} dw dz, \tag{37}$$

where Γ encircles the poles $y_1, \ldots, y_N$ and γ is a line parallel to the y-axis not crossing Γ.

At this point, it is worth mentioning that correlation functions and thus local eigenvalue statistics are invariant through conjugation of the correlation kernel. Indeed, one has that

$$\det(K_N(u_i, u_j))_{i,j=1}^m = \det\left(K_N(u_i, u_j)\frac{h(u_i)}{h(u_j)}\right)_{i,j=1}^m,$$

for any non vanishing function h. This fact will be used many times in this article.

Before starting the asymptotic analysis, we list some important facts and notations that are needed hereafter.

Let u_0 be given. Assume that both u and v satisfy $|u - u_0| \le N^{-\delta}$ for some $\delta > 0$. Let us set

$$F_{u_0}(z) := \frac{(z - u_0)^2}{2} + \int_{\mathbb{R}} \ln(z - y) d\nu(y).$$

Note that F_{u_0} is the first order approximation (as $N \to \infty$) of the true exponential term arising in both z and w integrals in the correlation kernel K_N. Indeed the true exponential term arising in both integrals is given by

$$F_{u_0,N}(z) := \frac{(z-u_0)^2}{2} + \frac{1}{N}\sum_{i=1}^{N} \ln(z-y_i).$$

We neglect for a while the fake singularity introduced by the logarithm (as $e^{F_{u_0,N}}$ is holomorphic). By definition, critical points satisfy

$$F'_{u_0,N}(z) = z - u_0 + \frac{1}{N}\sum_{i=1}^{N}\frac{1}{z-y_i} = 0$$

and one can note that $F''_{u_0,N} = 1 - \frac{1}{N}\sum_{i=1}^{N}\frac{1}{(z-y_i)^2}$ does not depend on u_0. It is also convenient for the following to define the curve of critical points of both F_u and $F_{u,N}$. Let us define

$$\mathcal{C} = \{x \pm iv(x), x \in \mathbb{R}\}.$$

One can check that a critical point of F_u with non null imaginary part lies on

$$\begin{aligned}\{x \pm iv(x), x \in U\} &= \mathcal{C} \cap \{z \in \mathbb{C}, \mathrm{Im} z \neq 0\} \\ &= \left\{z \in \mathbb{C}, \mathrm{Im} z \neq 0, \int \frac{1}{|z-y|^2} d\nu(y) = 1\right\}.\end{aligned}$$

For any $u \in \Psi(U)$, we denote by $z_c^{\pm}(u)$ these two critical points:

$$z_c^{\pm}(u) = \Psi^{-1}(u) \pm iv(\Psi^{-1}(u)).$$

Formula (15) due to Biane shows that $|\mathrm{Im} z_c(u)| = \pi p(u)$.

If instead F_u has no non real critical point, then $u \in \Psi(U^c)$. As a consequence there exists a unique $z_c(u) \in \mathcal{C} \cap \mathbb{R} = U^c$ such that $F'_u(z_c(u)) = 0$. The real numbers u and $z_c(u)$ are then related by the equation

$$u := z_c(u) + \int \frac{1}{z_c(u) - y} d\nu(y) \quad \text{i.e. } z_c(u) = \Psi^{-1}(u).$$

This follows from the fact that $\Psi : \mathbb{R} \to \mathbb{R}$ is one to one. In all cases $z_c^{\pm}(u)$, $z_c(u)$ and u are related by :

$$H(z_c^{(\pm)}(u)) = u.$$

Similarly we define

$$\mathcal{C}_N = \{x \pm iv_N(x), x \in \mathbb{R}\}.$$

A critical point of $F_{u,N}$ with non zero imaginary part lies on

$$\{x \pm i v_N(x), x \in U_N\} = \mathcal{C}_N \cap \{z \in \mathbb{C}, \mathrm{Im} z \neq 0\}$$
$$= \left\{ z \in \mathbb{C}, \mathrm{Im} z \neq 0, \frac{1}{N}\sum_{j=1}^{N} \frac{1}{|z-y_j|^2} = 1 \right\}.$$

For any $u \in \Psi_N(U_N)$, denote by $z_{c,N}^{\pm}(u)$ these two critical points of $F_{u,N}$:

$$z_{c,N}^{\pm}(u) = \Psi_N^{-1}(u) \pm i v_N(\Psi_N^{-1}(u)).$$

We note that $F_{u,N}$ necessarily admits $N-1$ other critical points, which are real interlaced with the y_i's. We disregard these critical points. Then one has that

$$H_N(z_{c,N}^{\pm}(u)) = u.$$

If instead $F_{u,N}$ has no non real critical points, $u \in \Psi_N(U_N^c)$ and there exists a unique $z_{c,N}(u) \in \mathbb{R} \cap \mathcal{C}_N = U_N^c$ such that $F'_{u,N}(z_{c,N}(u)) = 0$. Again one has that

$$u = H_N(z_{c,N}(u)) = z_{c,N}(u) + \frac{1}{N}\sum_{i=1}^{N} \frac{1}{z_{c,N}(u) - y_i}.$$

We emphasize that according to (18)

$$\forall z \in (\mathcal{C}_N \mathring{\cap} \mathbb{R}) = \overline{U_N}^c, \quad \frac{1}{N}\sum_{i=1}^{N} \frac{1}{(z-y_i)^2} < 1,$$

and that, according to Theorem 2.1 and Lemma 2.2, $u \mapsto \Re z_{c,N}^{(\pm)}(u) = \Psi_N^{-1}(u)$ is a strictly increasing function.

Actually in all the cases we study, it turns out that the critical points, that we here denote by $\mathbf{z_c}$, lie on the real axis. We may therefore need to modify $F_{u,N}$ so that there is no singularity in the logarithm. It may happen in particular that $\exists\, 1 \leq i \leq N$, $y_i < \mathbf{z_c} < y_{i+1}$. However by the assumptions we have made, in all cases there exists $\epsilon > 0$ such that $[\mathbf{z_c} - \epsilon, \mathbf{z_c} + \epsilon]$ contains no eigenvalue y_j, $j = 1, \ldots, N$. In that case we set

$$F_{u,N} = \frac{(z-u_0)^2}{2} + \frac{1}{N}\sum_{i:y_i < \mathbf{z_c}+\epsilon} \ln(z-y_i) + \frac{1}{N}\sum_{i:y_i > \mathbf{z_c}+\epsilon} \ln(y_i - z). \tag{38}$$

The contour Γ will be split into two parts: Γ_1 lying to the left of $\mathbf{z_c} + \epsilon$ and Γ_2 to its right (encircling all the eigenvalues $y_i > \mathbf{z_c} + \epsilon$). The contour γ will be chosen so that it lies to the left of $\mathbf{z_c} + \epsilon$. All these contours cross the real axis at a point where $F_{u,N}$ has no singularity. Note that with this new definition of $F_{u,N}$, it is still true that

$$F'_{u,N}(z) = z - u + \frac{1}{N} \sum_{i=1}^{N} \frac{1}{z - y_i}.$$

Thus all the subsequent derivatives and the curve $\mathcal{C}_N$ are unchanged with this new definition. The asymptotic exponential term at $\mathbf{z_c}$ is then given by

$$F_{u_0}(z) = \frac{(z-u_0)^2}{2} + \int_{(-\infty, \mathbf{z_c}+\epsilon)} \ln(z-y) d\nu(y) + \int_{(\mathbf{z_c}+\epsilon, +\infty)} \ln(-z+y) d\nu(y).$$

4.2 Asymptotics of the correlation kernel at the edges of the support

4.2.1 Proof of Theorem 1.1

We start from a right extremity point d of a connected component of $\text{supp}(\nu \boxplus \mu_{sc})$ so that $p(x) = 0, \forall x \in [d, d+\epsilon]$ for some small $\epsilon > 0$. We assume moreover that for any θ_j such that $\int \frac{d\nu(s)}{(\theta_i - s)^2} = 1$, we have $d \neq \theta_j + m_\nu(\theta_j)$. According to Proposition 3.1, such a point d satisfies $d = H(\mathbf{z_0})$ where $\mathbf{z_0}$ is a real solution of

$$F''_d(\mathbf{z_0}) = 0.$$

Since $\mathbf{z_0} \notin \text{supp}(\nu) \cup \Theta$, (H_3) implies that for all large N, one also has that $\inf_{k=1,\ldots,N} \text{dist}(\mathbf{z_0}, y_k) > 0$. By Proposition 3.1, there exists a unique extremity point d_N which is the right endpoint of a connected component of $\text{supp}(\mu_N \boxplus \mu_{sc})$ and such that $|d - d_N| \leq \epsilon$ for any ϵ. Then there exists a point $\mathbf{z_N}$ such that

$$H_N(\mathbf{z_N}) = d_N.$$

Let $F_{d_N,N}$ be defined as in (38) with $\mathbf{z_c} = \mathbf{z_N}$. By definition, one has that $\mathbf{z_N}$ is the real degenerate critical point associated to d_N:

$$F'_{d_N,N}(\mathbf{z_N}) = 0, \quad \text{and} \quad F''_{d_N,N}(\mathbf{z_N}) = 0. \tag{39}$$

We now turn to the asymptotics of the correlation kernel. Let $\alpha \in \mathbb{R}$ to be fixed later. Assume that

$$u_0 := d_N, \quad u = u_0 + \frac{\alpha x}{N^{\frac{2}{3}}}; \quad v = u_0 + \frac{\alpha y}{N^{\frac{2}{3}}} \tag{40}$$

We assume that there exists a real number $M_0 > 0$ such that $x, y \geq -M_0$. If u_0 is not the top edge of the support $\text{supp}(\mu_{A_N} \boxplus \mu_\sigma)$, then x and y shall be bounded from above by $\epsilon_0 N^{2/3}$ with ϵ_0 small enough so that $u_0 + \frac{\alpha x}{N^{2/3}}$ is smaller than the left edge of the next connected component of $\text{supp}(\mu_{A_N} \boxplus \mu_\sigma)$.

The associated rescaled correlation kernel is then

$$\frac{\alpha}{N^{\frac{2}{3}}} K_N(u, v).$$

We now consider the asymptotics of the correlation kernel and prove that the rescaled kernel $\frac{\alpha}{N^{2/3}} K_N(u, v)$ uniformly converges to the Airy kernel when $-M_0 \le x, y \le \epsilon_0 N^{2/3}$.

Theorem 1.1 is an easy consequence of the following Proposition. Set

$$\alpha = 2^{1/3} \frac{1}{|F^{(3)}_{u_0,N}(\mathbf{z_N})|^{1/3}}.$$

α is well defined using Lemma 2.3 (ii).

Proposition 4.2 *There exist constants $q, C, c > 0$ such that for any $x, y \in [-M_0, \epsilon_0 N^{2/3}]$,*

$$\left| \frac{\alpha}{N^{\frac{2}{3}}} K_N(u, v) e^{q(y-x)N^{\frac{1}{3}}} - \mathbf{A}(x, y) \right| \le \frac{C e^{-c(x+y)}}{N^{\frac{1}{3}}},$$

where $\mathbf{A}$ *denotes the Airy kernel.*

Proof of proposition 4.2 By Cauchy's theory and using the fact proved in Lemma 2.3 that

$$F^{(3)}_d(\mathbf{z_0}) = 2 \int \frac{1}{(\mathbf{z_0} - y)^3} d\nu(y) = a_i > 0,$$

one deduces that $F^{(3)}_{u_0,N}(\mathbf{z_N}) \ge a_i/2$ and that there exist $a > 0$, $M > 0$ and a small δ-neighborhood of $\mathbf{z_N}$ such that

$$\forall z,\ |z - \mathbf{z_N}| \le \delta,\ \Re F^{(3)}_{u_0,N}(z) > a \ \text{ and }\ |F^{(4)}_{u_0,N}(z)| \le M. \tag{41}$$

We now rewrite the correlation kernel. To this aim, we split Γ into two contours lying respectively to the left and to the right of $\mathbf{z_N}$. This is possible as we assume that $\Delta := \inf_{k=1,\dots,N} \mathrm{dist}(\mathbf{z_0}, y_k) > 0$ and $|z_N - z_0| < \Delta/2$ for N large enough. Denote by Γ_1 the part of the contour Γ lying to the left of $\mathbf{z_N}$ and set $\Gamma_2 := \Gamma \backslash \Gamma_1$. In the correlation kernel given by Proposition 4.1, along Γ_1, we first rewrite the singularity

$$1/(w - z) = \alpha N^{\frac{1}{3}} \int_{\mathbb{R}^+} e^{-N^{\frac{1}{3}} \alpha t_o (w-z)} dt_o,$$

which is valid provided the contour γ remains to the right of Γ_1. This then yields the following expression for the correlation kernel (up to a conjugation factor):

$$\frac{\alpha}{N^{\frac{2}{3}}} K_N(u, v) = \frac{\alpha^2 N^{2/3}}{(2i\pi)^2} \int_{\mathbb{R}^+} dt_o \int_{\Gamma_1} dz \int_{\gamma} dw$$

$$e^{-N^{\frac{1}{3}}\alpha t_o(w-z)} e^{N(\frac{w^2}{2}-wv-\frac{z^2}{2}+uz)} \prod_{i=1}^N \frac{w-y_i}{z-y_i} \tag{42}$$

$$+\frac{\alpha N^{\frac{1}{3}}}{(2i\pi)^2}\int_{\Gamma_2}\int_{\gamma} e^{N\frac{(w-v)^2}{2}-N\frac{(z-u)^2}{2}} \frac{1}{w-z}\prod_{i=1}^N \frac{w-y_i}{z-y_i} dwdz. \tag{43}$$

We denote by $K_N^{(l)}$ (resp. $K_N^{(r)}(u,v)$) the kernel arising in (42) [resp. (43) that we consider separately].
Note that it is enough to concentrate on $F_{u_0,N}$ for the saddle point analysis of the correlation kernel. Assume given $q \in \mathbb{R}$ that we will fix later. We rewrite the correlation kernel (and use conjugation thanks to q) as:

$$\begin{aligned} &K_N^{(l)}(u,v)e^{q(y-x)N^{\frac{1}{3}}} \\ &= \frac{1}{(2i\pi)^2}\int_{\mathbb{R}^+} dt_o \int_{\Gamma_1}\int_{\gamma} H(w, y+t_0)G(z, x+t_0)dwdz, \end{aligned} \tag{44}$$

where

$$\begin{aligned} H(w,y) &:= \alpha N^{\frac{1}{3}} e^{NF_{u_0,N}(w)-\alpha y(w-q)N^{\frac{1}{3}}}, \\ G(z,x) &:= \alpha N^{\frac{1}{3}} e^{-NF_{u_0,N}(z)+\alpha x(z-q)N^{\frac{1}{3}}}. \end{aligned} \tag{45}$$

Let us first consider the leading term in the exponential defining H and G that is $F_{u_0,N}$. By the choice of u_0, the two first derivatives of the exponential term vanish at the real point $\mathbf{z_N}$ so that standard saddle point analysis suggest that the ascent and descent contours shall be given by lines with direction $(2)i\pi/3$ through the critical point $\mathbf{z_N}$. This is true in a compact neighborhood of $\mathbf{z_N}$, as we see below. We ignore for a while the constraint that the contours do not cross each other.
We first check that Γ_1 and γ shall follow the directions $2i\pi/3$ or $i\pi/3$. To consider the constraint that they do not cross each other, we later modify these contours in a $N^{-1/3}$ neighborhood of $\mathbf{z_N}$. Using (41), there exists $\delta_0 > 0$ and $a = a(\delta_0)$ such that for any $|s| \leq \delta_0$

$$\begin{aligned} &\Re(F_{u_0,N}(\mathbf{z_N}+se^{i\pi/3}) - F_{u_0,N}(\mathbf{z_N})) \\ &= -\Re\left(s^3\int_0^1\int_0^1\int_0^1 dtdxdvF^{(3)}_{u_0,N}(\mathbf{z_N}+stxve^{i\pi/3})\right) \\ &= -s^3\int_0^1\int_0^1\int_0^1 dtdxdv\Re\frac{2}{N}\sum_{j=1}^N\frac{1}{(\mathbf{z_N}+stxve^{i\pi/3}-y_j)^3} < -as^3. \\ &\Re(F_{u_0,N}(\mathbf{z_N}+se^{i2\pi/3}) - F_{u_0,N}(\mathbf{z_N})) \\ &= \Re\left(s^3\int_0^1\int_0^1\int_0^1 dtdxdvF^{(3)}_{u_0,N}(\mathbf{z_N}+stxve^{2i\pi/3})\right) \\ &= s^3\int_0^1\int_0^1\int_0^1 dtdxdv\Re\frac{2}{N}\sum_{j=1}^N\frac{1}{(\mathbf{z_N}+stxve^{2i\pi/3}-y_j)^3} > as^3. \end{aligned} \tag{46}$$

One can then complete the w-contour by a line parallel to the imaginary axis. Indeed one can choose δ_0 small enough so that $\mathbf{z_N} + \delta_0 e^{i\pi/3}$ lies in the domain where $1 > \frac{1}{N}\sum \frac{1}{|z-y_i|^2}$. Thus there exists a constant $a' > 0$ such that

$$\frac{d\Re F_{u_0,N}(\mathbf{z_N} + \delta_0 e^{i\pi/3} + it)}{dt} < -a't, \ \ t > 0.$$

As a consequence $\Re F_{u_0,N}$ still decreases along the contour $t \mapsto \mathbf{z_N} + \delta_0 e^{i\pi/3} + it$, $t > 0$. This yields the descent path γ for the w-integral.

For the z-integral, we complete the contour as follows.

If $\mathbf{z_N} + \delta_0 e^{2i\pi/3}$ lies above the curve $\mathcal{C}_N$, we complete the contour by lines parallel to the real axis $x \mapsto \mathbf{z_N} + \delta_0 e^{2i\pi/3} + x$, $x < 0$, up to the moment one crosses the curve $\mathcal{C}_N$. Then this part of contour remains on the domain $\{z, \frac{1}{N}\sum_{j=1}^N \dfrac{1}{|z-y_j|^2} \leq 1\}$. Thus, one can check that there exists a constant $a'' > 0$ such that

$$\frac{d\Re F_{u_0,N}(\mathbf{z_N} + \delta_0 e^{2i\pi/3} - x)}{dx} > a''x.$$

This (part of) line is then an ascent path for $F_{u,N}$.

At the moment (if it exists) where the curve $x \mapsto \mathbf{z_N} + \delta_0 e^{2i\pi/3} - x$, $x < 0$ crosses $\mathcal{C}_N$, one follows $\mathcal{C}_N$ to the left direction up to the moment of time where $\mathrm{Im} z \leq \delta_0\sqrt{3}/2$ and then again follow a line parallel to the real axis. Due to the fact that $u \mapsto \Re z_c(u)$ is an increasing function, this part of the contour is also an ascent path.

If instead $\mathbf{z_N} + \delta_0 e^{2i\pi/3}$ lies below the curve $\mathcal{C}_N$, we first follow the contour $\mathbf{z_N} + \delta_0 e^{2i\pi/3} + it$, where $t \geq 0$ up to the moment one crosses $\mathcal{C}_N$. One then follows $\mathcal{C}_N$ to the left direction up to the moment of time where $\mathrm{Im} z \leq \delta_0\sqrt{3}/2$ and then again follow a line parallel to the real axis. It is an easy computation to check that this contour is also an ascent path.

Because d_N may not be the right edge of the support, we need to complete the z-contour Γ_2 to the right of $\mathbf{z_N}$ too. In this case, define

$$\mathbf{z'_N} = \inf\{x \in \mathbb{R}, x > \mathbf{z_N}, v_N(x) > 0\}.$$

Note that the contour $\mathcal{C}_N \cap \{z \in \mathbb{C}, \Re(z) \geq \mathbf{z'_N}\}$ is made of contours around y_i's. Let Z be the first point encountered on $\mathcal{C}_\mathcal{N}$ to the right of $\mathbf{z'_N}$ such that $\mathrm{Im}(Z)$ is a local maximum. The contour Γ_2 then follows $\mathcal{C}_N \cap \{z \in \mathbb{C}, \Re(Z) \geq \Re(z) \geq \mathbf{z'_N}\}$. Afterwards Γ_2 follows the highest of the two curves $\{Z + x, x > 0\}$ and $\mathcal{C}_N \cap \{\Re(z) > \Re(Z)\}$. The contour is completed by symmetry with respect to the real axis. Because $\mathcal{C}_N$ is the curve of critical points, along Γ_2 which lies above $\mathcal{C}_N$, one has that

$$\forall z \in \Gamma_2 \cap \mathcal{C}_N, \exists u > u_0, \ z = z_{c,N}(u) \ \text{and} \ \Re F_{u_0,N}(z) > \Re F_{u_0,N}(\mathbf{z_N});$$
$$\forall x > 0, \ \frac{\partial}{\partial x}\Re F_{u_0,N}(Z+x) > 0,$$

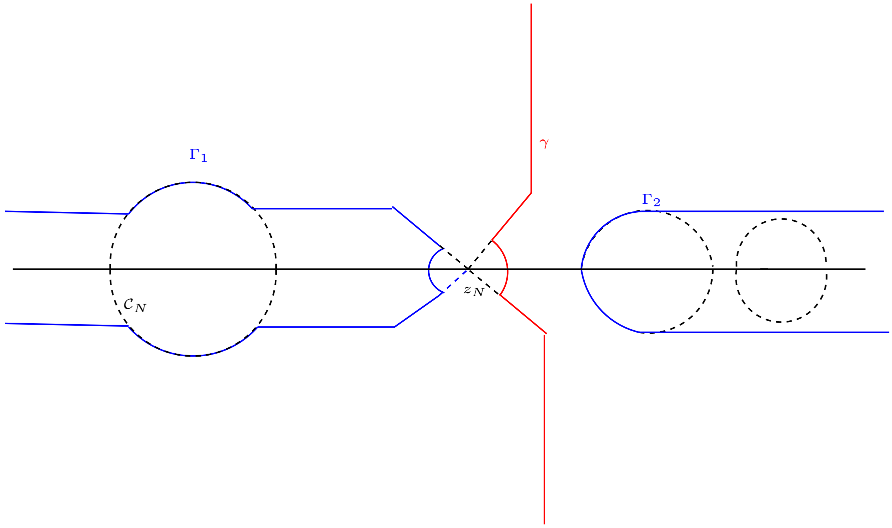

Fig. 2 The contour Γ and γ at an edge

as long as $Z+x$ lies above $\mathcal{C}_N$. This finishes the definition of the contours, apart from the constraint that the two contours cannot cross each other.

We now slightly modify the contours in a $N^{-\frac{1}{3}}$ neighborhood of $\mathbf{z_N}$ so that γ does not cross Γ_1. Let $\epsilon > 0$ (small) be fixed. The w and z contours do not go through $\mathbf{z_N}$ but instead follows an arc of circle of ray $\epsilon N^{-\frac{1}{3}}$ centered at $\mathbf{z_N}$ in order to avoid crossing each other (see Fig. 2). We now fix $q = \mathbf{z_N} + \frac{\epsilon}{2} N^{-\frac{1}{3}}$ where ϵ has been defined as above. By the estimates on the decay of $F_{u_0,N}$ given in (46), we deduce the following.

Assume first that $|x|, |y| \leq M_0$. Using (46), we first deduce that there exists $A > 0$ such that

$$\int_\gamma H(w,x)dw = \alpha \int_{|w-\mathbf{z_N}|\leq\delta_0} H(w,x)dw(1+O(e^{-AN})).$$

Let us now set $\gamma_0 := \{te^{i\pm\pi/3}, \epsilon \leq t \leq \delta_0 N^{1/3}\} \cup C_\epsilon$ where C_ϵ is the arc of circle centered at 0 joining $\epsilon e^{-i\pi/3}$ and $\epsilon e^{i\pi/3}$. This contour is oriented from bottom to top. We now make the change of variables $w = \mathbf{z_N} + sN^{-\frac{1}{3}}$ where $s \in \gamma_0$. We then obtain that

$$\begin{aligned}\int_\gamma & H(w,x)dw(1+O(e^{-AN})) \\ &= \alpha \int_{\gamma_0} e^{NF_{u_0,N}(\mathbf{z_N}+sN^{-\frac{1}{3}})-\alpha xs - x\epsilon/2} ds \\ &= \alpha \int_{\gamma_0} e^{F^{(3)}_{u_0,N}(\mathbf{z_N})\frac{s^3}{3!}-\alpha xs} e^{NF_{u_0,N}(\mathbf{z_N})-x\epsilon/2} ds(1+O(N^{-\frac{1}{3}})). \qquad (47)\end{aligned}$$

The last line is obtained by using the fact that

$$\begin{aligned}&|e^{NF_{u_0,N}(\mathbf{z_N}+sN^{-\frac{1}{3}})-NF_{u_0,N}(\mathbf{z_N})}-e^{F^{(3)}_{u_0,N}(\mathbf{z_N})\frac{s^3}{3!}}|\\&\quad\leq e^{-as^3}\frac{|s|^4\sup_{|z-\mathbf{z_N}|\leq\delta_0}|F^{(4)}_{u_0,N}(z)|}{N^{\frac{1}{3}}},\end{aligned}\tag{48}$$

for some constant $a > 0$. More detail can be found in [4, Section 3] and we do not develop the computations here.

Similarly we define $\Gamma_0 := \{te^{i2\pm\pi/3}, \epsilon \leq t \leq \delta_0 N^{1/3}\} \cup C'_\epsilon$ where C'_ϵ is the arc of circle centered at 0 joining $\epsilon e^{-2i\pi/3}$ and $\epsilon e^{2i\pi/3}$. This contour is again oriented from bottom to top.

$$\begin{aligned}\int_{\Gamma_1} G(z,x)dz &= \alpha\int_{|z-\mathbf{z_N}|\leq\delta_0} G(z,x)dz(1+O(e^{-AN}))\\&=\alpha\int_{\Gamma_0} e^{-NF_{u_0,N}(\mathbf{z_N}+tN^{-\frac{1}{3}})+\alpha xt+x\epsilon/2}dt(1+O(e^{-AN}))\\&=\alpha\int_{\Gamma_0} e^{-F^{(3)}_{u_0,N}(\mathbf{z_N})\frac{t^3}{3!}+\alpha xt}e^{-NF_{u_0,N}(\mathbf{z_N})+x\epsilon/2}dt(1+O(N^{-\frac{1}{3}})),\end{aligned}\tag{49}$$

where t describes the contour Γ_0 formed with the two half lines in the complex plane with angle $e^{\pm 2i\pi/3}$ with respect to the real axis. The contour is also oriented from bottom to top. We recall that α has been chosen as

$$\alpha = 2^{1/3}\frac{1}{|F^{(3)}_{u_0,N}(\mathbf{z_N})|^{1/3}}.$$

We then deduce that for $|x|, |y| \leq M_0$, one has that

$$\begin{aligned}\left|\frac{1}{2i\pi}\int_\gamma H(w,y)dw - Ai(y)e^{-y\epsilon/2}\right| &\leq \frac{C}{N^{\frac{1}{3}}},\\ \left|\frac{1}{2i\pi}\int_{\Gamma_1} G(z,y)dz - Ai(y)e^{y\epsilon/2}\right| &\leq \frac{C}{N^{\frac{1}{3}}}.\end{aligned}$$

We can now conclude to the asymptotic behavior of the rescaled correlation kernel $K^{(l)}_N(u,v)e^{q(y-x)N^{\frac{1}{3}}}$ when x and/or y are allowed to grow unboundedly positive. Indeed for this part of the kernel we do not need to bound x and y from above by $\epsilon_0 N^{2/3}$. As by construction the two contours Γ_1 and γ lie respectively to the left (resp. right) strictly of q, one can deduce (copying the arguments developed in [4, Section 3]) that there exist constants $C, c > 0$ such that

$$\begin{aligned}\left|\frac{1}{2i\pi}\int_\gamma H(w,y)dw - Ai(y)e^{-y\epsilon/2}\right| &\leq \frac{C}{N^{\frac{1}{3}}}e^{-cy},\\ \left|\frac{1}{2i\pi}\int_{\Gamma_1} G(z,y)dz - Ai(y)e^{y\epsilon/2}\right| &\leq \frac{C}{N^{\frac{1}{3}}}e^{-cy}.\end{aligned}\tag{50}$$

Note that (50) also holds true (modifying the constants C, c if needed) when $|x|, |y| \leq M_0$.

Last we need to consider the contribution of the contour $\Gamma_2 \cup \gamma$. We show that this contribution is negligible provided x and y are bounded from above by $\epsilon_0 N^{2/3}$ for some ϵ_0 small enough. Let us recall that $\Re F_u''(z) \geq 0$ for any z along Γ_2. Furthermore there exists $\eta > 0$ such that $\mathrm{dist}(\Gamma_2, \gamma) > \eta$. As a consequence the main contribution from Γ_2 comes from the closest point to $\mathbf{z_N}$, namely $\mathbf{z'_N}$. From this we deduce that

$$\left|\alpha \frac{N^{\frac{1}{3}}}{(2i\pi)^2} \int_{\Gamma_2} \int_\gamma e^{N(w-v)^2/2 - N(z-u)^2/2} \frac{e^{q(y-x)N^{\frac{1}{3}}}}{w-z} \prod_{i=1}^N \frac{w-y_i}{z-y_i} dw dz\right| \leq C e^{NF_{u_0,N}(\mathbf{z_N}) - NF_{u_0,N}(\mathbf{z'_N}) + q(y-x)N^{\frac{1}{3}}}, \tag{51}$$

for some constant $C > 0$. As $|y|, |x| \leq \epsilon_0 N^{2/3}$, we choose $\epsilon_0 > 0$ small enough so that there exists a constant $C' > 0$ so that

$$\Re(NF_{u_0,N}(\mathbf{z_N}) - NF_{u_0,N}(\mathbf{z'_N}) + q(y-x)N^{\frac{1}{3}}) < -C'N. \tag{52}$$

Combining (51), (52) and (50) then yields Proposition 4.2. □

4.3 Proof of Theorem 1.2

Consider a spike θ_{i_1} of multiplicity k_{i_1} such that $\int \frac{1}{(\theta_{i_1}-y)^2} d\nu(y) < 1$. Then θ_{i_1} makes k_{i_1} outliers separate from the bulk at $\rho(\theta_{i_1})$ asymptotically, with $\rho(z) := z + \int \frac{1}{z-y} d\nu(y)$. We recall that θ_{i_1} is such that $\mathrm{dist}(\rho(\theta_{i_1}), \mathrm{supp}(\mu_{sc} \boxplus \nu)) > 0$. Thus there exist (possibly) $\mathbf{z_N}$ and $\mathbf{w_N}$ such that $H_N(\mathbf{z_N})$ and $H_N(\mathbf{w_N})$ are respectively the right and left endpoints of the connected component of $\mathrm{supp}(\mu_{sc} \boxplus \mu_{A_N})$ which is respectively on the left hand side and right hand side of $\rho(\theta_{i_1})$ and we have $\mathbf{z_N} < \theta_{i_1} < \mathbf{w_N}$. If there is no connected component of $\mathrm{supp}(\mu_{sc} \boxplus \mu_{A_N})$ to the right respectively the left of $\rho(\theta_{i_1})$, we then set $\mathbf{w_N} = +\infty$, respectively $\mathbf{z_N} = -\infty$.

We first need some definitions to consider the asymptotic correlation functions close to an outlier. Let ρ_N be defined in (10). Let $c > 0$ be given (to be defined later). We set

$$u_0 := \rho_N(\theta_{i_1}), \quad u = u_0 + \frac{cx}{\sqrt{N}}, \quad v = u_0 + \frac{cy}{\sqrt{N}}.$$

Again we assume that x, y are bounded from below by $-M_0$ for some real number $M_0 > 0$. On the other side, x and y are not allowed to grow unboundedly. Let $\eta_1 > 0$ be given (small). We assume that $\eta_0 > 0$ is small enough so that

$$\rho_{\theta_{i_1}} + \eta_0 < \rho_N(\theta_i) - \eta_1, \quad \forall i \text{ s. t. } \theta_i > \theta_{i_1}, \quad \rho_{\theta_{i_1}} + \eta_0 < H_N(\mathbf{w_N}).$$

We assume that $x, y \leq \eta_0 N^{1/2}$. We now consider the asymptotics of the rescaled correlation kernel:

$$\frac{c}{\sqrt{N}} K_N(u, v).$$

Define

$$G_{u_0,N}(z) := \frac{z^2}{2} - u_0 z + \frac{1}{N}\left(\sum_{j:\, y_j<\theta_{i_1}} \ln(z-y_j)\right) + \frac{1}{N}\left(\sum_{j:\, y_j>\theta_{i_1}} \ln(y_j - z)\right). \tag{53}$$

We here set

$$c := \sqrt{G''_{u_0,N}(\theta_{i_1})} > 0.$$

Let K_H be the correlation kernel of a $k_{i_1} \times k_{i_1}$ GUE. We recall that K_H is the Christoffel Darboux kernel of some rescaled Hermite polynomials satisfying the orthogonality relationship $\int_{-\infty}^{\infty} p_m(x) p_n(x) e^{-\frac{1}{2}x^2} dx = \delta_{mn}$.

Proposition 4.3 *There exist constants $q, C,$ and $C' > 0$ such that for $x, y \in [-M_0, \eta_0 N^{1/2}]$*

$$\left| \frac{c}{\sqrt{N}} K_N(u, v) e^{qcN^{\frac{1}{2}}(y-x)} - K_H(x, y) \right| \leq \frac{Ce^{-C'(x+y)}}{N^{1/2}}.$$

Proof of Proposition 4.3 We again split the correlation kernel into two parts, by dividing the contour Γ into two parts. One contour, denoted by Γ_1 encircles the eigenvalues y_i such that $y_i \leq \theta_{i_1}$. The other contour Γ_2 then encircles all the eigenvalues y_j such that $y_j > \theta_{i_1}$. This is possible as we assume that spikes are independent of N. Note that Γ_1 can be chosen so that it lies to the left of $\theta_{i_1} + \eta N^{-1/2}$ for some small $\eta > 0$. Accordingly we define $K_N^{(l)}(u, v)$ and $K_N^{(r)}$ to be the corresponding contributions (from contours lying to the left or to the right of $\theta_{i_1} + \eta N^{-1/2}$) to the correlation kernel.

We first rewrite the singularity in the correlation kernel. Then, provided $\Re(w-z) > 0$, one has that

$$\frac{1}{w-z} = \int_{\mathbb{R}^+} dt_0 e^{-N^{\frac{1}{2}} ct_0(w-z)} cN^{\frac{1}{2}}.$$

Thus one can write that

$$\frac{c}{\sqrt{N}} K_N^{(l)}(u, v) = \frac{c^2 N}{(2i\pi)^2} \int_{\mathbb{R}^+} dt_0 \int_{\Gamma_1} \int_{\gamma} \prod_{i=1}^{N} \frac{w-y_i}{z-y_i} \times e^{N(\frac{w^2}{2} - wv) - N(\frac{z^2}{2} - zu) - N^{\frac{1}{2}} ct_0(w-z)} dw dz, \tag{54}$$

where γ is a line parallel to the y-axis not crossing Γ_1. We keep the other kernel unchanged:

$$\frac{c}{\sqrt{N}} K_N^{(r)}(u, v) = \frac{cN^{1/2}}{(2i\pi)^2} \int_{\Gamma_2} \int_{\gamma} \prod_{i=1}^{N} \frac{w - y_i}{z - y_i} \times e^{N(\frac{w^2}{2} - wv) - N(\frac{z^2}{2} - zu)} \frac{1}{w - z} dw dz. \tag{55}$$

Consider the rescaled correlation kernel $\frac{c}{\sqrt{N}} K_N(u, v) e^{qcN^{\frac{1}{2}}(y-x)}$ for some q to be defined. We now set, using the definition of $G_{u_0,N}$ given by (53):

$$\begin{aligned} H(w, y) &= c\sqrt{N} \left(\sqrt{N}\right)^{k_{i_1}} e^{NG_{u_0,N}(w) - N^{\frac{1}{2}} cy(w-q)} (w - \theta_{i_1})^{k_{i_1}}, \\ G(z, x) &= c \frac{\sqrt{N}}{\left(\sqrt{N}\right)^{k_{i_1}}} e^{-NG_{u_0,N}(z) + N^{\frac{1}{2}} cx(z-q)} \times (z - \theta_{i_1})^{-k_{i_1}}. \end{aligned} \tag{56}$$

Then one has that

$$\frac{c}{\sqrt{N}} K_N^{(l)}(u, v) e^{qcN^{\frac{1}{2}}(y-x)} = \int_{\mathbb{R}^+} dt_0 \int_{\gamma} dw \int_{\Gamma_1} dz H(w, y + t_0) G(z, x + t_0). \tag{57}$$

Note that the measure

$$\tilde{\nu}_N = \frac{1}{N - k_{i_1}} \sum_{j:\, y_j \neq \theta_{i_1}} \frac{1}{z - y_j}$$

still converges to ν. Let us define

$$\begin{aligned} &\tilde{v}_N : \mathbb{R} \mapsto \mathbb{R}, \quad \tilde{v}_N(x) = \inf \left\{ v \geq 0, \int \frac{d\tilde{\nu}_N(s)}{(x - s)^2 + v^2} > \frac{N}{N - k_{i_1}} \right\}, \\ &\tilde{U}_N = \{x \in \mathbb{R}, \tilde{v}_N(x) > 0\} \end{aligned}$$

and

$$\mathcal{C}'_N = \{x \pm i\tilde{v}_N(x), x \in \mathbb{R}\}.$$

In addition θ_{i_1} is a critical point of $G_{u_0,N}$, which is the leading term in the exponential term defining both G and H. An easy computation shows that $G''_{u_0,N}(\theta_{i_1}) > 0$. Furthermore one can check that there exist $\delta > 0$ and constants $c(\delta) > 0$, $M(\delta) > 0$ such that

$$\forall\ z, |z - \theta_{i_1}| \leq \delta, \ \ |G''_{u_0,N}(z)| \geq c(\delta), \quad \text{and} \quad |G^{(3)}_{u_0,N}(z)| \leq M(\delta).$$

In order to perform the asymptotic analysis of the correlation kernel, we now choose

$$q = \theta_{i_1} + \frac{\epsilon}{2cN^{\frac{1}{2}}}.$$

We start with the kernel $K_N^{(l)}$. We first consider the asymptotics of the function H. We first consider the case where $|x|, |y| \leq M_0$. The other case will be considered hereafter. Let $\epsilon > 0$ be small. Define $\gamma = \theta'_{i_1} + it, t \in \mathbb{R}$ oriented from bottom to top where $\theta'_{i_1} = \theta_{i_1} + \frac{\epsilon}{c} N^{-\frac{1}{2}}$. One has that

$$\frac{d}{dt}\Re(G_{u_0N}(\theta'_{i_1} + it)) = -t\left(1 - \frac{1}{N}\sum_{j:\, y_j \neq \theta_{i_1}} \frac{1}{|\theta'_{i_1} - y_j + it|^2}\right) \leq -Ct,$$

for some constant $C > 0$. This follows from the fact that the second derivative of $G_{u_0,N}$ does not vanish in a neighborhood of θ_{i_1} in particular. Note also that the variation of $G_{u_0,N}(\theta'_{i_1}) - G_{u_0,N}(\theta_{i_1})$ is of the order of $1/N$. We now use the same arguments as in Sect. 4.2.1. As we see just below, we can deform Γ_1 so that γ lies strictly to the right of Γ_1. Assuming this holds true, one gets that there exists a constant $A > 0$ such that

$$\int_\gamma H(w, y)dw$$
$$= c(\sqrt{N})^{k_{i_1}+1} \int_{|w-\theta'_{i_1}|\leq\delta} e^{NG_{u_0,N}(w) - N^{\frac{1}{2}}cy(w-q)} \times (w - \theta_{i_1})^{k_{i_1}}(1 + O(e^{-AN})).$$

Making the change of variables $w = \theta_{i_1} + i\frac{t}{c\sqrt{N}}$, and setting $\mathbb{R}_{\text{def}} = \mathbb{R} - i\epsilon$ one obtains that

$$\int_\gamma H(w, y)dw(1 + O(e^{-AN}))$$
$$= ce^{NG_{u_0,N}(\theta_{i_1})}e^{y\epsilon/2} \int_{\mathbb{R}_{\text{def}}} \frac{i}{c} e^{-\frac{t^2}{2} - yit} \left(i\frac{t}{c}\right)^{k_{i_1}} (1 + O(N^{-\frac{1}{2}}))$$
$$= e^{NG_{u_0,N}(\theta_{i_1})} \int_{\mathbb{R}_{\text{def}}} ie^{-\frac{t^2}{2} - y(it - \epsilon/2)} \left(\frac{it}{c}\right)^{k_{i_1}} (1 + O(N^{-\frac{1}{2}})).$$

We consider now the case where y can be as large as $\epsilon_0 N^{1/2}$. We use the fact that the contour γ remains to the right of q strictly. In particular, one can show that there exist constants $C, C' > 0$ such that

$$\left| \int_\gamma \frac{H(w, y)}{e^{NG_{u_0,N}(\theta_{i_1})}} dw - \int_{\mathbb{R}_{\text{def}}} ie^{-\frac{t^2}{2} - y(it - \epsilon/2)} \left(\frac{it}{c}\right)^{k_{i_1}} \right| \leq \frac{Ce^{-C'y}}{\sqrt{N}}. \tag{58}$$

We now turn to the asymptotics of $\int_{\Gamma_1} G(z, y)dz$. Similarly for the z contour, we use the following contour Γ_1 (see Fig. 3).

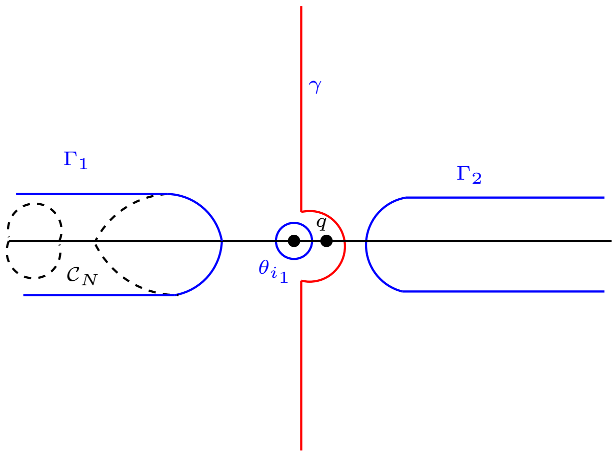

Fig. 3 The contour Γ and γ at a spike

First Γ_1 contains a circle of ray $\frac{\epsilon}{4cN^{\frac{1}{2}}}$ around θ_{i_1}. Γ_1 then has to encircle all the eigenvalues to the left of θ_{i_1}. Note that there exists $\eta > 0$

$$\sup\{x \in \tilde{U}_N, x < \theta_{i_1}\} =: \mathbf{w'_N} \leq \theta_{i_1} - \eta$$

and

$$\inf\{x \in \tilde{U}_N, x > \theta_{i_1}\} =: \mathbf{z'_N} \geq \theta_{i_1} + \eta.$$

Let then Z' be the first point along $\mathcal{C}'_N$ to the left of $\mathbf{w'_N}$ such that $\mathrm{Im}(Z')$ is a local maximum. Γ_1 then follows $\mathcal{C}'_N$ from $\mathbf{w'_N}$ to the left direction up to Z'. Then to the left of Z', Γ_1 follows the highest of the two curves $\mathcal{C}'_N$ and $Z' - x, x > 0$. The contour is completed by symmetry with respect to the real axis. Computing residues, one easily gets that the asymptotics for $G(z, y)$ splits into two parts

- the residue at θ_{i_1} that yields by a straightforward Taylor approximation:

$$e^{-\frac{\epsilon}{2}y} e^{-NG_{u_0,N}(\theta_{i_1})} \mathrm{Res}_{a=0} \left(\left(\frac{c}{a}\right)^{k_{i_1}} e^{-NG_{u_0}\left(\theta_{i_1} + \frac{a}{c\sqrt{N}}\right) + NG_{u_0,N}(\theta_{i_1}) + ya} \right).$$

- The contribution of the rest of the contour $\Gamma_1 \cap \{z, \in \mathbb{C}, \Re z < \theta_{i_1} - \eta\}$ which, by a small extension of the previous subsection, is in the order of

$$e^{-NG_{u_0,N}(\mathbf{w'_N})} << e^{-NG_{u_0,N}(\theta_{i_1})}.$$

This is also exponentially negligible in the large N limit.

To finish the asymptotic analysis of G, we show that the first term is indeed in the order of $e^{-NG_{u_0,N}(\theta_{i_1})}$. By a straightforward Taylor expansion one obtains that

$$\begin{aligned} e^{-\frac{\epsilon y}{2}} \Bigg| &\mathrm{Res}_{a=0}\left(\left(\frac{c}{a}\right)^{k_{i_1}} e^{N\left(G_{u_0,N}(\theta_{i_1})-G_{u_0,N}\left(\theta_{i_1}+\frac{a}{cN^{\frac{1}{2}}}\right)\right)+ya}\right) \\ &-\mathrm{Res}_{a=0}\left(\left(\frac{c}{a}\right)^{k_{i_1}} e^{ay-\frac{a^2}{2}}\right)\Bigg| \\ &\leq \frac{Ce^{-C'y}}{\sqrt{N}}, \end{aligned} \tag{59}$$

for some constants $C, C' > 0$. The exponential decay for large y follows again from the fact that the residue is computed on a circle of ray $\epsilon/4c$ lying to the left strictly of $\epsilon/2c$.

We now turn to the asymptotic analysis of $K_N^{(r)}(u, v)$. Let us define the contour Γ_2 as in the preceding section. Let Z be the first point along $\mathcal{C}'_N$ to the left of $\mathbf{z}'_{\mathbf{N}}$ such that $\mathrm{Im}(Z)$ is a local maximum. Γ_2 first follows the part $\mathcal{C}'_N$ lying to the right of $\mathbf{z}'_{\mathbf{N}}$ up to the moment where it reaches Z. Then Γ_2 is pursued to the right by following the highest of the two curves $\mathcal{C}'_N$ and $Z + x, x > 0$. Again it is completed by symmetry with respect to the real axis. It is an easy computation to check that $\Re G_{u_0}(z)$ achieves its minimum on Γ_2 at $\mathbf{z}'_{\mathbf{N}}$. The contour γ is chosen as before. Note that the function $\frac{1}{w-z}$ remains bounded along $\gamma \cup \Gamma_2$. We then deduce that

$$\begin{aligned} \left| \frac{c}{\sqrt{N}} K_N^{(r)}(u, v) e^{N^{1/2}c(y-x)q} \right| &= \leq C e^{N\Re G_{u_0,N}(\theta_{i_1}) - G_{u_0,N}(\mathbf{z}'_{\mathbf{N}}) + N^{1/2}c(y-x)q} \\ &\leq Ce^{-C'N}, \end{aligned} \tag{60}$$

provided ϵ_0 is small enough. Thus the kernel $\frac{c}{\sqrt{N}} K_N^{(r)}(u, v) e^{N^{1/2}c(y-x)q}$ converges uniformly to 0 on $[-M_0, \epsilon_0 N^{1/2}]$. Combining (58), (59) and (60) then yields Proposition 4.3 using the expression of the correlation functions of $k_{i_1} \times k_{i_1}$ GUE given in Section 4.3 of [4]. □

4.4 At a point where two connected components merge

Let now consider a point $u \in \mathrm{supp}(\mu_{sc} \boxplus \nu)$ such that the density p of $\mu_{sc} \boxplus \nu$ verifies

$$p(u) = 0, \quad p(x) > 0 \ \forall\, x \in [u - \epsilon/2, u + \epsilon/2] \backslash \{u\} \quad \text{for some } \epsilon > 0.$$

This means that the critical point $z_c(u)$ associated to $u = H(z_c(u))$ is unique, real and lies at the “intersection” of two complex curves (see Fig. 4 below). Because $z_c(u) \notin \mathrm{supp}(\nu)$, we deduce from Lemma 2.1 that

$$F''(z_c(u)) = 0 \ \text{ and that } \ F^{(3)}(z_c(u)) = 0.$$

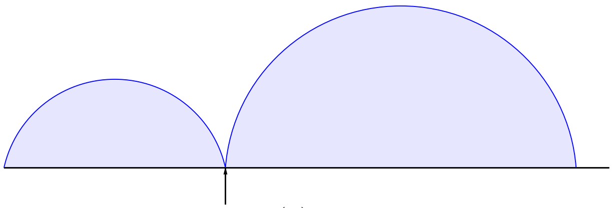

Fig. 4 A point in the bulk with vanishing density

The first order derivative which does not vanish at $z_c(u)$ is then the fourth one: $F^{(4)}(z_c(u)) < 0$. For the asymptotic exponential term F, $z_c(u)$ is a doubly degenerate critical point. Thanks to Proposition 3.3, one can transmit this double degeneracy to the true exponential term $F_{u,N}$. There exists a unique point $z_{c,N}$ in a η-neighborhood of z_c (for any $\eta > 0$) such that

$$F''_{u,N}(z_{c,N}) = F^{(3)}_{u,N}(z_{c,N}) = 0.$$

At such a point, one obviously has that

$$F^{(4)}_{u,N}(z_{c,N}) < 0.$$

Here $F_{u,N}$ is defined by (38) with $\mathbf{z_c} = z_{c,N}$. Set $u_0 = H_N(z_{c,N})$. We here show that the asymptotic correlation functions in the vicinity of u_0 are determined by the so-called Pearcey kernel defined by (13).

Proposition 4.4 *Set* $\kappa = |F^{(4)}(z_{c,N})|^{1/4}$. *Uniformly for* x, y *in a fixed compact interval, one has that*

$$\lim_{N\to\infty} \frac{\kappa}{N^{\frac{3}{4}}} K_N\left(u_0 + \frac{\kappa x}{N^{\frac{3}{4}}}, u_0 + \frac{\kappa y}{N^{\frac{3}{4}}}\right) = K_P(x, y).$$

Proof of Proposition 4.4 We start from the expression for the correlation kernel given in Proposition 4.1, where the contours are as shown on Fig. 5.

One has that $F_N^{(4)}(z_{c,N}) < 0$ and it is not difficult to see that, given $\delta > 0$ small, there exists a constant M such that $|F_N^{(5)}(z)| \leq M$ for all complex numbers z, such that $|z - z_{c,N}| \leq \delta$. From this we deduce that for any real t such that $|t| \leq \delta$

$$\left| F_{u,N}(z_{c,N} + te^{i\frac{\pi}{4}}) - F_{u,N}(z_{c,N}) + F^{(4)}_{u,N}(z_{c,N}) \frac{t^4}{4!} \right| \leq \frac{M|t|^5}{5!}.$$

Assume that $|t| \leq \delta$, then one has that

$$\Re(F_{u,N}(z_{c,N} + te^{i\frac{\pi}{4}}) - F_{u,N}(z_{c,N})) \geq |F_N^{(4)}(z_{c,N})| t^4/8!,$$

provided δ is small enough. This ensures that the (z-)contour made of two lines with direction $\pm\pi/4$ with the real axis is an ascent contour for $F_{u,N}$, at least in a δ neighborhood of $z_{c,N}$. To complete the z-contour, we need to encircle all the remaining

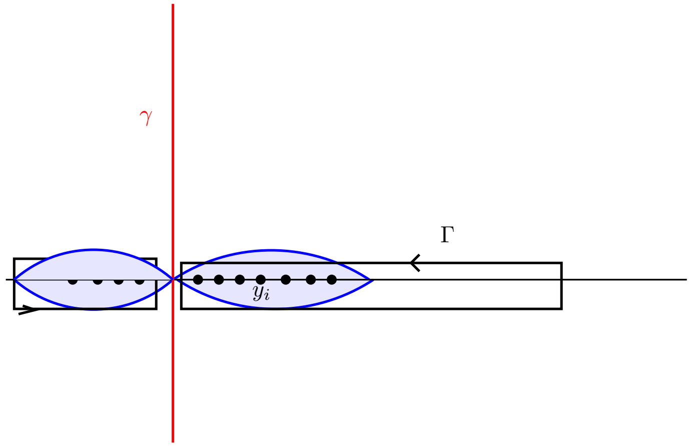

Fig. 5 Initial contours γ and Γ, which do not cross

eigenvalues. We pursue the contour as before. If $z_{c,N} + \delta e^{i\frac{\pi}{4}}$ (resp. $z_{c,N} + \delta e^{3i\frac{\pi}{4}}$) lies above $\mathcal{C}_N$, the contour goes parallely to the real axis to the right (resp. left) up to the moment of time one crosses the curve $\mathcal{C}_N$. Then it follows $\mathcal{C}_N$ to the right (resp. left) direction up to the moment where it crosses the line $\mathrm{Im} z = \delta\sqrt{2}/2$ and so on. If instead $z_{c,N} + \delta e^{i\frac{\pi}{4}}$ (resp. $z_{c,N} + \delta e^{3i\frac{\pi}{4}}$) lies below $\mathcal{C}_N$, then one first joins $\mathcal{C}_N$ along $z_{c,N} + \delta e^{i\frac{\pi}{4}} + it, t \geq 0$ (resp. $z_{c,N} + \delta e^{3i\frac{\pi}{4}} + it, t \geq 0$) and then follows $\mathcal{C}_N$ to the right (resp. left) direction (not going below the line $\mathrm{Im} z = \delta\sqrt{2}/2$). The contour is then completed by symmetry with respect to the real axis.

For the w contour it is an easy computation that the curve $z_{c,N} + it, t \in \mathbb{R}$ satisfies the descent assumption. Last, so that the w and z contours do not cross each other, we deform the z contour in a small neighbordhood of $z_{c,N}$ to the new contour Γ_0 as on Fig. 1.

We can now conclude to the asymptotic behavior of the kernel. We make the change of variables $w = z_{c,N} + sN^{-1/4}$, $z = z_{c,N} + tN^{-1/4}$, neglecting the part of the contour where $|w - z_{c,N}| \geq \delta$ or $|z - z_{c,N}| \geq \delta$.
One has that (up to a conjugation factor)

$$\begin{aligned} &\frac{1}{N^{\frac{3}{4}}} K_N\left(u + \frac{x}{N^{\frac{3}{4}}}, u + \frac{y}{N^{\frac{3}{4}}}\right) \\ &= \frac{1}{(2i\pi)^2} \int_{\Gamma_0} dt \int_{i\mathbb{R}} ds e^{F^{(4)}(z_{c,N})\frac{s^4 - t^4}{4!} - sy + tx} \frac{1}{s-t}(1 + O(N^{-\frac{1}{4}})), \end{aligned} \tag{61}$$

where we first neglected the parts of the contour lying at a distance $\delta > 0$ of $z_{c,N}$ and then performed a Taylor expansion, using the boundedness of the fifth derivative $F_{u,N}^{(5)}$ in a compact neighborhood of $z_{c,N}$. The last estimate holds uniformly for x, y in a fixed compact real interval. Then making the change of variables $s = |F^{(4)}(z_{c,N})|^{1/4} s'$ yields the desired result. □

References

1. Adler, M., Cafasso, M., van Moerbeke, P.: From the Pearcey to the Airy process. Electron. J. Probab. **16**(36), 1048–1064 (2011)
2. Adler, M., van Moerbeke, P.: PDEs for the Gaussian ensemble with external source and the Pearcey distribution. Commun. Pure Appl. Math. **60**(9), 1261–1292 (2007)
3. Anderson, G., Guionnet, A., Zeitouni, O.: An Introduction to Random Matrices. Cambridge University Press, Cambridge (2009)
4. Baik, J., Ben Arous, G., Péché, S.: Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. Ann. Probab. **33**(5), 1643–1697 (2005)
5. Baik, J., Nadakuditi, R.R.: Batch latency analysis and phase transitions for a tandem of queues with exponentially distributed service times. Queueing Syst. **78**(3), 255–285 (2014)
6. Baik, J., Silverstein, J.: Eigenvalues of large sample covariance matrices of spiked population models. J. Multivar. Anal. **97**(6), 1382–1408 (2006)
7. Belinschi, S.T., Bercovici, H.: A new approach to subordination results in free probability. J. Anal. Math. **101**, 357–365 (2007)
8. Biane, P.: On the free convolution with a semi-circular distribution. Indiana Univ. Math. J. **46**(3), 705–718 (1997)
9. Biane, P.: Processes with free increments. Math. Z. **227**(1), 143–174 (1998)
10. Brézin, E., Hikami, S.: Correlations of nearby levels induced by a random potential. Nuclear Phys. B **479**(3), 697–706 (1996)
11. Capitaine, M., Donati-Martin, C., Féral, D.: The largest eigenvalues of finite rank deformation of large Wigner matrices: convergence and nonuniversality of the fluctuations. Ann. Probab. **37**(1), 1–47 (2009)
12. Capitaine, M., Donati-Martin, C., Féral, D., Février, M.: Free convolution with a semi-circular distribution and eigenvalues of spiked deformations of Wigner matrices. Electron. J. Probab. **16**, 1750–1792 (2011)
13. Erdòs, L., Yau, H.-T., Yin, J.: Rigidity of eigenvalues of generalized Wigner matrices. Adv. Math. **229**(3), 1435–1515 (2012)
14. Erdòs, L., Yau, H.-T., Yin, J.: Universality of random matrices and local relaxation flow. Invent. Math. **185**(1), 75–119 (2011)
15. Hiai, F., Petz, D.: The semicircle law, free random variables and entropy. In: Mathematical Surveys and Monographs, Vol. 77. AMS, New York (2000)
16. Johansson, K.: Universality of the local spacing distribution in certain ensembles of Hermitian Wigner matrices. Commun. Math. Phys. **215**(3), 683–705 (2001)
17. Knowles, A., Yin, J.: The isotropic semicircle law and deformation of Wigner matrices. arXiv:1110.6449 (2011) (preprint, to appear in Commun Math Phys)
18. Knowles, A., Yin, J.: The outliers of a deformed Wigner matrix. arXiv:1207.5619 (2012) (preprint, to appear in Ann Prob)
19. Mehta, M.: Random Matrices, 2nd edn. Academic Press, San Diago (1991)
20. Nica, A., Speicher, R.: Lectures on the Combinatorics of Free Probability, vol. 13. Cambridge University Press, Cambridge (2006)
21. O'Rourke, S., Vu, V.: Universality of local eigenvalue statistics in random matrices with external source. arXiv:1308.1057 (2013)
22. Péché, S.: The largest eigenvalue of small rank perturbations of Hermitian random matrices. Probab. Theory Relat. Fields **134**(1), 127–173 (2006)
23. Pizzo, A., Renfrew, D., Soshnikov, A.: On finite rank deformations of Wigner matrices. Ann. Inst. Henri Poincaré Probab. Stat. **49**(1), 64–94 (2013)
24. Renfrew, D., Soshnikov, A.: On finite rank deformations of Wigner matrices II. Delocalized perturbations. arXiv:1203.5130 (2013)
25. Shcherbina, T.: On universality of bulk local regime of the deformed Gaussian unitary ensemble. Zh. Mat. Fiz. Anal. Geom. **5**(4), 396–433 (2009)
26. Shcherbina, T.: On universality of local edge regime for the deformed Gaussian unitary ensemble. J. Stat. Phys. **143**(3), 455–481 (2011)
27. Tao, T., Vu, V.: Random matrices: universality of local eigenvalue statistics up to the edge. Commun. Math. Phys. **298**(2), 549–572 (2010)
28. Tracy, C., Widom, H.: Level spacing distributions and the Airy kernel. Commun. Math. Phys. **159**, 33–72 (1994)

29. Tracy, C., Widom, H.: The Pearcey process. Commun. Math. Phys. **263**(2), 381–400 (2006)
30. Voiculescu, D.: Limit laws for random matrices and free products. Invent. Math. **104**, 201–220 (1991)
31. Voiculescu, D.: The analogues of entropy and of Fisher's information measure in free probability theory I. Commun. Math. Phys. **155**(1), 71–92 (1993)
32. Voiculescu, D.: Lectures on free probability theory, pp. 279–349. In: Lectures on Probability Theory and Statistics, Saint-Flour 1998. Lecture Notes in Mathematics, Vol. 1738. Springer, New York (2000)
33. Voiculescu, D.V., Dykema, K., Nica, A.: CRM Monograph Series No. 1. Free random variables. American Mathematical Society, Providence (1992)