Probab. Theory Relat. Fields 108, 517–542 (1997)

Probability Theory and Related Fields

Large deviations for Wigner's law and Voiculescu's non-commutative entropy

G. Ben Arous[1], **A. Guionnet**[2]

[1] URA 762, CNRS, DMI, Ecole Normale Superieure, F-75230 Paris, France
[2] URA 743, CNRS, Bat. 425, Université de Paris Sud, F-91405 Orsay, France

Received: 3 April 1995 / In revised form: 14 December 1996

Summary. We study the spectral measure of Gaussian Wigner's matrices and prove that it satisfies a large deviation principle. We show that the good rate function which governs this principle achieves its minimum value at Wigner's semicircular law, which entails the convergence of the spectral measure to the semicircular law. As a conclusion, we give some further examples of random matrices with spectral measure satisfying a large deviation principle and argue about Voiculescu's non commutative entropy.

Mathematics Subject of Classification: 60F10, 15A18, 15A52

1 Introduction

If $X^N = \left(X^N_{ij}\right)$ is a $N \times N$ Wigner's matrix, that is a $N \times N$ real symmetric random matrix with centered independent entries with covariance $(1/2N)$, then the celebrated Wigner theorem asserts that the spectral measure

$$\hat{\mu}^N = \frac{1}{N}\sum_{i=1}^{N} \delta_{\lambda_i}$$

converges almost surely to the semicircular law

$$\sigma_1(dx) = \frac{1}{\pi}\mathbb{1}_{|x|<\sqrt{2}}\sqrt{(2-x^2)}dx \ .$$

Recently, Voiculescu showed how this theorem can be understood in the framework of his free probability theory (see [13], Theorem 2.2). In [14], he introduced a concept of non commutative entropy

$$\Sigma(\mu) = \int\int \log|x-y| d\mu(x)\ d\mu(y)$$

and gave some heuristics to relate it to large random matrices. For instance, he argued that Σ could be seen as a normalized limit of relative entropies of the law of the eigenvalues of large random matrices.

We will show here a large deviation principle for the law of the spectral measure $\hat{\mu}^N$ of the Gaussian Wigner's matrix with a good rate function where Σ indeed plays the role of the relative entropy in Sanov's Theorem. We will as well recall some rigorous aspects of Voiculescu's heuristics (see the end of section 5).

More precisely, if X^N is the $N \times N$ symmetric random matrix with independent entries and with law given by

$$\text{Law of } \left(X_{ij}^N\right) = N\left(0, \frac{1}{2N}\right) \text{ if } i < j \qquad \text{Law of} \left(X_{ii}^N\right) = N\left(0, \frac{1}{N}\right) , \tag{1}$$

then, X^N has N real eigenvalues $(\lambda_i)_{1 \le i \le N}$ and its spectral measure $\hat{\mu}^N = \frac{1}{N}\sum_{i=1}^N \delta_{\lambda_i}$ is a probability measure on $\mathbb{R}$. In the following, we will denote $\mathcal{M}_1^+(\mathbb{R})$ the space of probability measure on $\mathbb{R}$ and will endow $\mathcal{M}_1^+(\mathbb{R})$ with its usual weak topology. We now can state one of the main results of this paper.

Theorem 1.1

1) Let $I_1(\mu) = \frac{1}{2}\left(\int x^2 \, d\mu(x) - \Sigma(\mu)\right) - \frac{3}{8} - \frac{1}{4}\log 2$.
Then:

- *a.* I_1 *is well defined on* $\mathcal{M}_1^+(\mathbb{R})$ *and takes its values in* $[0, +\infty]$.
- *b.* $I_1(\mu)$ *is infinite as soon as* μ *satisfies one of the following conditions:*
  b.1: $\int x^2 d\mu(x) = +\infty$.
  b.2: There exists a subset A *of* R *of positive* μ *mass but null loga rithmic capacity, i.e. a set* A *such that:*
  $\mu(A) > 0 \quad \gamma(A) := \exp\left\{-\inf_{\nu \in \mathcal{M}_1^+(A)} \int\int \log \frac{1}{|x-y|} d\nu(x) d\nu(y)\right\} = 0$.
- *c.* I_1 *is a good rate function, i.e* $\{I_1 \le M\}$ *is a compact subset of* $\mathcal{M}_1^+(\mathbb{R})$ *for* $M \ge 0$.
- *d.* I_1 *is a convex function on* $\mathcal{M}_1^+(\mathbb{R})$.
- *e.* I_1 *achieves its minimum value at a unique probability measure on* $\mathbb{R}$ *which is described as the Wigner's semicircular law* σ_1.

2) The law of the spectral measure $\hat{\mu}^N = \frac{1}{N}\sum_{i=1}^N \delta_{\lambda_i}$ *satisfies a full large deviation principle with good rate function* I_1 *and in the scale* N^2, *that is*

for any open subset O *of* $\mathcal{M}_1^+(\mathbb{R})$,

$$\underline{\lim}_{N\to\infty} \frac{1}{N^2} \log P\left(\frac{1}{N}\sum_{i=1}^N \delta_{\lambda_i} \in O\right) \ge -\inf_O I_1,$$

for any closed subset F *of* $\mathcal{M}_1^+(\mathbb{R})$,

$$\overline{\lim}_{N\to\infty} \frac{1}{N^2} \log P\left(\frac{1}{N}\sum_{i=1}^N \delta_{\lambda_i} \in F\right) \le -\inf_F I_1.$$

This result shows that the convergence of the spectral measure to the Wigner's semicircular law σ_1 is exponentially fast in the sense that, for any open neighborhood $U(\sigma_1)$ of σ_1, we find that:

$$\overline{\lim}_{N\to\infty} \tfrac{1}{N^2} \log P\left(\tfrac{1}{N}\sum_{i=1}^{N} \delta_{\lambda_i} \in U(\sigma_1)^c\right) < 0 \ .$$

Let us also mention that some results are known about the fluctuations in Wigner's Theorem; Bai [2] proved a rate of $0(N^{-1/4})$ (in a much more general setting than ours since he does not impose that the entries are Gaussian). Bai also remarked that the fluctuations are expected to be on the scale $0(N^{-1})$.

Furthermore, Theorem 1.1 shows that the non commutative entropy Σ is in fact related to a commutative entropy i.e the rate function of a large deviation principle. We will see in section 5 that this is valid for more general random matrices. We want to recall another relation between Σ and a commutative entropy, more precisely with a specific entropy. This interpretation is due to Voiculescu. Namely, we will recall in section 5 that:

Theorem 1.2 *Let $p\colon \mathbb{R} \to \mathbb{R}$ be any diffeomorphism such that there exists positive real numbers ϵ and M so that*

$$\epsilon < \left\| (p^{-1})' \right\|_\infty < M \tag{2}$$

and a positive real number α, $\alpha > 2$, such that

$$\sup_N E\left[\left|\frac{1}{N}\sum_{i=1}^{N} p(X^N)_{ii}\right|^{\alpha}\right] < +\infty \ , \tag{3}$$

then, if $Q^N_{(p)}$ is the law of the eigenvalues of $p(X^N)$ and $I(\mu|v)$ the relative entropy of μ with respect to μ, i.e:

$$I(\mu|v) = \begin{cases} \int \log(\frac{d\mu}{dv}) dv & \text{if } \mu \ll v, \\ +\infty & \text{otherwise,} \end{cases}$$

there exists a finite constant C such that

$$\Sigma(\sigma_1 \circ p^{-1}) = 2 \lim_{N\to\infty} \frac{1}{N^2} I\left(Q^N_{(p)} | Q^N_1\right) + \int p(x)^2 d\sigma_1(x) + C \ .$$

To prove Theorem 1.1, we use the fact that the law of the spectrum can be described completely using the invariance by rotation of the law of X^N. Indeed, it is well known that the law Q^N_1 of the eigenvalues of X^N is given by

$$Q^N_1(d\lambda_1, \cdots, d\lambda_N) = \frac{1}{Z^N_1} \prod_{1\le i<j\le N} |\lambda_i - \lambda_j| \exp\left\{-\frac{1}{2} N \sum_{i=1}^{N} \lambda_i^2\right\} \prod_{i=1}^{N} d\lambda_i \ ,$$

where Z_1^N is the normalizing constant

$$Z_1^N = \int \cdot \int \prod_{1 \le i < j \le N} |\lambda_i - \lambda_j| \exp\left\{ -\frac{1}{2} N \sum_{i=1}^N \lambda_i^2 \right\} \prod_{i=1}^N d\lambda_i \ .$$

In fact, we will prove a more general result than Theorem 1.1.
Let

$$Q_\beta^N(d\lambda_1, \cdots, d\lambda_N) = \frac{1}{Z_\beta^N} \prod_{1 \le i < j \le N} |\lambda_i - \lambda_j|^\beta \exp\left\{ -\frac{1}{2} N \sum_{i=1}^N \lambda_i^2 \right\} \prod_{i=1}^N d\lambda_i,$$

where β is a positive real number and Z_β^N is the normalizing constant

$$Z_\beta^N = \int \cdot \int \prod_{1 \le i < j \le N} |\lambda_i - \lambda_j|^\beta \exp\left\{ -\frac{1}{2} N \sum_{i=1}^N \lambda_i^2 \right\} \prod_{i=1}^N d\lambda_i \ .$$

Then:

Theorem 1.3

1) Let $I_\beta(\mu) = \frac{1}{2}\left(\int x^2 d\mu(x) - \beta\Sigma(\mu) + \frac{\beta}{2}\log\frac{\beta}{2} - \frac{3}{4}\beta \right)$. Then:

a. I_β is well defined on $\mathscr{M}_1^+(\mathbb{R})$ and takes its values in $[0, +\infty]$.
b. $I_\beta(\mu)$ is infinite as soon as μ satisfies one of the following conditions:
b.1: $\int x^2 d\mu(x) = +\infty$.
b.2: There exists a subset A of $\mathbb{R}$ of positive μ mass but null logarithmic capacity.
c. I_β is a good rate function.
d. I_β is a convex function on $\mathscr{M}_1^+(\mathbb{R})$.
e. I_β achieves its minimum value at a unique probability measure which is described as the semi circular law:

$$\sigma_\beta(dx) = \frac{1}{\beta\pi} 1\!\!1_{|x| < \sqrt{2\beta}} \sqrt{(2\beta - x^2)} dx \ .$$

2) The law of the spectral measure $\hat{\mu}^N$ under Q_β^N satisfies a full large deviation principle with good rate function I_β, i.e:

for any open subset O of $\mathscr{M}_1^+(\mathbb{R})$,
$\varliminf_{N\to\infty} \frac{1}{N^2} \log Q_\beta^N \left(\frac{1}{N} \sum_{i=1}^N \delta_{\lambda_i} \in O \right) \ge -\inf_O I_\beta,$

for any closed subset F of $\mathscr{M}_1^+(\mathbb{R})$,
$\varlimsup_{N\to\infty} \frac{1}{N^2} \log Q_\beta^N \left(\frac{1}{N} \sum_{i=1}^N \delta_{\lambda_i} \in F \right) \le -\inf_F I_\beta.$

This generalization applies to other Gaussian matrices than Wigner's matrices since Q_β^N is the law of the eigenvalues of the Gaussian matrices of the symplectic ensemble when $\beta = 4$ and of the unitary ensemble when $\beta = 2$ (see [9]). Moreover, this generalization might be of some interest for other problems than large random matrices, for instance for particles interacting via a two dimensional Coulomb force.

We must point out that Theorem 1.1 has already been stated by T. Chan in [4]. Unfortunately, we do not understand the proof (see Remark 3.5).

It is the reason why we give here a proof of this large deviation principle, proof which appears to be quite simple. Indeed, we do not need precise large deviations estimates, as one could guess in view of the singularities of the density, mainly because of the scale N^2 of this large deviation principle. Furthermore, a crucial point of the proof is given by the results of Selberg which have been recently recalled by Mehta (see [9], Ch. 17) and which allow us to prove the convergence of the free energy $(1/N^2)\log Z_\beta^N$ and to compute its limit.

The organization of the paper is as follows:

In section 2 we study the rate function I_β and prove Theorem 1.3.(1).

In section 3 we state and prove a weak large deviation principle.

In section 4 we prove that the law of the empirical measure under Q_β^N is exponentially tight, which ends, with section 3, the proof of the full large deviation principle stated in Theorem 1.3.(2).

In section 5, we show how Theorem 1.1 can be used to study the spectral measure of more general sequences of large random matrices than $(X^N)_{N\in\mathbb{N}}$ by means of Laplace's method or contraction principle. We then derive Voiculescu's heuristics.

We thank A. Ancona for useful conversations about Potential Theory, P. Gerard for helping us to go through tricky integrals and O. Zeitouni for crucial remarks about the uniqueness of the minima. Alice Guionnet wishes to thank the Courant Institute for welcoming her for most of the research period, and in particular S. R. S. Varadhan and H. T. Yau.

2 Study of the rate function I_β

The aim of this section is first to show that I_β is a good rate function so that it achieves its minimum value and then prove that this minimum value is achieved at a unique probability measure which is characterized as a semicircular law.

2.1 I_β is a good rate function

To prove that I_β is a good rate function (i.e Theorem 1.3.1), we shall use the obvious identification between the definition of I_β given in Theorem 1.3 and the new definition described below:
Let Δ denote the diagonal of $\mathbb{R}\times\mathbb{R}$:

$$\Delta = \{(x,y)\in\mathbb{R}\times\mathbb{R} |\ x = y\ \}$$

Let f_β be the real function on $\mathbb{R} \times \mathbb{R}$ with value in $\mathbb{R} \cup \{+\infty\}$ defined by:

$$f_\beta(x,y) = \frac{1}{2}\left(x^2 + y^2\right) - \beta \log|y - x| \qquad \text{if } (x,y) \in \mathbb{R} \times \mathbb{R}\backslash\Delta$$
$$= +\infty \qquad \text{on } \Delta.$$

Then, if we define, for any probability measure μ on $\mathbb{R}$:

$$H_\beta(\mu) = \frac{1}{2}\int\int_{\mathbb{R}\times\mathbb{R}} f_\beta(x,y)d\mu(x)d\mu(y) \ ,$$

and F the function on $\mathbb{R}^+$ such that:

$$F(\beta) = \frac{\beta}{4}\log\frac{\beta}{2} - \frac{3}{8}\beta \ ,$$

I_β is equal to

$$I_\beta(\mu) = H_\beta(\mu) + F(\beta) \ . \tag{4}$$

Hence, we now prove that H_β is well defined and has compact level sets which will prove Theorem 1.3.1), except for the positivity of I_β. We will not prove this last point here since the positivity of I_β can be viewed as a consequence of the statement of the large deviation lower bound of Theorem 1.3, which will be proven independently.

Property 2.1

(1) H_β is well defined on $\mathscr{M}_1^+(\mathbb{R})$ and takes its values in $[(\beta/4)\ (1 - \log 2\beta), +\infty]$.

(2) $H_\beta(\mu)$ is infinite as soon as μ satisfies one of the following conditions:
2.1: $\int x^2 d\mu(x) = +\infty$.
2.2: There exists a subset A of $\mathbb{R}$ of positive μ mass but null logarithmic capacity, i.e.

$$\mu(A) > 0 \ \gamma(A) = \exp\left\{ - \inf_{\nu \in \mathscr{M}_1^+(A)} \int\int \log\frac{1}{|x-y|} d\nu(x)d\nu(y)\right\} = 0 \ .$$

Moreover, if μ has no atom,

$$H_\beta(\mu) = \int\int_{x<y} f_\beta(x,y)d\mu(x)d\mu(y). \tag{5}$$

(3) $\{H_\beta \leq M\}$ is a compact subset of $\mathscr{M}_1^+(\mathbb{R})$ for any real number M.
(4) H_β is a strictly convex function on $\mathscr{M}_1^+(\mathbb{R})$.

Remark 2.2. According to the lower bound stated in Theorem 3.2, $(\beta/4)$ $(1 - \log 2\beta)$ is not a sharp lower bound of H_β, but in fact the minimum value of H_β is $-F(\beta) = (\beta/4)((3/2) - \log(\beta/2))$.

Proof. (1) To prove that H_β is well defined, it is clearly enough to prove that f_β is a measurable function which is bounded from below. But it is clear that f_β is a continuous function. Moreover, the following decomposition holds for any $(x,y) \in \mathbb{R} \times \mathbb{R}\backslash\Delta$

$$f_\beta(x,y) = \frac{1}{2}\left((x^2+y^2) - \beta\log(x^2+y^2)\right) + \beta\log\frac{\sqrt{(x^2+y^2)}}{|y-x|} \quad . \tag{6}$$

But, a trivial computation shows that, for any positive real number z,

$$z - \beta\log z \geq \beta(1-\log\beta) \tag{7}$$

and it is well known that, for any $(x,y) \in \mathbb{R}\times\mathbb{R}\backslash\Delta$:

$$\frac{\sqrt{(x^2+y^2)}}{|y-x|} \geq \frac{1}{\sqrt{2}} \quad . \tag{8}$$

Hence, we conclude, using (6), that, for any $(x,y) \in \mathbb{R}^2$,

$$f_\beta(x,y) \geq (\beta/2)(1-\log 2\beta) \quad .$$

This lower bound shows that H_β is well defined and is bounded from below by $(\beta/4)(1-\log 2\beta)$.

(2) Let us first notice that, since x^2+y^2 grows to infinity much faster than $\log|x-y|$, it is clear that $H_\beta(\mu) = +\infty$ for any probability measure μ such that $\int x^2 d\mu(x)$ is infinite.

Moreover, since f_β is bounded from below by $m_\beta = (\beta/2)(1-\log 2\beta)$, we have, for any probability measure μ and any measurable subset B of $\mathbb{R}^2$, that

$$H_\beta(\mu) \geq \iint_B f_\beta(x,y)d\mu(x)d\mu(y) + (1-\mu^{\otimes 2}(B))m_\beta \tag{9}$$

But on the other hand, if we take $B = A\times A$,

$$\begin{aligned}\iint_B f_\beta(x,y)d\mu(x)d\mu(y) &\geq -\beta\iint_B \log|x-y|d\mu(x)d\mu(y)\\ &\geq -\beta\mu(A)^2\log(\gamma(A)) \quad ,\end{aligned}$$

so that, if μ fulfills 2.2 for some A, then $H_\beta(\mu) = +\infty$.

As a special case, it is clear that $H_\beta(\mu)$ is infinite as soon as μ has an atom. Finally, if μ has no atom, $\mu^{\otimes 2}(\Delta) = 0$, so that one finds the new definition (5) of $H_\beta(\mu)$ thanks to the symmetry of f_β.

(3) To prove the third point, we first define, for any positive real number M, a function f_β^M on $\mathbb{R}\times\mathbb{R}$ by:

$$f_\beta^M(x,y) = \min\{f_\beta(x,y), M\} \quad .$$

Then, it is obvious that f_β^M is a bounded continuous function. As a consequence,

$$H_\beta^M(\mu) = \frac{1}{2}\iint_{\mathbb{R}\times\mathbb{R}} f_\beta^M(x,y)d\mu(x)d\mu(y) \tag{10}$$

is a bounded continuous function on $\mathscr{M}_1^+(\mathbb{R})$.

Moreover, f_β^M grows to f_β pointwise so that, by monotone convergence Theorem, for any probability measure $\mu \in \mathscr{M}_1^+(\mathbb{R})$, $H_\beta^M(\mu)$ grows to $H_\beta(\mu)$.

Thus, H_β is lower semi continuous. In other words, for any real number L, $\{H_\beta \le L\}$ is a closed subset of $\mathcal{M}_1^+(\mathbb{R})$.

Hence, to prove that $\{H_\beta \le L\}$ is compact, it is enough to show that $\{H_\beta \le L\}$ is included in a compact subset of $\mathcal{M}_1^+(\mathbb{R})$. But, for any positive real number A large enough so that $h_\beta(z) = z - \beta \log z$ is positive and increases on $[2A^2, +\infty[$ (i.e $2A^2 - \beta \log 2A^2 > 0$ and $A \ge \sqrt{(\beta/2)}$), for any probability measure $\mu \in \mathcal{M}_1^+(\mathbb{R})$,

$$\begin{aligned}
[\mu([-A,+A]^c)]^2 &= \mu^{\otimes 2}(|x| > A, |y| > A) \\
&\le \mu^{\otimes 2}\big(x^2 + y^2 > 2A^2\big) = \mu^{\otimes 2}\big(h_\beta\big(x^2+y^2\big) > h_\beta\big(2A^2\big)\big) \\
&\le \frac{1}{h_\beta(2A^2)} \int\int \big(\big(x^2+y^2\big) - \beta \log\big(x^2+y^2\big)\big) d\mu(x) d\mu(y) \\
&\le \frac{1}{h_\beta(2A^2)} \left(4H_\beta(\mu) + \frac{\beta}{2}\log 2\right) \qquad \text{by (6) and (8)} \ .
\end{aligned}$$

As a consequence, there exists an integer number n_β^0 such that:

$$\{H_\beta \le L\} \subset \cap_{n \ge n_\beta^0} \left\{ \mu \in \mathcal{M}_1^+(\mathbb{R}) \mid \mu([-n,n]^c) \le \sqrt{\frac{2L + (\beta/4)\log 2}{n^2 - \beta \log 2n}} \right\} \ . \tag{11}$$

Since the subset of $\mathcal{M}_1^+(\mathbb{R})$ in the right hand side of (11) is compact, we conclude that $\{H_\beta \le L\}$ is compact.

(4) It is clearly enough to prove that Σ is a concave function to get that H_β is convex. This point can be understood thanks to the following equality:

$$\log|x-y| = \int_0^\infty \frac{1}{2t}\left(\exp\left\{-\frac{1}{2t}\right\} - \exp\left\{-\frac{|x-y|^2}{2t}\right\}\right) dt \ , \tag{12}$$

which can be checked easily. Let us then define, for integers numbers $T > 1$, a function Σ_T such that, for any $\mu \in \mathcal{M}_1^+(\mathbb{R})$,

$$\Sigma_T(\mu) \equiv \int\int\int_{\frac{1}{T}}^{T} \frac{1}{2t}\left(\exp\left\{-\frac{1}{2t}\right\} - \exp\left\{-\frac{|x-y|^2}{2t}\right\}\right) dt d\mu(x) d\mu(y) \ .$$

Then, since

$$v_t(x,y) \equiv \exp\left\{-\frac{1}{2t}\right\} - \exp\left\{-\frac{|x-y|^2}{2t}\right\}$$

has a constant sign on $\{(x,y) \in \mathbb{R}^2 \mid |x-y| \le 1\}$ and on its complement $\{(x,y) \in \mathbb{R}^2 \mid |x-y| > 1\}$, it is not hard to see that the monotone convergence theorem implies that, for any $\mu \in \mathcal{M}_1^+(\mathbb{R})$:

$$\lim_{T \uparrow \infty} \Sigma_T(\mu) = \Sigma(\mu) \ .$$

Moreover, Fubini Theorem applied to Σ_T gives:

$$\Sigma_T(\mu) = \int_{\frac{1}{T}}^{T} \frac{1}{2t} \exp\left\{-\frac{1}{2t}\right\} dt - \int_{\frac{1}{T}}^{T} \frac{1}{2t}\left(\int\int \exp\left\{-\frac{|x-y|^2}{2t}\right\} d\mu(x) d\mu(y)\right) dt \ . \tag{13}$$

It is well known that the last function in the right hand side of (13) is concave. In fact, one can even compute:

$$\begin{aligned}&\int\int \exp\left\{-\frac{|x-y|^2}{2t}\right\} d\mu(x) d\mu(y) \\ &= \sqrt{\frac{t}{2\pi}} \int_{-\infty}^{+\infty} \left|\int \exp\{i\lambda x\} d\mu(x)\right|^2 \exp - \left\{\frac{1}{2} t\lambda^2\right\} d\lambda \ .\end{aligned}$$

Therefore, Σ_T is a concave function so that its limit Σ is concave.

Furthermore, if μ_1 and μ_2 are two probability measures in $\{v \in \mathcal{M}_1^+(\mathbb{R}) / \Sigma(v) > -\infty\}$, our computation also shows that

$$\Sigma(\mu_1 - \mu_2) = -\int_0^{\infty} \frac{1}{2t}\left(\int\int \exp\left\{-\frac{|x-y|^2}{2t}\right\} d(\mu_1 - \mu_2)(x) d(\mu_1 - \mu_2)(y)\right) dt$$

which is obviously negative if $\mu_1 \neq \mu_2$. Therefore, H_β is in fact strictly convex. □

2.2 Study of the minima of the rate function I_β

In this part, we prove Theorem 1.3.(1).e, i.e:

Theorem 2.3 *I_β achieves its minimum value at a unique probability measure σ_β, the semi circular law given by:*

$$\sigma_\beta(dx) = \frac{1}{\beta\pi} \mathbb{1}_{|x|<\sqrt{2\beta}} \sqrt{(2\beta - x^2)} dx \ .$$

This theorem is in fact already proved in the literature. Indeed, the uniqueness result is proved in [8] and the characterization of the minimum as the semicircular law is due to Wigner's theorem (at least for $\beta = 1$). For the sake of completeness, we would like to prove that σ_β is a minimum of I_β independently of Wigner's Theorem. To this end, let us first recall theorem 2. 3. of [8] (applied to our setting where what the authors called admissible weight function w is $w(x) = (1/\beta)x^2$) which gives the uniqueness of the minima of I_β but as well some characterization of this minimum.

Theorem 2.4

(a) There exists a unique probability measure μ_β on $\mathbb{R}$ such that:

$$I_\beta(\mu_\beta) = 0 \ .$$

(b) The support $\mathscr{S}_\beta$ of μ_β is compact and with positive logarithmic capacity.
(c) The inequality

$$\beta \int \log|x-y| d\mu_\beta(y) \leq \frac{1}{2}x^2 - 2\inf H_\beta + \frac{1}{2}\int y^2 d\mu_\beta(y)$$

holds for any real number x except on a set A with null logarithmic capacity, i.e such that:

$$\gamma(A) = \exp\left\{ \sup_{\nu \in \mathscr{M}_1^+(A)} \int\int \log|x-y| d\nu(x) d\nu(y) \right\} = 0 \ .$$

(d) The inequality

$$\beta \int \log|x-y| d\mu_\beta(y) \geq \frac{1}{2}x^2 - 2\inf H_\beta + \frac{1}{2}\int y^2 d\mu_\beta(y)$$

holds for all x in $\mathscr{S}_\beta$.

Let us mention about the proof that the uniqueness of the minima is a direct consequence of the strict convexity of I_β (see property 2.1.4). The characterization of the minimum is not difficult but necessitates to pay special attention on the sets with null logarithmic energy.

As a consequence of Theorem 2.4. c) and d), the minimum of I_β is characterized by:

Lemma 2.5 *μ minimizes I_β if and only if for μ-almost all x:*

$$\beta \int \log|x-y| d\mu(y) = \frac{1}{2}x^2 - 2\inf H_\beta + \frac{1}{2}\int y^2 d\mu(y) \ . \tag{14}$$

Proof. Indeed, we saw in Theorem 1.3.b.2 that any probability measure μ such that $I_\beta(\mu)$ is finite does not put mass on sets of zero logarithmic capacity. Thus Theorem 2.4 c) and d) imply that the minimum of I_β satisfies (14).

Reversely, if μ satisfies (14), then it is clear that $2H_\beta(\mu) = \int\int f_\beta(z,y) d\mu(z)d\mu(y) = 2\inf H_\beta$, so that μ minimizes H_β, that is I_β. Hence, the proof of Lemma 2.5 is complete. □

As a consequence, we only need to show that σ_β satisfies (14) to achieve the proof of Theorem 2.3. Let us first notice that it is enough to concentrate on the case where $\beta = 1$ since:

Corollary 2.6 *μ_β minimizes I_β iff the probability measure $\tilde{\mu}_\beta$ defined by:*

$$\tilde{\mu}_\beta(x \in .) = \mu_\beta\left(\sqrt{\beta}^{-1} x \in .\right)$$

minimizes I_1.

Proof. Indeed, with this definition of $\tilde{\mu}_\beta$, Lemma 2.5 shows that μ_β minimizes I_β iff for $\tilde{\mu}_\beta$- almost all x, we have:

$$\frac{\beta}{2}x^2 + \frac{\beta}{2}\int y^2 d\widetilde{\mu}_\beta(y) - \beta\int \log\left|\sqrt{\beta}x - \sqrt{\beta}y\right| d\widetilde{\mu}_\beta(y) = 2\inf H_\beta \ .$$

Since $\inf H_\beta = -F(\beta)$ (see Remark 2.2) and according to the definition of $F(\beta)$, we deduce

$$\begin{aligned}\frac{1}{2}x^2 + \frac{1}{2}\int y^2 d\widetilde{\mu}_\beta(y) - \int \log|x-y| d\widetilde{\mu}_\beta(y) &= \frac{2}{\beta}\inf H_\beta + \log\sqrt{\beta} \\ &= -\frac{2}{\beta}F(\beta) + \log\sqrt{\beta} = \frac{3}{4} - \frac{1}{2}\log\frac{1}{2} \\ &= -2F(1) = 2\inf H_1,\end{aligned}$$

so that, again by Lemma 2.5, this is equivalent to say that $\widetilde{\mu}_\beta$ minimizes I_1. □

Therefore the problem boils down to show that, for any x in $[-\sqrt{2}, +\sqrt{2}]$:

$$\frac{1}{2}x^2 + \frac{1}{2}\int y^2 d\sigma_1(y) - \int \log|x-y| d\sigma_1(y) = 2\inf H_1 \ .$$

which resumed to prove (since $\int y^2 d\sigma_1(y) = (1/2)$ and $2\inf H_1 = (3/4) + (1/2)\log 2$) that

Lemma 2.7 *For any x in $[-\sqrt{2}, +\sqrt{2}]$,*

$$\int \log|x-y| d\sigma_1(y) = \frac{1}{2}x^2 - \frac{1}{2}\log 2 - \frac{1}{2} \ . \tag{15}$$

Proof. It appears to us that it was easier to compute the derivative (in the sense of distribution) of the left hand side of (15) and then the constant term. This strategy follows the idea of Mehta ([9], p. 74) who computed the principal value of this derivative. Indeed, the derivative of $\int \log|x-y| d\sigma_1(y)$ on the support of σ_1 can only be computed in the sense of distribution and is then called a principal value. There are at least two ways to compute it. The first one (used by Mehta) is to write:

$$PV\int \frac{1}{x-y} d\sigma_1(y) = \lim_{\epsilon\downarrow 0}\int_{|y-x|\geq\epsilon} \frac{1}{x-y} d\sigma_1(y) \ .$$

Metha found that, for any $x \in [-\sqrt{2}, +\sqrt{2}]$:

$$PV\int \frac{1}{x-y} d\sigma_1(y) = x \ . \tag{16}$$

For the sake of completness, let us mention the second method (which gives some information on the behavior of the derivative on the whole real line). Indeed, the second method is based on the remark that, if one knows the value of

$$D(z) = \int \frac{1}{z-y} d\sigma_1(y)$$

for complex number z in $[-\sqrt{2}, +\sqrt{2}]^c$, the principal value is then given by:

$$PV \int \frac{1}{x-y} d\sigma_1(y) = \frac{1}{2}(D(x+i0) + D(x-i0)) \ .$$

One can in fact compute D in $[-\sqrt{2}, +\sqrt{2}]^c$ using the residue method. Indeed,

$$D(z) = \frac{1}{\pi}\int_{-\sqrt{2}}^{\sqrt{2}} \frac{1}{z-y}\sqrt{2-y^2}dy \tag{17}$$

$$= \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{1}{(1+t^2)\frac{z}{\sqrt{2}} - 2t}\left(\frac{1-t^2}{1+t^2}\right)^2 dt \tag{18}$$

where we have used several changes of variables. For $z \in [-\sqrt{2}, +\sqrt{2}]^c$, we can apply the residue method to compute the right hand side of (18). We then find that:

$$D(z) = z - \sqrt{z^2 - 2}$$

where one should take care to choose the good determination of the square root. We then recover (16) since $\sqrt{z^2-2} = -\sqrt{\bar{z}^2 - 2}$.

As a consequence, (16) implies that there exists a finite constant C such that:

$$\int \log|x-y| d\sigma_1(y) = \frac{1}{2}x^2 + C. \tag{19}$$

Finally, we prove that $C = -(1/2)(1 + \log 2)$. Indeed, taking $x \in (0, -\sqrt{2}, \sqrt{2})$ and using the change of variables $y = \sqrt{2}\sin(\theta)$, we find that C must verify the three following equalities:

$$C = \log\sqrt{2} + \frac{4}{\pi}\int_0^{\frac{\pi}{2}} \log(\sin(\theta))\cos^2(\theta)d\theta \ , \tag{20}$$

$$C = \log\sqrt{2} - 1 + \frac{4}{\pi}\int_0^{\frac{\pi}{2}} \log(1-\sin(\theta))\cos^2(\theta)d\theta \ , \tag{21}$$

$$C = \log\sqrt{2} - 1 + \frac{4}{\pi}\int_0^{\frac{\pi}{2}} \log(1+\sin(\theta))\cos^2(\theta)d\theta \ . \tag{22}$$

Summing the two last equations, we get:

$$C = \log\sqrt{2} - 1 + \frac{4}{\pi}\int_0^{\frac{\pi}{2}} \log(\sin(\theta))\sin^2(\theta)d\theta \ . \tag{23}$$

Let us now sum this last equality (20) with the first one (20) so that we find:

$$2C + 1 = \log 2 + \frac{4}{\pi}\int_0^{\frac{\pi}{2}} \log(\sin(\theta))d\theta \ . \tag{24}$$

The last integral $I = \frac{4}{\pi}\int_0^{\frac{\pi}{2}} \log(\sin(\theta))d\theta$ can easily be computed. Indeed,

$$I = \frac{4}{\pi}\int_0^{\frac{\pi}{2}} \log(\sin(\theta))d\theta = \frac{2}{\pi}\int_0^{\frac{\pi}{2}} \log\left(\frac{\sin(2\theta)}{2}\right)d\theta = \frac{1}{2}I - \log 2$$

so that $I = -2\log 2$. Thus, we have proved that:

$$C = -\frac{1}{2}(1 + \log 2) \ ,$$

which finishes to prove that:

$$\int \log|x - y| d\sigma_1(y) = \frac{1}{2}x^2 - \frac{1}{2}(1 + \log 2) \ . \quad \square$$

3 Weak large deviation principle

We start with the study of the normalizing factor Z_β^N and prove that:

Property 3.1 *For any positive real number* β,

$$\lim_{N\to\infty} \frac{1}{N^2} \log Z_\beta^N = F(\beta)$$

Proof. According to Selberg (see also [9], theorem 4.1.1), for any positive integer number N and real or complex β we have, identically,

$$\begin{aligned}\psi_\beta^N &:= \int\int \prod_{1\le i<j\le N} |\lambda_i - \lambda_j|^\beta \exp\left\{-\frac{1}{2}\beta\sum_{i=1}^N \lambda_i^2\right\}\prod_{i=1}^N d\lambda_i \\ &= (2\pi)^{N/2}\beta^{-(N/2)-\beta N(N-1)/4}\left[\Gamma\left(1+\frac{\beta}{2}\right)\right]^{-N}\prod_{j=1}^N \Gamma\left(1+\frac{\beta}{2}j\right) \ . \end{aligned} \quad (25)$$

But

$$Z_\beta^N = \left(\sqrt{\frac{\beta}{N}}\right)^{\beta N(N-1)/2+N} \psi_\beta^N,$$

so that (25) gives

$$Z_\beta^N = (2\pi)^{\frac{N}{2}}(N)^{-(N/2)-\beta N(N-1)/4}\left[\Gamma\left(1+\frac{\beta}{2}\right)\right]^{-N}\prod_{j=1}^N \Gamma\left(1+\frac{\beta}{2}j\right) \ . \quad (26)$$

Moreover, for j large enough,

$$\Gamma\left(1+\frac{\beta}{2}j\right) \underset{N\to\infty}{\simeq} \left(\frac{1}{e}+\frac{\beta j}{2e}\right)^{1+\beta j/2} \sqrt{2\pi(1+\beta j/2)}$$

So that (26) gives:

$$\frac{1}{N^2}\log Z_\beta^N \underset{N\to\infty}{\simeq} -\frac{\beta}{4}\log N + \frac{\beta}{2N^2}\sum_{j=1}^N j\log\left(\frac{\beta j}{2e}\right)$$
$$\underset{N\to\infty}{\simeq} \frac{\beta}{2N^2}\sum_{j=1}^N j\log\left(\frac{\beta j}{2eN}\right) \underset{N\to\infty}{\simeq} \frac{\beta}{2}\int_0^1 x\log\frac{\beta x}{2e}dx \ . \tag{27}$$

Finally, it is not hard to compute the right hand side of (27) so that:

$$\frac{\beta}{2}\int_0^1 x\log\frac{\beta x}{2e}dx = \frac{\beta}{4}\log\frac{\beta}{2} - \frac{3}{8}\beta = F(\beta) \ ,$$

which achieves, according to (27), the proof of property 3.1. □
From this, we derive the weak large deviation principle:

Theorem 3.2 *Q_β^N satifies a weak large deviation principle with good rate function $I_\beta = H_\beta + F(\beta)$, i.e:*

for any open subset O of $\mathcal{M}_1^+(\mathbb{R})$,

$$\underline{\lim}_{N\to\infty} \frac{1}{N^2}\log Q_\beta^N\left(\hat{\mu}^N \in O\right) \geq -\inf_O H_\beta - F(\beta) \ ,$$

for any compact subset K of $\mathcal{M}_1^+(\mathbb{R})$,

$$\overline{\lim}_{N\to\infty} \frac{1}{N^2}\log Q_\beta^N\left(\hat{\mu}^N \in K\right) \leq -\inf_K H_\beta - F(\beta) \ .$$

Proof of theorem 3.2. According to Exercise 2.1.14 (v) of [6] (or Theorem 4.1.11 of [5]), it is enough to prove that for any probability measure $\nu \in \mathcal{M}_1^+(\mathbb{R})$:

$$-H_\beta(\nu) - F(\beta) = \lim_{\delta\searrow 0}\underline{\lim}_{N\to\infty}\frac{1}{N^2}\log Q_\beta^N\left(\hat{\mu}^N \in B(\nu,\delta)\right)$$
$$= \lim_{\delta\searrow 0}\overline{\lim}_{N\to\infty}\frac{1}{N^2}\log Q_\beta^N\left(\hat{\mu}^N \in B(\nu,\delta)\right) \ , \tag{28}$$

if $B(\nu,\delta)$ is the open ball centered at ν and with radius δ for the weak topology, i.e:

$$B(\nu,\delta) = \left\{\mu \in \mathcal{M}_1^+(\mathbb{R}) \mid d(\mu,\nu) < \delta\right\} \ ,$$

where we can take d to be the distance defined by:

$$d(\mu,\nu) = \inf\left|\int f d\mu - \int f d\nu\right| \ ,$$

where the infimum is taken over the Lipschitz functions f with Lipschitz constant bounded by one.

Moreover, according to property 3.1, it only remains to prove that, if $\overline{Q}_\beta^N = Z_\beta^N Q_\beta^N$ is the non normalized measure defined by:

$$d\overline{Q}_\beta^N = \prod_{1\le i<j\le N} |\lambda_i-\lambda_j|^\beta \exp\left\{-\frac{1}{2}N\sum_{i=1}^N \lambda_i^2\right\}\prod_{i=1}^N d\lambda_i \ ,$$

then:

$$\lim_{\delta\searrow 0}\overline{\lim_{N\to\infty}}\frac{1}{N^2}\log\overline{Q}_\beta^N\left(\hat{\mu}^N\in B(v,\delta)\right)\le -H_\beta(v) \tag{29}$$

and

$$\lim_{\delta\searrow 0}\underline{\lim_{N\to\infty}}\frac{1}{N^2}\log\overline{Q}_\beta^N\left(\hat{\mu}^N\in B(v,\delta)\right)\ge -H_\beta(v)\ . \tag{30}$$

We will first prove the upper bound (29) and then show that the lower bound (30) holds.

3.1 *Proof of the upper bound*

To prove (29), we notice that if $\Delta_N=\{\lambda_1<\lambda_2<\cdots<\lambda_N\}$ then:

$$\begin{aligned}
&\overline{Q}_\beta^N\left(\hat{\mu}^N\in B(v,\delta)\right)\\
&\quad=\int_{\{\hat{\mu}^N\in B(v,\delta)\}}\prod_{1\le i<j\le N}|\lambda_i-\lambda_j|^\beta\exp\left\{-\frac{1}{2}N\sum_{i=1}^N\lambda_i^2\right\}\prod_{i=1}^N d\lambda_i\\
&\quad= N!\int_{\{\hat{\mu}^N\in B(v,\delta)\}\cap\Delta_N}\exp\left\{\beta\sum_{\lambda_i<\lambda_j}\log|\lambda_i-\lambda_j|-\frac{1}{2}\sum_{\lambda_i<\lambda_j}(\lambda_i^2+\lambda_j^2)\right\}\\
&\qquad\qquad\times\exp\left\{-\frac{1}{4}\sum_{i=1}^N\lambda_i^2\right\}\prod_{i=1}^N d\lambda_i\\
&\quad= N!\int_{\{\hat{\mu}^N\in B(v,\delta)\}\cap\Delta_N}\exp\left\{-N^2\iint_{x<y}f_\beta(x,y)d\hat{\mu}^N(x)d\hat{\mu}^N(y)\right\}\\
&\qquad\qquad\times\exp\left\{-\frac{1}{4}\sum_{i=1}^N\lambda_i^2\right\}\prod_{i=1}^N d\lambda_i\ .
\end{aligned}$$

Moreover, under $\overline{Q}_\beta^N$, the λ_i's are almost surely distinct so that

$$(\hat{\mu}^N)^{\otimes 2}(\Delta_N)=\frac{1}{N}\qquad \text{a.s}$$

which allows us to conclude that, if H_β^M is defined as in (10), for any positive real number M,

$$\begin{aligned}
&\overline{Q}_\beta^N\left(\hat{\mu}^N\in B(v,\delta)\right)\\
&\quad= N!\int_{\{\hat{\mu}^N\in B(v,\delta)\}\cap\Delta_N}\exp\left\{-N^2\iint_{x<y}f_\beta(x,y)d\hat{\mu}^N(x)d\hat{\mu}^N(y)\right.
\end{aligned}$$

$$- N^2 M\left((\hat{\mu}^N)^{\otimes 2}(\Delta_N) - \frac{1}{N} \right) \Big\} \times \exp\left\{ -\frac{1}{4}\sum_{i=1}^N \lambda_i^2 \right\} \prod_{i=1}^N d\lambda_i$$

$$\leq \exp\left\{ -N^2 \inf_{\mu \in B(v,\delta)} H_\beta^{2M}(\mu) + MN \right\} N! \int_{\Delta_N} \exp\left\{ -\frac{1}{4}\sum_{i=1}^N \lambda_i^2 \right\} \prod_{i=1}^N d\lambda_i$$

$$= \exp\left\{ -\frac{N^2}{2} \inf_{\mu \in B(v,\delta)} H_\beta^{2M}(\mu) + MN \right\} (\sqrt{4\pi})^N \ .$$

So that, for any positive real number M, we get

$$\overline{\lim_{N\to\infty}} \frac{1}{N^2} \log \overline{Q}_\beta^N \left(\hat{\mu}^N \in B(v,\delta) \right) \leq - \inf_{\mu \in B(v,\delta)} H_\beta^M(\mu) \ . \tag{31}$$

But we have seen in section 2 that $H_\beta^M(\mu)$ is continuous so that it is clear that

$$\lim_{\delta \searrow 0} \overline{\lim_{N\to\infty}} \frac{1}{N^2} \log \overline{Q}_\beta^N \left(\hat{\mu}^N \in B(v,\delta) \right) \leq -H_\beta^M(v) \ . \tag{32}$$

Using the fact that $H_\beta^M(v)$ grows to $H_\beta(v)$ when M goes to infinity gives (29).

3.2 Proof of the lower bound

To prove (30), we first give a technical lemma:

Lemma 3.3 *Let v be a probability measure on $\mathbb{R}$ such that v has no atom. Let $(x^{i,N})_{1\leq i\leq N}$ be the sequence of real numbers defined by:*

$$x^{1,N} = \inf\left\{ x \mid v(]-\infty, x]) \geq \frac{1}{N+1} \right\}$$

$$x^{i+1,N} = \inf\left\{ x \geq x^{i,N} \mid v\left(]x^{i,N}, x]\right) \geq \frac{1}{N+1} \right\} \qquad 1 \leq i \leq N-1 \ .$$

Then

$$-\infty < x^{1,N} < x^{2,N} < \cdots < x^{N,N} < +\infty \ ,$$

and, for any real number η, there exists an integer number $N(\eta)$ such that, for any N larger than $N(\eta)$,

$$d\left(v, \frac{1}{N}\sum_{i=1}^N \delta_{x^{i,N}} \right) < \eta \ .$$

The proof is simple and is left to the reader.
We now turn to the proof of (30).

It is clear that we can assume without loss of generality that $H(v)$ is finite so that v has no atom (see property 2.1). Then, we can use Lemma 3.3 to see that, for $N \geq N(\frac{\delta}{2})$:

$$\left\{(\lambda_i)_{1\le i\le N} \mid |\lambda_i - x^{i,N}| < \frac{\delta}{2}\ \forall i \in [1,N]\right\} \subset \left\{(\lambda_i)_{1\le i\le N} \mid \hat{\mu}^N \in B(v,\delta)\right\} .$$

Thus,

$$\begin{aligned}
&\overline{Q}_\beta^N\left(\hat{\mu}^N \in B(v,\delta)\right) \\
&\ge \int_{\cap_{1\le i\le N}\left\{\lambda_i : |\lambda_i - x^{i,N}| < \frac{\delta}{2}\right\}} \prod_{i<j} |\lambda_i - \lambda_j|^\beta \exp\left\{-\frac{N}{2}\sum_{i=1}^N \lambda_i^2\right\} \prod_{i=1}^N d\lambda_i \\
&\ge \int_{[-\frac{\delta}{2},\frac{\delta}{2}]^N \cap \Delta_N} \prod_{i<j} \left|x^{i,N} - x^{j,N} + \lambda_i - \lambda_j\right|^\beta \exp\left\{-\frac{N}{2}\sum_{i=1}^N (x^{i,N} + \lambda_i)^2\right\} \prod_{i=1}^N d\lambda_i \quad (33)
\end{aligned}$$

But, when $\lambda_1 < \lambda_2 \cdots < \lambda_N$ and since we have constructed the $x^{i,N}$'s such that $x^{1,N} < x^{2,N} < \cdots < x^{N,N}$, we have, for any integer numbers (i,j), the lower bound

$$\left|x^{i,N} - x^{j,N} + \lambda_i - \lambda_j\right| > \max\left\{\left|x^{j,N} - x^{i,N}\right|, |\lambda_j - \lambda_i|\right\} .$$

As a consequence, (33) gives:

$$\begin{aligned}
\overline{Q}_\beta^N\left(\hat{\mu}^N \in B(v,\delta)\right) \ge& \prod_{i+1<j} |x^{i,N} - x^{j,N}|^\beta \prod_{i=1}^{N-1} |x^{i+1,N} - x^{i,N}|^{\frac{\beta}{2}} \\
&\times \exp\left\{-\frac{N}{2}\sum_{i=1}^N (|x^{i,N}| + \delta)^2\right\} \\
&\times \int_{[-\frac{\delta}{2},\frac{\delta}{2}]^N \cap \Delta_N} \prod_{i=1}^{N-1} |\lambda_{i+1} - \lambda_i|^{\frac{\beta}{2}} \prod_{i=1}^N d\lambda_i \qquad (34)
\end{aligned}$$

Moreover, one can easily bound from below the last term in the right hand side of (34). Indeed, if one uses the change of variables

$$u_1 = \lambda_1, \qquad u_{i+1} = \lambda_{i+1} - \lambda_i \quad 1 \le i \le N-1 ,$$

Then,

$$\int_{[-\frac{\delta}{2},\frac{\delta}{2}]^N \cap \Delta_N} \prod_{i=1}^{N-1} (\lambda_{i+1} - \lambda_i)^{\frac{\beta}{2}} \prod_{i=1}^N d\lambda_i = \int_{A_N} \prod_{i=2}^N (u_i)^{\frac{\beta}{2}} \prod_{i=1}^N du_i,$$

where $A_N = \{(u_i)_{1\le i\le N} / u_1 \in \left[-\frac{\delta}{2}, +\frac{\delta}{2}\right],\ u_i \in \left[0, +\frac{\delta}{2}\right],\ 2 \le i \le N,\ \left|\sum_{i=1}^N u_i\right| \le \frac{\delta}{2}\}$. But $\cap_{1\le i\le N}\{u_i \in [0, +\frac{\delta}{2N}]\} \subset A_N$, so that one can get the lower bound.

$$\int_{[-\frac{\delta}{2},\frac{\delta}{2}]^N \cap \Delta_N} \prod_{i=1}^{N-1} (\lambda_{i+1} - \lambda_i)^{\frac{\beta}{2}} \prod_{i=1}^N d\lambda_i \ge \left(\frac{1}{\beta/2+1}\right)^{(N-1)} \left(\frac{\delta}{2N}\right)^{(\beta/2+1)(N-1)+1} . \quad (35)$$

Hence, (34) implies:

$$\begin{aligned}\overline{Q}^N_\beta\left(\hat{\mu}^N \in B(\nu,\delta)\right) \geq & \prod_{i+1<j} |x^{i,N}-x^{j,N}|^\beta \\ & \times \prod_{i=1}^{N-1} |x^{i+1,N}{}^{i,N}|^{\frac{\beta}{2}} \exp\left\{-\frac{N}{2}\sum_{i=1}^N (x^{i,N})^2\right\} \\ & \times \left(\frac{1}{\beta/2+1}\right)^{(N-1)} \left(\frac{\delta}{2N}\right)^{(\beta/2+1)(N-1)+1} \\ & \times \exp\left\{-N\delta\sum_{i=1}^N |x^{i,N}| - N^2\delta^2\right\} .\end{aligned} \tag{36}$$

Moreover

$$\begin{aligned}\frac{1}{2(N+1)}\sum_{i=1}^N \left(x^{i,N}\right)^2 - & \frac{\beta}{(N+1)^2}\sum_{i+1<j}\log\left|x^{i,N}-x^{j,N}\right| \\ & -\frac{\beta}{2(N+1)^2}\sum_{i=1}^{N-1}\log\left|x^{i+1,N}-x^{i,N}\right| \\ & \leq \frac{1}{2}\int x^2 d\nu(x) - \beta\int_{x^{1,N}\leq x<y\leq x^{N,N}} \log(y-x)d\nu(x)d\nu(y)\end{aligned} \tag{37}$$

Indeed, it is easy to see that, with our choice of the $x^{i,N}$'s,

$$\frac{1}{N+1}\sum_{i=1}^N \left(x^{i,N}\right)^2 \leq \int x^2 d\nu(x) \ . \tag{38}$$

Similarly, since $x \to \log(x)$ increases on $\mathbb{R}^+$, we notice that

$$\begin{aligned}& \int_{x^{1,N}\leq x<y\leq x^{N,N}} \log(y-x)d\nu(x)d\nu(y) \\ & \leq \sum_{1\leq i\leq j\leq N-1} \log\left(x^{j+1,N}-x^{i,N}\right)\nu^{\otimes 2}\left(x^{i,N}\leq x\leq x^{i+1,N}; x<y; x^{j,N}\leq y\leq x^{j+1,N}\right) \\ & = \frac{1}{(N+1)^2}\sum_{i<j}\log\left|x^{i,N}-x^{j+1,N}\right| + \frac{1}{2(N+1)^2}\sum_{i=1}^{N-1}\log\left|x^{i+1,N}-x^{i,N}\right| \ .\end{aligned} \tag{39}$$

(38) and (39) give (37).

On the other hand, we can see as in (38) that $1/(N+1)\sum_{i=1}^N |x^{i,N}|$ is bounded by $\int |x| d\nu(x)$ so that (36) finally gives:

$$\begin{aligned}& \overline{Q}^N_\beta\left(\hat{\mu}^N \in B(\nu,\delta)\right) \\ & \geq \exp\left\{(N+1)^2\left\{\beta\int_{x^{1,N}\leq x<y\leq x^{N,N}} \log(y-x)d\nu(x)d\nu(y) - \frac{1}{2}\int x^2 d\nu(x)\right\}\right\} \\ & \quad \times \exp\left\{-(N^2+N)\left\{\delta\int |x|d\nu(x) + \frac{1}{N}\int x^2 d\nu(x) + \delta^2\right\}\right\} \\ & \quad \times \left(\frac{2}{\beta+2}\right)^{N-1} \times \left(\frac{\delta}{2N}\right)^{(\beta/2+1)(N-1)+1} .\end{aligned} \tag{40}$$

But it is not hard to see that since $H_\beta(v)$ is finite, $\int x^2 dv(x)$ is also finite so that we can conclude from (40) that

$$\begin{aligned}\varliminf_{N\to\infty} \frac{1}{N^2} \log \overline{Q}_\beta^N \left(\hat{\mu}^N \in B(v,\delta)\right) \\ \geq -\delta \int |x| dv(x) - \delta^2 + \varliminf_{N\to\infty} \left\{\beta \int_{x^{1,N} \leq x < y \leq x^{N,N}} \log(y-x) dv(x) dv(y) \right. \\ \left. - \frac{1}{2} \int x^2 dv(x) \right\} .\end{aligned} \tag{41}$$

It is clear that the last term in the right hand side of (41) converges so that we get that for any positive real number δ,

$$\varliminf_{N\to\infty} \frac{1}{N^2} \log \overline{Q}_\beta^N \left(\hat{\mu}^N \in B(v,\delta)\right) \geq -\delta \int |x| dv(x) - \delta^2 - H_\beta(v) \tag{42}$$

so that, since $\int |x| dv(x) \leq \sqrt{\int x^2 dv(x)}$ is finite, the proof of (30) is complete once we let δ tends to zero. $\square$

Remark 3.4. The arguments used for the proof of Theorem 3.2 seem rather unusual among large deviations techniques. In fact, the very crude estimates that we make here are only possible because of the scale N^2 of the weak large deviation principle 3.2 which allows us to neglect even reordering in the eigenvalues (of weight $N!$). Nevertheless, we never use any continuity property of H_β or Σ. Indeed, even if H_β is lower semi continuous, it has no reason to be upper semi continuous since f_β is not bounded. In fact, if one considers the sequence of probability measures $(\mu_n)_{n\in\mathbb{N}}$ defined by:

$$\mu_n(dx) = \left(n\alpha_n \mathbb{1}_{[0,\frac{1}{n}]}(x) + (1-\alpha_n)\mathbb{1}_{[0,1]}(x)\right) dx \ ,$$

where α_n is a sequence of positive real numbers which goes to zero when n goes to infinity, then one can check that (μ_n) converges to $\mathbb{1}_{[0,1]}(x)dx$ when n goes to infinity. On the other hand, if α_n goes to zero more slowly than $(1/\log n)^{\frac{1}{2}}$, $\Sigma(\mu_n)$ goes to $-\infty$ when n goes to infinity whereas $\Sigma(\mathbb{1}_{[0,1]}(x)dx)$ is obviously finite.

Remark 3.5. It is the main reason why we do not understand the proof given by Chan [4]. Indeed, the author assumes there (see p. 187 and Lemma 3.1) that Σ is continuous at any probability measure v absolutely continuous with respect to Lebesgue measure and with bounded density, which does not hold under the weak topology according to the above example.

4 Exponential tightness

We finally prove that Q_β^N is exponentially tight.

Property 4.1 *For any positive real number L, there exists a compact subset K_L of $\mathscr{M}_1^+(\mathbb{R})$ such that*

$$\overline{\lim_{N\to\infty}} \frac{1}{N^2} \log Q_\beta^N\left(\hat{\mu}^N \in K_L^c\right) \le -L \ .$$

Proof. Using Cauchy Schwarz's inequality, one finds that for any measurable subset K of $\mathscr{M}_1^+(\mathbb{R})$,

$$Q_\beta^N\left(\hat{\mu}^N \in K^c\right) \le \left(\frac{Z_{2\beta}^N}{\left(Z_\beta^N\right)^2}\right)^{\frac{1}{2}} \left(\int_{\hat{\mu}^N \in K^c} \exp\left\{-\frac{1}{2}N\sum_{i=1}^N \lambda_i^2\right\} \prod_{i=1}^N d\lambda_i\right)^{\frac{1}{2}} \ .$$

Hence, property 3.1 implies that

$$\begin{aligned}&\overline{\lim_{N\to\infty}} \frac{1}{N^2} \log Q_\beta^N\left(\hat{\mu}^N \in K^c\right) \\ &\quad \le \frac{1}{2}(F(2\beta) - 2F(\beta)) + \overline{\lim_{N\to\infty}} \frac{1}{2N^2} \log \int_{\hat{\mu}^N \in K^c} \exp\left\{-\frac{1}{2}N\sum_{i=1}^N \lambda_i^2\right\} \prod_{i=1}^N d\lambda_i \ .\end{aligned} \tag{43}$$

Moreover, for any positive real number η, if we choose $K = \cap_{i\in\mathbb{N}} \left\{\mu \in \mathscr{M}_1^+(\mathbb{R}) \Big/ \mu\left(\left[-\left(i+\frac{4}{\eta}\right), +\left(i+\frac{4}{\eta}\right)\right]^c\right) < \left(i+\frac{4}{\eta}\right)^{-1}\right\}$, then K is a compact subset of $\mathscr{M}_1^+(\mathbb{R})$ and:

$$\begin{aligned}&\int_{\hat{\mu}^N \in K^c} \exp\left\{-\frac{1}{2}N\sum_{i=1}^N \lambda_i^2\right\} \prod_{i=1}^N d\lambda_i \\ &\quad \le \sum_{i\in\mathbb{N}} \int_{\left\{\hat{\mu}^N([-n^i,+n^i]^c) < \left(i+\frac{4}{\eta}\right)^{-1}\right\}^c} \exp\left\{-\frac{1}{2}N\sum_{i=1}^N \lambda_i^2\right\} \prod_{i=1}^N d\lambda_i \ .\end{aligned}$$

But one can check that

$$\begin{aligned}&\int_{\left\{\hat{\mu}^N\left(\left[-\left(i+\frac{4}{\eta}\right),+\left(i+\frac{4}{\eta}\right)\right]^c\right) \ge \left(i+\frac{4}{\eta}\right)^{-1}\right\}^c} \exp\left\{-\frac{1}{2}N\sum_{j=1}^N \lambda_i^2\right\} \prod_{j=1}^N d\lambda_j \\ &\quad \le \left(2\sqrt{4\pi/N}\right)^N \exp\left\{-\frac{1}{4}\left(i+\frac{4}{\eta}\right)N^2\right\}\end{aligned}$$

so that we get

$$\int_{K^c} \exp\left\{-\frac{1}{2}N\sum_{i=1}^N \lambda_i^2\right\} \prod_{i=1}^N d\lambda_i \le \frac{\left(2\sqrt{4\pi/N}\right)^N}{1-\exp\left\{-\frac{N^2}{4}\right\}} \exp\left\{-\frac{1}{\eta}N^2\right\} \ .$$

Choosing $\eta^{-1} = 2L + (F(2\beta) - 2F(\beta))$ ends, with (43), the proof of property 4.1. $\square$

5 Large deviation principles for more general large random matrices

In this section, we would like to point out how the large deviation principle stated in Theorem 1.3 for the Gaussian ensembles can be extended to other random matrices.

As a first generalization, we mention the case of Hermitian (resp. symmetric) random matrices $(M_N)_{N\geq 0}$ which appear in quantum field theory and quantum chaology (see [3], [11] and [10]) and which law can be described by

$$p_N(M_N)dM_N = \frac{1}{Z^N_{\beta,V}}\exp\{-N^2\tau(V(M_N))\}dM_N \ , \tag{44}$$

where τ is the normalized trace $\tau(A) = (1/N)\sum_{i=1}^N A_{ii}$, V is a real function, $Z^N_{\beta,V}$ is the normalizing constant and dM_N is the Lebesgue measure on the space of $N \times N$ complex (resp. real) matrices:

$$dM_N = \prod_{i\leq j} d\mathrm{Im}\ M_{ij}\ d\mathrm{Re}\ M_{ij} \qquad \left(\text{resp. } dM_N = \prod_{i\leq j} dM_{ij}\right) \ .$$

Then, if M_N is assumed to be invariant either under the action of the unitary group or of the symplectic group in the complex setting, or under the action of the orthogonal group in the real setting, the joint law of the eigenvalues $(\lambda_i)_{1\leq i\leq N}$ of the matrix M_N can be seen to be of the form:

$$Q^N_{\beta,V}(d\lambda) = \frac{1}{Z^N_{\beta,V}}\prod_{i<j}|\lambda_i - \lambda_j|^\beta \exp\left\{-\frac{N}{2}\sum_{i=1}^N V(\lambda_i)\right\}\prod_{i=1}^N d\lambda_i,$$

with $\beta = 1$ (resp. $\beta = 2$, resp. $\beta = 4$) for the orthogonal (resp. unitary, resp. symplectic) case. Of course, the Gaussian ensembles we studied above can be recovered by choosing $V(\lambda) = (1/2)\lambda^2$. We now wish to give some extension to other potentials. The first obvious generalization can be deduced from Theorem 1.3 thanks to Laplace-Varadhan's Lemma (see [5], Theorem 4.3.1):

Corollary 5.1 *If there exists a finite positive real number a such that $f(x) := V(x) - ax^2$ is a bounded continuous function on $\mathbb{R}$, the law of the empirical measure $\frac{1}{N}\sum_{i=1}^N \delta_{\lambda_i}$ under $Q^N_{\beta,V}$ satisfies a large deviation principle with good rate function*

$$I^V_\beta(\mu) = aI_{\frac{\beta}{a}}(\mu) + \frac{1}{2}\int f d\mu - \inf_{\nu\in\mathscr{M}_1^+(\mathbb{R})}\left(aI_{\frac{\beta}{a}}(\nu) + \frac{1}{2}\int f d\nu\right) \ .$$

More generally, Laplace-Varadhan's Lemma shows that, for any bounded continuous function ϕ: $\mathscr{M}_1^+(\mathbb{R}) \to \mathbb{R}$, the law of the spectral measure under

$$Q^N_{\beta,\phi}(d\lambda) = \frac{1}{Z^N_{\beta,\phi}}\prod_{i<j}|\lambda_i - \lambda_j|^\beta \exp\left\{-N^2\phi\left(\frac{1}{N}\sum_{i=1}^N \delta_{\lambda_i}\right) - \frac{N}{2}\sum_{i=1}^N \lambda_i^2\right\}\prod_{i=1}^N d\lambda_i$$

satisfies a large deviation principle with good rate function

$$I_\beta^\phi(\mu) = I_\beta(\mu) + \phi - \inf_{v \in \mathcal{M}_1^+(\mathbb{R})} (I_\beta(v) + \phi) \ .$$

On the other hand, one could be interested in matrices satisfying (44) but with a potential V which is not a bounded modification of x^2 (for instance $V(x) = x^4$). Then, it is worth noticing that the method used to prove Theorem 1.3 can be applied in many of these cases. For instance, one can easily see that:

Theorem 5.2 *If V is a continuously differentiable function on $\mathbb{R}$ going to infinity sufficiently fast with $|x|$ (faster than logarithmicaly) but not too fast so that:*

$$\overline{\lim_{\delta \to 0}} \ \overline{\lim_{|x| \to \infty}} \sup_{|y| \le \delta} \left| \frac{V'(x+y)}{V(x)} \right| < \infty,$$

then:

1) Let $I_\beta^V(\mu) = \frac{1}{2}\left(\int V d\mu - \beta\Sigma(\mu)\right) - \frac{1}{2} \inf_{v \in \mathcal{M}_1^+(\mathbb{R})} \left(\int V dv - \beta\Sigma(v)\right).$

a. I_β^V *is a good rate function.*

b. I_β^V *achieves its minimal value at a unique probability measure* σ_β^V *which is characterized by:*

$$\frac{1}{2}V(x) - \beta \int \log|y - x| d\sigma_\beta^V(y) = \frac{1}{2} \inf_{v \in \mathcal{M}_1^+(\mathbb{R})} \left(\int V dv - \beta\Sigma(v)\right) \qquad \sigma_\beta^V a.s \ .$$

and, for all x except possibly on a set with null logarithmic capacity,

$$\frac{1}{2}V(x) - \beta \int \log|y - x| d\sigma_\beta^V(y) \ge \frac{1}{2} \inf_{v \in \mathcal{M}_1^+(\mathbb{R})} \left(\int V dv - \beta\Sigma(v)\right) \ .$$

2) The law of the spectral measure $\frac{1}{N}\sum_{i=1}^N \delta_{\lambda_i}$ *under* $Q_{\beta,V}^N$ *satisfies a large deviation principle in the scale* N^2 *with good rate function* I_β^V*.*

Let us remark that the assumptions on V have two different origins: the first one is needed at least to insure the existence of the probability measure $Q_{\beta,V}^N$ and the second one is needed on a technical level to prove the lower bound of the large deviation principle. It implies that we can not consider potential V which are going to infinity faster than exponentially at infinity.

In a slightly different context, one could wonder what our approach could tell about the eigenvalues of the circular ensembles which are known to have complex eigenvalues $(e^{i\theta_j})_{1 \le j \le N}$ with real numbers θ_j in $[0, 2\pi]$ satisfying:

$$Q_{\beta,c}^N(d\theta) = \frac{1}{Z_{\beta,c}^N} \prod_{k<j} |e^{i\theta_j} - e^{i\theta_k}|^\beta \prod_{k=1}^N d\theta_k \ .$$

In that case, it is clear that:

Property 5.3 *For any probability measure μ on* $[0, 2\pi]$, *let*

$$I_\beta^c = -\beta \int_0^{2\pi} \int_0^{2\pi} \log|e^{i\alpha} - e^{i\alpha'}| d\mu(\alpha) d\mu(\alpha') \ .$$

Then

1) I^c_β *is a good rate function which achieves its minimal value at a unique probability measure which is the uniform law on* $[0, 2\pi]$.
2) *The law of the empirical measure of the arguments* $(\theta_k)_{1\le k\le N}$ *under* $Q^N_{\beta,c}$ *satisfies a large deviation upper bound with good rate function* I^c_β, *i.e:*

$$\text{for any closed set } F \qquad \overline{\lim_{N\to\infty}} \frac{1}{N^2} \log Q^N_{\beta,c}\left(\frac{1}{N}\sum_{i=1}^N \delta_{\lambda_i} \in F\right) \le -\inf_F I^c_\beta \ .$$

Of course, this result is not as satisfactory as a full large deviation principle since it does not give the lower bound (which proof would need some new argument). Nevertheless, Property 5.3 is already enough to prove the convergence of the empirical measure $(1/N)\sum \delta_{\theta_i}$ to the uniform law on $[0, 2\pi]$ and thus the convergence of the spectral measure. Moreover, the proof is straightforward. Indeed, the first point is obvious since I^c_β can be seen to be lower semicontinous because:

$$g_\beta(\alpha, \alpha') = \begin{cases} -\log|e^{i\alpha} - e^{i\alpha'}| & \text{if } \alpha \neq \alpha' \\ +\infty & \text{otherwise} \end{cases}$$

is a positive continuous function. But since the space of probability measures on $[0, 2\pi]$ is compact, it is then a good rate function. Moreover, the uniqueness of the minimum is here a direct consequence of logarithmic capacity problem. Finally, the proof of the upper bound is exactly similar to that given for the Gaussian setting, since the normalizing constant $Z^N_{\beta,c}$ has already been computed (see [9], Theorem 11.1.1).

The second generalization of Theorem 1.3 that we would like to point out is given by the contraction principle (see Thm 4.2.1 of [5]). For simplicity, let us focus on the real case where $\beta = 1$ and consider the random matrices X^N defined in (1). Then, the contraction principle allows us to deduce from Theorem 1.1 that the spectral measure of any continuous image of X^N satisfies a large deviation principle. More precisely:

Theorem 5.4 *If* p *is a continuous map, then the law of the spectral measure of* $Y^N = p(X^N)$ *satisfies a large deviation principle with good rate function* $\mathcal{I}^p$ *given by*

$$\mathcal{I}^p(\mu) = \inf\{I_1(\nu)/\nu(p(x) \in .) = \mu(x \in .)\} \ .$$

Remark 5.5 Of course, if p is a continuous bijective map, Theorem 5.4 applies. Then, $\mathcal{I}^p(\mu) = I_1(\mu \circ p)$. As a consequence, Theorem 2.3 shows that $\mathcal{I}^p$ achieves its minimal value at a unique probability measure which is $\sigma_1 \circ p^{-1}$ so that the spectral measure of Y^N converges to $\sigma_1 \circ p^{-1}$.

For all the large deviations principles we have been stating, $-\Sigma$ plays the role of an entropy. We would like, as a final remark, recall how Voiculescu interpreted $-\Sigma$ as an entropy in [14]. In this paper, he indeed considered the relative entropy $I\left(Q^N_{(p)}|L\right)$ of the law $Q^N_{(p)}$ of the eigenvalues of $p(X^N)$ with respect to the Lebesgue measure:

$$L^N(d\lambda) = \prod_{1\le i<j\le N} |\lambda_i - \lambda_j| \prod_i d\lambda_i \ .$$

Voiculescu then stated that:

$$\lim_{N\to\infty} \frac{1}{N^2} I\left(Q^N_{(p)}|L^N\right) = \Sigma\left(\sigma_1 \circ p^{-1}\right) + \text{const.}$$

so that Σ appears as a normalized limit of entropies of laws of random matrices and thus as a good candidate for the definition of a non commutative entropy. Of course, the above heuristics are not really rigorous since the "Lebesgue" measures L^N are not probability measures. With this remark in mind, we finally want to give a rigorous statement of Voiculescu's heuristics and generalize them to other random matrices as follows:

Voiculescu's heuristics 5.6. *For any positive β, if $p\colon \mathbb{R} \to \mathbb{R}$ is any diffeomorphism such that there exists positive real numbers ϵ and M so that*

$$\epsilon < \left|\left|(p^{-1})'\right|\right|_\infty < M \tag{45}$$

and if V is any continuous function going to infinity fast enough and so that there exists a positive real number α, $\alpha > 1$, such that

$$\sup_N E\left[\left|\frac{1}{N}\sum_{i=1}^N V\left(p(X^N)_{ii}\right)\right|^\alpha\right] < +\infty \ , \tag{46}$$

the relative entropy of the law of the eigenvalues of $Y^N = p(X^N_{\beta,V})$ with respect to the law $Q^N_{\beta,V}$ of the eigenvalues of $X^N_{\beta,V}$ converges, once divided by N^2, to $I^V_\beta\left(\sigma^V_\beta\right) - I^V_\beta\left(\sigma^V_\beta \circ p^{-1}\right)$.

Hence, leaving asides the constants, $\beta\Sigma(\mu) - \int V(x)d\mu(x)$ appears as a limit of normalized relative entropies of the spectral law of a set of large random matrices. In this sum of two terms, Σ is the most relevant term since it is closely related to the matrix origin of our problem.

Moreover, we must emphasize that the hypotheses (45) has only be been assumed to make the proof shorter. Furthermore, we here need assumption (46) because we do not consider the relative entropy with respect to the Lebesgue measure but with respect to $Q^N_{\beta,V}$.

The proof of 5.6 follows that given by Voiculescu in [14].

Indeed, if one considers the relative entropy of the law $Q^N_{(p)}$ of the eigenvalues of $Y^N = p(X^N)$ with respect to the law Q^N_1 of the eigenvalues of X^N, then, denoting $q = p^{-1}$, one gets:

$$
\begin{aligned}
I\left(Q^N_{(p)}|Q^N_1\right) &= \int \log\left(\frac{dQ^N_{(p)}}{dQ^N_1}\right) dQ^N_{(p)} \\
&= \beta \int \left(\sum_{i<j} \log\left|\frac{\lambda_i - \lambda_j}{q(\lambda_i) - q(\lambda_j)}\right| + \frac{N}{2}\sum_i \{V(q(\lambda_i)) - V(\lambda_i)\} \right. \\
&\qquad \left. - \sum_i \log q'(\lambda_i) \right) dQ^N_{(p)}.
\end{aligned}
$$

Hence

$$
\begin{aligned}
\frac{1}{N^2} I\left(Q^N_{(p)}|Q^N_1\right) = \int \Bigg(& -\frac{\beta}{2} \int \int_{x \neq y} \log\left|\frac{q(x) - q(y)}{x - y}\right| d\hat{\mu}^N(x) d\hat{\mu}^N(y) \\
& + \frac{1}{2} \int \left\{ V(q(x))^2 - V(x)^2 \right\} d\hat{\mu}^N(x) \\
& - \frac{1}{N} \int \log q'(x) d\hat{\mu}^N(x) \Bigg) dQ^N_{(p)} \ . \qquad (47)
\end{aligned}
$$

But, under assumption (45),

$$
\mu \to \iint_{x \neq y} \log\left|\frac{q(x) - q(y)}{x - y}\right| d\mu(x) d\mu(y)
$$

is a bounded continuous function so that, according to Remark 5.5,

$$
\begin{aligned}
& \lim_{N \to \infty} \int \left(\iint_{x \neq y} \log\left|\frac{q(x) - q(y)}{x - y}\right| d\hat{\mu}^N(x) d\hat{\mu}^N(y) \right) dQ^N_{(p)} \\
&= \iint_{x \neq y} \log\left|\frac{q(x) - q(y)}{x - y}\right| d\sigma_1 \circ q(x) d\sigma_1 \circ q(y) \\
&= \Sigma(\sigma_1) - \Sigma(\sigma_1 \circ q) \ .
\end{aligned}
$$

Moreover, assumption (45) implies that $\log q'$ is uniformly bounded so that:

$$
\lim_{N \to \infty} \frac{1}{N} \int \left[\int \log q'(x) d\hat{\mu}^N(x) \right] dQ^N_{(p)} = 0 \ .
$$

And finally, one can see that assumption (46) and Theorem 1.1 imply that:

$$
\begin{aligned}
& \lim_{N \to \infty} \int \left[\int \{ V(q(x)) - V(x) \} d\hat{\mu}^N(x) \right] dQ^N_{(p)} \\
&= \int V(x) d\sigma_1(x) \ - \int V(x) d\sigma_1 \circ q(x) \ .
\end{aligned}
$$

Thus (47) gives:

$$
\begin{aligned}
\lim_{N \to \infty} \frac{2}{N^2} I\left(Q^N_{(p)}|Q^N_1\right) &= \beta\Sigma(\sigma_1 \circ q) - \beta\Sigma(\sigma_1) \\
&\quad + \left(\int V(x) d\sigma_1(x) - \int V(x) d\sigma_1 \circ q(x) \right)
\end{aligned}
$$

i.e Voiculescu's heuristics 5.6. □

References

[1] Arnold, L.: On the asymptotic distribution of the eigenvalues of random matrices. J. Math. Anal. Appl. **20**, 262–268 (1967)

[2] Bai, Z. D.: Convergence rate of expected spectral distributions of large random matrices. Part I. Wigner Matrices. Annal. Probab. **21**, No. 2, 625–648 (1993)

[3] Bohigas, O.: Chaos and Quantum Physics. Giannoni, M.J. Voros, A. Zinn-Justin J. (eds), North Holland; Amsterdam (1991)

[4] Chan, T.: Large deviations for empirical measure with degenerate limiting distribution. Probability Theory and Related Fields, **97**, 179–193 (1993)

[5] Dembo, A., Zeitouni, O.: Large deviations techniques. Jones and Bartlett

[6] Deuschel, J-D., Stroock, D.W.: Larges deviations. Academic Press

[7] Hormander, L.: The analysis of linear partial differential operators, 1. Springer (1983)

[8] Mhaskar, H. N., Saff, E. B.: Where does the sup norm of a weighted Polynomial live ? (A generalization of Incomplete Polynomials) Constr. Approx. 1: 71–91 (1985)

[9] Mehta, M. A.: Random matrices, revised and enlarged (second edition). Academic Press

[10] Pastur, L.A.: Spectral and Probabilistic aspects of matrix models. Preprint (1995)

[11] Pichard, J.L.: Quantum coherence in Mesoscopic systems. B. Kramer (Ed) Plenum, NY (1991)

[12] Tsuji, M.: Potential theory in modern function theory. Marisen (Tokyo) (1959)

[13] Voiculescu, D.: Limit laws for random matrices and free products. Invent. math. **104**, 201–220 (1991)

[14] Voiculescu, D.: The analogues of entropy and of Fisher's information measure in free probability theory, I. Communications in mathematical physics, **155**, 70–92 (1993)

[15] Wigner, E. P.: On the distributions of the roots of certain symmetric matrices. Ann. of Math., **67**, 325–327 (1958)