Large Deviations for the Largest Eigenvalue of Sub-Gaussian Matrices

Fanny Augeri, Alice Guionnet, Jonathan Husson

LARGE DEVIATIONS FOR THE LARGEST EIGENVALUE OF SUB-GAUSSIAN MATRICES

FANNY AUGERI, ALICE GUIONNET, AND JONATHAN HUSSON

Abstract: We establish large deviations estimates for the largest eigenvalue of Wigner matrices with sub-Gaussian entries.

1. Introduction

Understanding the large deviation behavior of the largest eigenvalue of a random matrix is a challenging question, with many applications in statistics and mobile communications systems, e.g [10, 7]. However, it is in general a difficult question and very few results are known. It is known since [19] that the empirical distribution of the eigenvalues of a Wigner matrix converges to the semi-circle law provided the off-diagonal entries have a finite second moment. Following the pioneering work of Kómlos and Fűredi [11], we know by [4] that assuming the Wigner matrix has centered entries, the largest eigenvalue converges to the right edge of the support of the semi-circle law if and only if the fourth moment of the off-diagonal entries is finite. Given these two results, one can wonder what is the probability that the empirical measure or the largest eigenvalue have an unexpected behavior. Large deviation principles were derived for the empirical distribution of the eigenvalues and the largest eigenvalue of classical Gaussian ensembles, as the Gaussian Unitary Ensemble (GUE) and Gaussian Orthogonal Ensemble (GOE) in [6] and [5]. Indeed, in this case, the joint law of the eigenvalues is explicit and large deviations estimates can be derived by Laplace's method, up to taking care of the singularity of the interaction. In a breakthrough paper, Bordenave and Caputo [8] showed that large deviations for the empirical measure of the eigenvalues can be estimated when the tails of the entries are heavier than in the Gaussian case. These large deviations have a smaller speed than in the cases of classical Gaussian Ensembles and are due to a relatively small number of entries of order one. This phenomenon was shown to hold as well for the largest eigenvalue by one of the authors [2].

Yet, the case of sub-Gaussian entries remained still mysterious. Last year, two of the authors showed that if the Laplace transform of the Wigner matrix is pointwise bounded from above by the one of the GUE or GOE, then a large deviations principle holds with the same rate function as in the Gaussian case. This special case of Wigner matrices, which was called *with sharp sub-Gaussian tails*, was shown to include matrices with Rademacher variables and variables uniformly sampled in an interval. Yet, many Wigner matrices with sub-Gaussian entries are not with sharp sub-Gaussian tails, as for Gaussian sparse matrices which are obtained by multiplying entrywise a GOE matrix with an independent

This project was partially supported by Labex MILYON and has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 692452).

Bernoulli random variable. In this article, we investigate this general setting and derive large deviations estimates for the largest eigenvalue of Wigner matrices with sub-Gaussian entries. In particular, we show that the rate function of this large deviations principle is different from the one of the GOE.

We will consider hereafter a $N \times N$ symmetric random matrix X_N with independent entries $(X_{ij})_{i \leq j}$ above the diagonal so that $\sqrt{N}X_{ij}$ has law μ for all $i < j$ and $\sqrt{N/2}X_{ii}$ has law μ for all i. In particular, the variance profile is the same as the one of the GOE. We assume that μ is centered and has a variance equal to 1. Let

$$\forall x \in \mathbb{R},\ \psi(x) = \frac{1}{x^2} \log \int e^{xt} d\mu(t)\,.$$

$\psi(0) = 1/2$ and is a continuous function on $\mathbb{R}$. Assume that

$$\frac{A}{2} := \sup_{x \in \mathbb{R}} \psi(x) < +\infty. \tag{1}$$

The case where $A = 1$ is the case of *sharp sub-Gaussian tails* which was studied in [12]. We investigate here the case where $A > 1$ and we show the following result.

Theorem 1.1. *Denote by λ_{X_N} the largest eigenvalue of X_N. Under some technical assumptions, there exist a good rate function $I_\mu : \mathbb{R} \to [0, +\infty]$ and a set $\mathcal{O}_\mu \subset \mathbb{R}$ such that $(-\infty, 2] \cup [x_\mu, +\infty) \subset \mathcal{O}_\mu$ for some $x_\mu \geq 2$ and such that for any $x \in \mathcal{O}_\mu$,*

$$\lim_{\delta \to 0} \liminf_{N \to +\infty} \frac{1}{N} \log \mathbb{P}\left(|\lambda_{X_N} - x| \leq \delta\right) = \lim_{\delta \to 0} \limsup_{N \to +\infty} \frac{1}{N} \log \mathbb{P}\left(|\lambda_{X_N} - x| \leq \delta\right) = -I_\mu(x).$$

The rate function I_μ is infinite on $(-\infty, 2)$ and satisfies,

$$I_\mu(x) \underset{x \to +\infty}{\sim} \frac{1}{4A} x^2.$$

If $A \in (1, 2)$, then $[2, \sqrt{A-1} + 1/\sqrt{A-1}] \subset \mathcal{O}_\mu$ and I_μ coincides on this interval to the rate function of the GOE, that is,

$$I_\mu(x) = \frac{1}{2} \int_2^x \sqrt{y^2 - 4} dy =: I_{GOE}(x). \tag{2}$$

Moreover, for all $x \geq 2$, $I_\mu(x) \leq I_{GOE}(x)$.

The technical assumptions include the case where ψ is increasing (which holds in the case of sparse Gaussian entries) and the case where the maximum of ψ is achieved on $\mathbb{R}$ at a unique point in a neighborhood of which it is strictly concave. In the later case, $I_\mu(x)$ only depends on A for x large enough.

1.1. **Assumptions.**

We now describe more precisely our assumptions.

Assumption 1.1. *Let $\mu \in \mathcal{P}(\mathbb{R})$ be a symmetric probability measure with unit variance. We denote by L its log-Laplace transform,*

$$\forall x \in \mathbb{R},\ L(x) = \log \int e^{xt} d\mu(t),$$

and $\psi(x) = L(x)/x^2$. *We assume that* μ *is sub-Gaussian in the sense that*

$$\frac{A}{2} := \sup_{x\in\mathbb{R}} \psi(x) < +\infty,$$

and we define $B \geq 0$ *by,*

$$\frac{B}{2} := \lim_{|x|\to+\infty} \psi(x).$$

We assume moreover that $L(\sqrt{.})$ *is a Lipschitz function and that* μ *does not have sharp sub-Gaussian tails, meaning that* $A > 1$.

We describe below a few examples of probability measures μ which satisfy the above assumptions. In each of these cases, the fact that $L(\sqrt{.})$ is Lipschitz is clear and left to the reader.

Example 1.2.

- *(Combination of Gaussian and Rademacher laws). Let*

$$\mu(dx) = a\frac{e^{-\frac{1}{2B}x^2}}{\sqrt{2\pi B}}dx + (1-a)\frac{1}{2}(\delta_{-b} + \delta_{+b})$$

where a, b, B *are non negative real numbers such that* $a \in (0,1)$ *and* $aB+(1-a)b^2 = 1$. *Then, for all* $x \in \mathbb{R}$,

$$L_\mu(x) = \log\left(ae^{\frac{B}{2}x^2} + (1-a)\cosh(bx)\right).$$

If $B > 1$ *and* $b \in (0,1)$ *we see that our conditions are fulfilled and* $A = B$.

- *(Sparse Gaussian case). Let* μ *be the law of* $\zeta\Gamma$ *with* ζ *a Bernoulli variable of parameter* $p \in (0,1)$ *and* Γ *a centered Gaussian variable with variance* $1/p$. *For any* $x \in \mathbb{R}$,

$$L_\mu(x) = \log\left(pe^{\frac{x^2}{2p}} + 1 - p\right)$$

so that $A = B = \frac{1}{p}$.

- *(Combination of Rademacher laws). Let*

$$\mu = \sum_{i=1}^{p} \frac{\alpha_i}{2}(\delta_{\beta_i} + \delta_{-\beta_i})$$

with $\alpha_i \geq 0$, $\beta_i \in \mathbb{R}$ *and* $p \in \mathbb{N}$ *so that* $\sum \alpha_i = 1, \sum \alpha_i\beta_i^2 = 1$. *Since* μ *is compactly supported* $B = 0$. *The fact that* μ *does not have sharp sub-Gaussian tails means that there exist some* t *and* $A > 1$ *such that*

$$\sum_{i=1}^{p} \alpha_i \cosh(\beta_i t) \geq e^{A\frac{t^2}{2}}.$$

The latter is equivalent to

$$\sum_{i=1}^{p} \frac{\alpha_i}{2} e^{\frac{\beta_i^2}{2A}}\left(e^{-\frac{A}{2}(t-\frac{\beta_i}{A})^2} + e^{-\frac{A}{2}(t+\frac{\beta_i}{A})^2}\right) \geq 1.$$

This inequality holds as soon as $\alpha_i e^{\frac{\beta_i^2}{2A}} \geq 2$ *for some* $i \in \{1, \ldots, p\}$ *by taking* $t = \frac{\beta_i}{A}$. *This can be fulfilled if* β_i *is large enough while* $\alpha_i\beta_i^2 < 1$. *We also see with this*

family of examples that A can be taken arbitrarily large even if $B = 0$ (take e.g $p = 2$, $A = \beta_1, t = 1, \alpha_1 = (2\beta_1^2)^{-1}$, $e^{\beta_1/2} \geq 4\beta_1^2, \beta_2^2 = (2 - \beta_1^{-2})^{-1}, \alpha_2 = 1 - \alpha_1$).

Let $\mathcal{H}_N$ be the set of real symmetric matrices of size N. We denote for any $A \in \mathcal{H}_N$ by λ_A its largest eigenvalue, $||A||$ is spectral radius and by $\hat{\mu}_A$ the empirical distribution of its eigenvalues, that is

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

where $\lambda_1, \ldots, \lambda_N$ are the eigenvalues of A. We make the following assumption of exponential tightness of the spectral radius and of concentration of the empirical distribution of the eigenvalues at the scale N.

Assumption 1.2. *The spectral radius of X_N, $||X_N||$, is exponentially tight at the scale N:*

$$\lim_{K \to +\infty} \limsup_{N \to +\infty} \frac{1}{N} \log \mathbb{P}\big(||X_N|| > K\big) = -\infty. \tag{3}$$

Moreover, the empirical distribution of the eigenvalues $\hat{\mu}_{X_N}$ concentrates at the scale N:

$$\limsup_{N \to +\infty} \frac{1}{N} \log \mathbb{P}\left(d(\hat{\mu}_{X_N}, \sigma) > N^{-\kappa}\right) = -\infty, \tag{4}$$

for some $\kappa > 0$, where d is a distance compatible with the weak topology and σ is the semi-circle law, defined by

$$\sigma(dx) = \frac{1}{2\pi} \sqrt{4 - x^2} \mathbb{1}_{|x| \leq 2} dx.$$

Remark 1.3. (1) *From* [12, Lemmas 1.8, 1.11]*, we know that Assumption 1.2 is fulfilled if μ is either compactly supported, or if μ satisfies a logarithmic Sobolev inequality in the sense that there exists $c > 0$ so that for any smooth function $f : \mathbb{R} \to \mathbb{R}$, such that $\int f^2 d\mu = 1$,*

$$\int f^2 \log f^2 d\mu \leq c \int \|\nabla f\|_2^2 d\mu \,.$$

(2) *If μ is a symmetric sub-Gaussian probability measure on $\mathbb{R}$ with log-concave tails in the sense that $t \mapsto \mu(|x| \geq t)$ is a log-concave function, then the Wigner matrix X_N satisfies Assumption 1.2. In particular, if B is a Wigner matrix with Bernoulli entries with parameter p and Γ is a GOE matrix, then the sparse Gaussian matrix $B \circ \Gamma/\sqrt{p}$, where $\circ$ the Hadamard product, satisfies Assumption 1.2. We refer the reader to section 7.1 of the appendix for more details.*

1.2. **Statement of the results and scheme of the proof.** As in [12], our approach to derive large deviations estimates is based on a tilting of the law of the Wigner matrix X_N by spherical integrals. Let us recall the definition of spherical integrals. For any $\theta \geq 0$, we define

$$I_N(X_N, \theta) = \mathbb{E}_e[e^{\theta N \langle e, X_N e \rangle}]$$

where e is uniformly sampled on the sphere $\mathbb{S}^{N-1}$ with radius one. The asymptotics of

$$J_N(X_N, \theta) = \frac{1}{N} \log I_N(X_N, \theta)$$

were studied in [13] where the following result was proved.

Theorem 1.4. [13, Theorem 6] *Let $(E_N)_{N\in\mathbb{N}}$ be a sequence of $N\times N$ real symmetric matrices such that:*

- *The sequence of empirical measures $\hat{\mu}_{E_N}$ converges weakly to a compactly supported measure μ.*
- *There is a real number λ_E such that the sequence of largest eigenvalues λ_{E_N} converges to λ_E.*
- $\sup_N ||E_N|| < +\infty$.

For any $\theta \geq 0$,

$$\lim_{N\to+\infty} J_N(E_N,\theta) = J(\mu,\theta,\lambda_E)$$

The limit J is defined as follows. For a compactly supported probability measure $\mu\in\mathcal{P}(\mathbb{R})$ we define its Stieltjes transform G_μ by

$$\forall z\notin \mathrm{supp}(\mu),\ G_\mu(z) := \int_{\mathbb{R}} \frac{1}{z-t} d\mu(t),$$

where $\mathrm{supp}(\mu)$ is the support of μ. For any compactly supported probability measure μ, we denote by r_μ the right edge of the support of μ. Then G_μ is a bijection from $(r_\mu,+\infty)$ to $\big(0, G_\mu(r_\mu)\big)$ where

$$G_\mu(r_\mu) = \lim_{t\downarrow r_\mu} G_\mu(t).$$

Let K_μ be the inverse of G_μ on $(0, G_\mu(r_\mu))$ and let

$$\forall z\in(0,G_\mu(r_\mu)),\ R_\mu(z) := K_\mu(z) - 1/z,$$

be the R-transform of μ as defined by Voiculescu in [18]. Then, the limit of spherical integrals is defined for any $\theta\geq 0$ and $x\geq r_\mu$ by,

$$J(\mu,\theta,x) := \theta v(\mu,\theta,x) - \frac{1}{2}\int \log\big(1+2\theta v(\mu,\theta,x) - 2\theta y\big) d\mu(y),$$

with

$$v(\mu,\theta,x) := \begin{cases} R_\mu(2\theta) & \text{if } 0\leq 2\theta\leq G_\mu(x),\\ x-\frac{1}{2\theta} & \text{if } 2\theta > G_\mu(x).\end{cases}$$

In the case of the semi-circle law, we have

$$G_\sigma(x) = \frac{1}{2}(x-\sqrt{x^2-4}),\ R_\sigma(x) = x.$$

We denote by $J(\lambda,\theta)$ as a short-hand for $J(\sigma,\theta,\lambda)$. In the next lemma we compute explicitly $J(x,\lambda)$, whose proof is left to the reader.

Lemma 1.5. *Let $\theta\geq 0$ and $x\geq 2$. For $\theta\leq \frac{1}{2}G_\sigma(x)$,*

$$J(x,\theta) = \theta^2.$$

Whereas for $\theta\geq\frac{1}{2}G_\sigma(x)$,

$$J(x,\theta) = \theta x - \frac{1}{2} - \frac{1}{2}\log 2\theta - \frac{1}{2}\int \log(x-y) d\sigma(y).$$

Moreover,for any $x\geq 2$, $J(x,.)$ is continuously differentiable.

To derive large deviations estimates using a tilt by spherical integrals, it is central to obtain the asymptotics of the annealed spherical integral $F_N(\theta)$ defined as,

$$F_N(\theta) = \frac{1}{N} \log \mathbb{E}_{X_N} \mathbb{E}_e[\exp(N\theta \langle e, X_N e\rangle)] \,.$$

In the following lemma, we obtain the limit of F_N as the solution of a certain variational problem. We denote by $\overline{F}(\theta)$ and $\underline{F}(\theta)$ its upper and lower limits:

$$\overline{F}(\theta) = \limsup_{N\to+\infty} F_N(\theta),$$

$$\underline{F}(\theta) = \liminf_{N\to+\infty} F_N(\theta).$$

For any measurable subset $I \subset \mathbb{R}$, we denote by $\mathcal{M}(I)$ and $\mathcal{P}(I)$ respectively the set of measures and the set of probability measures supported on I.

Proposition 1.6. *Assume X_N satisfies Assumptions 1.1 and 1.2.*

$$\overline{F}(\theta) = \limsup_{\substack{\delta\to 0, K\to+\infty \\ \delta K\to 0}} \sup_{\substack{\alpha_1+\alpha_2+\alpha_3=1 \\ \alpha_i\geq 0}} \limsup_{N\to+\infty} \mathcal{F}^N_{\alpha_1,\alpha_2,\alpha_3}(\delta, K),$$

$$\underline{F}(\theta) = \sup_{\substack{\alpha_1+\alpha_2+\alpha_3=1 \\ \alpha_i\geq 0}} \liminf_{\substack{\delta\to 0, K\to+\infty \\ \delta K\to 0}} \limsup_{N\to+\infty} \mathcal{F}^N_{\alpha_1,\alpha_2,\alpha_3}(\delta, K) \,.$$

$\mathcal{F}^N_{\alpha_1,\alpha_2,\alpha_3}(\delta, K)$ is the function given by:

$$\begin{aligned}
\mathcal{F}^N_{\alpha_1,\alpha_2,\alpha_3}(\delta, K) &= \theta^2\big(\alpha_1^2 + 2\alpha_1\alpha_2 + B\alpha_3^2\big) \\
&+ \sup_{\substack{t_i\in I_2, i\leq l \\ |\sum_i t_i^2 - N\alpha_2|\leq \delta N}} \sup_{\substack{s_i\in I_3, i\leq k \\ |\sum_i s_i^2 - N\alpha_3|\leq \delta N}} \Big\{ \frac{1}{N}\sum_{i=1}^{k}\sum_{j=1}^{l} L\Big(\frac{2\theta s_i t_j}{\sqrt{N}}\Big) + \frac{1}{2N}\sum_{i,j=1}^{l} L\Big(\frac{2\theta t_i t_j}{\sqrt{N}}\Big) \\
&+ \sup_{\substack{\nu_1\in\mathcal{P}(I_1) \\ \int x^2 d\nu_1(x) = \alpha_1}} \Big\{ \sum_{i=1}^{k} \int L\Big(\frac{2\theta s_i x}{\sqrt{N}}\Big) d\nu_1(x) - H(\nu_1)\Big\} - \frac{1}{2}\log(2\pi) - \frac{1}{2}\Big\},
\end{aligned}$$

where $I_1 = \{x : |x| \leq \delta^{1/2} N^{1/4}\}$, $I_2 = \{x : \delta^{1/2} N^{1/4} \leq |x| \leq K^{1/2} N^{1/4}\}$, $I_3 = \{K^{1/2} N^{1/4} \leq |x| \leq \sqrt{N\alpha_3}\}$, and

$$H(\nu) = \int \log \frac{d\nu}{dx} d\nu(x),$$

if ν is absolutely continuous with respect to the Lebesgue measure, whereas $H(\nu)$ is infinite otherwise.

Remark 1.7. *Note that $\underline{F}$ and $\overline{F}$ are convex by Hölder inequality. Since the entries of X_N are sub-Gaussian, $\overline{F}(\theta) \leq A\theta^2$. In particular $\overline{F}, \underline{F}$ are finite convex functions and therefore they are continuous on $\mathbb{R}_+$.*

The above proposition gives quite an intricate definition for the limit of the annealed spherical integrals. Yet, for small enough θ it can be computed explicitly.

Lemma 1.8. *For any $\theta \leq \frac{1}{2\sqrt{A-1}}$,*

$$\underline{F}(\theta) = \overline{F}(\theta) = \theta^2.$$

Note that for large θ this formula is not valid anymore when $A > 1$ since $\underline{F}$ grows like $A\theta^2$ at infinity (see the proof of Proposition 1.10).

Proof. Using the bound $L(x) \leq Ax^2/2$ for any $x \geq 0$ and the notation of Proposition 1.6, we have

$$\mathcal{F}^N_{\alpha_1,\alpha_2,\alpha_3}(\delta, K) \leq \sup_{\alpha_1+\alpha_2+\alpha_3=1} \left\{ \theta^2\big(\alpha_1^2 + 2\alpha_1\alpha_2 + B\alpha_3^2 + 2A\alpha_3\alpha_2 + A\alpha_2^2 + 2A\alpha_1\alpha_3\big) + \frac{1}{2}\log\alpha_1 \right\}.$$

Here we used the fact that the infimum

$$\inf\{H(\nu_1) : \int x^2 d\nu_1 = \alpha_1, \nu_1 \in \mathcal{P}(I_1)\} \geq \inf\{H(\nu_1) : \int x^2 d\nu_1 = \alpha_1, \nu_1 \in \mathcal{P}(\mathbb{R})\},$$

where the infimum in the right hand side is achieved at $\nu_1(dx) = (2\pi\alpha_1)^{-1/2} e^{-\frac{x^2}{2\alpha_1}} dx$ and hence equals $-1/2(1 + \log(2\pi\alpha_1))$.

As $A \geq 1$ and $B \leq A$, we have the upper bound,

$$\begin{aligned}\overline{F}(\theta) &\leq \sup_{\alpha\in[0,1]} \left\{ \theta^2\big(\alpha^2 + 2A\alpha(1-\alpha) + A(1-\alpha)^2\big) + \frac{1}{2}\log\alpha \right\} \\ &= \sup_{\alpha\in[0,1]} \left\{ \theta^2\big(A - (A-1)\alpha^2\big) + \frac{1}{2}\log\alpha \right\}.\end{aligned}$$

Hence for all $\theta \geq 0$, (and as we could have seen directly from the uniform upper bound $L(\theta) \leq \frac{A}{2}\theta^2$)

$$\overline{F}(\theta) \leq A\theta^2\,. \tag{5}$$

We see that if $2\theta\sqrt{A-1} \leq 1$ then the function

$$\alpha \mapsto \theta^2\big(A - (A-1)\alpha^2\big) + \frac{1}{2}\log\alpha,$$

is increasing on $[0, 1]$. Thus the supremum is achieved at $\alpha = 1$, and $\overline{F}(\theta) \leq \theta^2$. Moreover, taking $\alpha_1 = 1, \alpha_2 = \alpha_3 = 0$, and ν_1 the standard Gaussian restricted to I_1 $\nu_1(dx) = \mathbb{1}_{I_1} e^{-\frac{x^2}{2}} dx/Z$ we find that

$$\underline{F}(\theta) \geq \theta^2. \tag{6}$$

Thus, if $2\theta\sqrt{A-1} \leq 1$, we get that $\overline{F}(\theta) = \underline{F}(\theta) = \theta^2$.

□

Although the limit of the annealed spherical integrals may not be explicit for all θ, we can still use it to obtain large deviations upper bounds as we describe now in the following theorem.

Theorem 1.9. *Under the Assumptions 1.1 and 1.2, the law of the largest eigenvalue λ_{X_N} satisfies a large deviation upper bound with good rate function $\bar{I}$ which is infinite on $(-\infty, 2)$ and otherwise given by:*

$$\forall y \geq 2,\ \bar{I}(y) = \sup_{\theta\geq 0}\{J(y,\theta) - \overline{F}(\theta)\}\,. \tag{7}$$

Moreover, $\bar{I}(y) \leq I_{GOE}(y)$ for all $y \geq 2$.

Proof. From Assumption 1.2 we know that the law of the largest eigenvalue is exponentially tight at the scale N. Therefore, it is sufficient to prove a weak large deviations upper bound by [9, Lemma 1.2.18]. Let $\delta > 0$. We have,

$$\mathbb{P}(\lambda_{X_N} < 2 - \delta) \leq \mathbb{P}(\hat{\mu}_{X_N}(f) = 0),$$

where f is a smooth compactly supported function with support in $(2-\delta, 2)$. Since $\mathrm{supp}(\sigma) = [-2, 2]$, we deduce that,

$$\mathbb{P}(\lambda_{X_N} < 2 - \delta) \leq \mathbb{P}(d(\hat{\mu}_{X_N}, \sigma) > \varepsilon),$$

for some $\varepsilon > 0$. As the empirical distribution of the eigenvalues concentrates at the scale N according to (4), we conclude that

$$\lim_{N \to +\infty} \frac{1}{N} \log \mathbb{P}(\lambda_{X_N} < 2 - \delta) = -\infty.$$

Let now $x \geq 2$ and $\delta > 0$. Recall from (6) that $\underline{F}(\theta) \geq \theta^2$ for any $\theta \geq 0$. Therefore,

$$\bar{I}(x) \leq \sup_{\theta \geq 0} \{J(x, \theta) - \theta^2\}.$$

From [12, Section 4.1], we know that

$$\sup_{\theta \geq 0} \{J(x, \theta) - \theta^2\} = I_{GOE}(x),$$

where I_{GOE} is the rate function of the largest eigenvalue of a GOE matrix. Therefore we have proved that

$$\bar{I}(x) \leq I_{GOE}(x), \forall x \geq 2\,.$$

In particular $\bar{I}(2) = 0$ since $I_{GOE}(2)$. Therefore we only need to estimate small ball probabilities around $x \neq 2$. As $\hat{\mu}_{X_N}$ concentrates at the scale N, and $||X_N||$ is exponentially tight at the scale N by Assumption 1.2 it is enough to show that for any $K > 0$,

$$\limsup_{\delta \to 0} \limsup_{N \to +\infty} \frac{1}{N} \log \mathbb{P}(X_N \in V_{\delta,x}^K) \leq -\bar{I}(x),$$

where $V_{\delta,x}^K = \{Y \in \mathcal{H}_N : |\lambda_Y - x| < \delta, d(\hat{\mu}_Y, \sigma) < N^{-\kappa}, ||Y|| \leq K\}$, for some $\kappa > 0$. Let $\theta \geq 0$. From [16, Proposition 2.1], we know that the spherical integral is continuous, more precisely, for N large enough and any $X_N \in V_{\delta,x}$,

$$|J_N(X_N, \theta) - J(x, \theta)| < g(\delta),$$

for some function $g(\delta)$ going to 0 as $\delta \to 0$. Therefore,

$$\mathbb{P}(X_N \in V_{\delta,x}^K) = \mathbb{E}\Big(\mathbb{1}_{X_N \in V_{\delta,x}^K} \frac{I_N(X_N, \theta)}{I_N(X_N, \theta)}\Big) \leq \mathbb{E}[I_N(X_N, \theta)] e^{-NJ(\theta,x) - Ng(\delta)}.$$

Taking the limsup as $N \to 0$ and $\delta \to 0$ at the logarithmic scale, we deduce

$$\limsup_{\delta \to 0} \limsup_{N \to +\infty} \frac{1}{N} \log \mathbb{P}(X_N \in V_{\delta,x}^K) \leq \overline{F}(\theta) - J(\theta, x).$$

Opimizing over $\theta \geq 0$, we get the claim. □

Proposition 1.10. *Under Assumption 1.1, the rate function $\bar{I}$ defined in Theorem 1.9 is lower semi-continuous, and growing at infinity like $x^2/4A$. In particular, $\bar{I}$ is a good rate function.*

Proof. $\bar{I}$ is lower semi-continuous as a supremum of continuous functions (recall here that $J(\theta,.)$ is continuous by Lemma 1.5 and $\overline{F}$ is continuous by Remark 1.5). It remains to show that its level sets are compact, for which it is sufficient to prove that $\bar{I}$ goes to infinity at infinity. Let $x > 2$. Let $C > 0$ be a constant to be chosen later such that $Cx \geq 1/2$. We have by taking $\theta = Cx$ and using (1.9), that

$$\begin{aligned}\bar{I}(x) &\geq J(Cx, x) - \overline{F}(Cx) \\ &\geq Cx^2 - \frac{1}{2} - \frac{1}{2}\log\frac{x}{A} - \frac{1}{2}\log x - AC^2x^2. \end{aligned} \tag{8}$$

Taking $C = 1/2A$, and assuming that $x > A$, we obtain that

$$\bar{I}(x) \geq \frac{x^2}{4A} - o(x^2). \tag{9}$$

To get the converse bound, we show that as θ goes to infinity, $\overline{F}$ goes to infinity like $A\theta^2$. We distinguish two cases. First, we consider the case $A = B$. Using Proposition 1.6, we get the lower bound for $\theta \geq 1$,

$$\underline{F}(\theta) \geq A\theta^2\Big(1 - \frac{1}{\theta^2}\Big) - \frac{1}{4}\log\theta,$$

by taking $\alpha_2 = 0$, $\alpha_3 = 1 - \theta^{-2}$, $\alpha_1 = \theta^{-2}$ and ν_1 the Gaussian law restricted to I_1 with variance α_1. In the case $A > B$, we define m_* such that $\psi(m_*) = A/2$. Taking $\alpha_3 = 0$, $\alpha_2 = 1 - \theta^{-2}$, $\alpha_1 = \theta^{-2}$, $\nu_2 = \frac{2\theta\alpha_2}{m_*}\delta_{\sqrt{\frac{m_*}{2\theta}}}$, and ν_1 the Gaussian law restricted to I_1 with variance α_1, we obtain,

$$\underline{F}(\theta) \geq A\theta^2\Big(1 - \frac{1}{\theta^2}\Big) - \log\theta. \tag{10}$$

It follows that for any $\varepsilon > 0$, there exists $M < \infty$ such that for $\theta \geq M$,

$$\underline{F}(\theta) \geq (1-\varepsilon)A\theta^2.$$

Therefore

$$\bar{I}(x) \leq \max\left\{\sup_{\theta\geq M}\{J(\theta,x) - (1-\varepsilon)A\theta^2\}, \sup_{\theta\leq M}\{J(\theta,x) - \overline{F}(\theta)\}\right\}.$$

But from Lemma 1.5 one can see that the second term in the above right-hand side is bounded by $Mx + C$ where C is a numerical constant. Besides, using the same argument as in (8), we get

$$\sup_{\theta\geq M}\{J(\theta,x) - (1-\varepsilon)A\theta^2\} \geq \frac{x^2}{4(1-\varepsilon)A} - o(x^2).$$

Hence, for x large enough,

$$\bar{I}(x) \leq \sup_{\theta\geq M}\{J(\theta,x) - (1-\varepsilon)A\theta^2\}.$$

But, for x large enough and $\theta \geq 1/2$, $J(\theta,x) \leq \theta x$. Thus,

$$\sup_{\theta\geq M}\{J(\theta,x) - (1-\varepsilon)A\theta^2\} \leq \sup_{\theta\geq 0}\{\theta x - (1-\varepsilon)A\theta^2\} = \frac{x^2}{4(1-\varepsilon)A},$$

which ends the proof.

□

Proposition 1.11. *For any $\theta \geq 0$, $J(\theta, .)$ is a convex function. Therefore, $\bar{I}$ is also convex.*

Proof. Let $x, y \geq 2$ and $t \in (0,1)$. Let E_N be a sequence of diagonal matrices such that $||E_N|| \leq 2$ and such that $\hat{\mu}_{E_N}$ converges weakly to σ. Let E_N^x and E_N^y be such that $(E_N^x)_{i,i} = (E_N^y)_{i,i} = (E_N)_{i,i}$ for any $i \in \{1, \dots, N-1\}$, and

$$(E_N^x)_{N,N} = x \quad (E_N^y)_{N,N} = y.$$

We have $\lambda_{E_N^x} = x$ and $\lambda_{E_N^y} = y$. Then, $H_N = tE_N^x + (1-t)E_N^y$ is such that its empirical distribution of eigenvalues converges to σ, and $\lambda_{H_N} = tx + (1-t)y$. By Hölder's inequality we have,

$$\log I_N(H_N, \theta) \leq t \log I_N(E_N, \theta) + (1-t) \log I_N(D_N, \theta).$$

Taking the limit as $N \to +\infty$, we get,

$$J(tx + (1-t)y, \theta) \leq tJ(x, \theta) + (1-t)J(y, \theta).$$

Therefore, $J(\theta, .)$ is convex and I is convex as a supremum of convex functions. □

To derive the large deviation lower bound, we denote by $\mathcal{C}_\mu$ the set of $\theta \in \mathbb{R}^+$ such that

$$\underline{F}(\theta) = \overline{F}(\theta) =: F(\theta)\,.$$

By Lemma 1.8, $\mathcal{C}_\mu$ is not empty. We observe also that by continuity of both $\underline{F}$ and $\overline{F}$ (see Remark 1.7), $\mathcal{C}_\mu$ is closed. Let

$$\forall x \geq 2,\ I(x) = \sup_{\theta \in \mathcal{C}_\mu} \{J(x, \theta) - F(\theta)\}$$

Theorem 1.12. *For any $x \geq 2$, denote by*

$$\Theta_x = \{\theta \geq 0 : \bar{I}(x) = J(x, \theta) - F(\theta)\},$$

where $\bar{I}$ is defined in (7). Let $x \geq 2$ such that there exists $\theta \in \Theta_x \cap \mathcal{C}_\mu$ and $\theta \notin \Theta_y$ for any $y \neq x$. Then, $I(x) = \bar{I}(x)$ and

$$\lim_{\delta \to 0} \liminf_{N \to +\infty} \frac{1}{N} \log \mathbb{P}\left(|\lambda_{X_N} - x| \leq \delta\right) \geq -I(x).$$

We apply this general theorem in two cases. We first investigate the case where the function ψ is increasing, case for which we can check that our hypotheses on the sets Θ_x holds for x large enough. This includes the case where μ is the sparse Gaussian law, see Example 1.2.

Proposition 1.13. *Suppose that Assumptions 1.1 and 1.2 hold. If ψ is increasing on $\mathbb{R}_+$, then $\mathcal{C}_\mu = \mathbb{R}^+$. Moreover, there exists $x_\mu \geq 2$ such that for any $x \geq x_\mu$, the large deviation lower bound holds with rate function I.*

We then consider the case where μ is such that $B < A$. This includes any compactly supported measure μ since then $B = 0$. We prove in this case the following result.

Proposition 1.14. *Suppose that Assumptions 1.1 and 1.2 hold. If μ is such that $B < A$ and such that the maximum of ψ is attained on $\mathbb{R}^+$ for a unique m_* such that $\psi''(m_*) < 0$, then there exists a positive finite real number θ_0 such that for $[\theta_0, +\infty[\subset \mathcal{C}_\mu$. Therefore, there exists a finite constant x_μ such that for $x \geq x_\mu$, the large deviation lower bound holds with rate function I. Furthermore on the interval $[x_\mu, +\infty)$ the rate function I depends only on A.*

In the case where A is sufficiently small, we can show without any additional assumption that the large deviation lower bound holds in a vicinity of 2 and the rate function I is equal to the one of the GOE. This contrasts with Proposition 1.10 which shows that the rate function $\bar{I}$ goes to infinity like $x^2/4A$ at infinity and therefore depends on A. In other words the "heavy tails" only kicks in above a certain threshold.

Proposition 1.15. *Assume $A < 2$. The large deviation lower bound holds with rate function $\bar{I}$ on $[2, 1/\sqrt{A-1} + \sqrt{A-1}]$. Moreover, $\bar{I}$ coincides with the rate function in the GOE case I_{GOE}, defined in (2), on this interval. As a consequence, for all $x \in [2, 1/\sqrt{A-1} + \sqrt{A-1}]$*

$$\lim_{\delta \to 0} \liminf_{N \to +\infty} \frac{1}{N} \log \mathbb{P}\left(|\lambda_{X_N} - x| \leq \delta\right) = \lim_{\delta \to 0} \limsup_{N \to +\infty} \frac{1}{N} \log \mathbb{P}\left(|\lambda_{X_N} - x| \leq \delta\right) = -I_{GOE}(x).$$

In the next section 2, we detail our approach to prove large deviations lower bounds. Since Proposition 1.6 is crucial to all our results, we prove it in the next section 3. Then, we will apply these results to prove the large deviations lower bounds close to the bulk in section 4, that is, we give a proof of Proposition 1.15. To prove the large deviations lower bounds for large x, we consider first the case of increasing ψ in section 5 and then the case of $B < A$ in section 6. Indeed, the variational formulas for the limiting annealed spherical integrals differ in these two cases, as $B = A$ in the first case whereas $B < A$ in the second.

2. A general large deviation lower bound

We first prove Theorem 1.12 and will then give more practical descriptions of the sets Θ_x in order to apply it.

Proof of Theorem 1.12. By assumption, there exists $\theta \in \Theta_x \cap \mathcal{C}_\mu$ such that $\theta \notin \Theta_y$ for $y \neq x$. In particular, it entails that $I(x) = \bar{I}(x)$. Introducing the spherical integral with parameter $\theta \geq 0$, we have

$$\mathbb{P}\left(|\lambda_{X_N} - x| \leq \delta\right) \geq \mathbb{E}\Big(\mathbb{1}_{X_N \in V_{\delta,x}^K} \frac{I_N(X_N, \theta)}{I_N(X_N, \theta)}\Big),$$

where $V_{\delta,x}^K = \{Y \in \mathcal{H}_N : |\lambda_Y - x| \leq \delta, d(\hat{\mu}_Y, \sigma) < N^{-\kappa}, ||X_N|| \leq K\}$ for some $K > 0$ and $\kappa > 0$. Using the continuity of the spherical integral (see [16, Proposition 2.1]), we get

$$\mathbb{P}\left(|\lambda_{X_N} - x| \leq \delta\right) \geq \frac{\mathbb{E}\big(\mathbb{1}_{X_N \in V_{\delta,x}^K} I_N(X_N, \theta)\big)}{\mathbb{E} I_N(X_N, \theta)} e^{N\underline{F}(\theta) - NJ(x,\theta) - Ng(\delta) - o(N)}, \tag{11}$$

where g is a function such that $g(\delta) \to 0$ as $\delta \to 0$. We claim that

$$\liminf_{N\to+\infty} \frac{1}{N} \log \frac{\mathbb{E}\big(\mathbb{1}_{X_N \in V_{\delta,x}^K} I_N(X_N, \theta_x)\big)}{\mathbb{E} I_N(X_N, \theta_x)} \geq 0.$$

To this end we will use our large deviation upper bound. Since $\hat{\mu}_{X_N}$ concentrates at scales faster than N, and $||X_N||$ is exponentially tight at the scale N by Assumption 1.2, this remains true under the measure tilted by the spherical integral since the logarithm of its density grows at most like N. Hence, it suffices to prove that for all $y \neq x$, for δ small enough, and K large enough,

$$\limsup_{N} \frac{1}{N} \log \frac{\mathbb{E}[\mathbb{1}_{X_N \in V_{\delta,y}^K} I_N(X_N, \theta)]}{\mathbb{E} I_N(X_N, \theta)} < 0.$$

By assumption, there exists $\theta \in \Theta_x \cap \mathcal{C}_\mu$ such that $\theta \notin \Theta_y$ for $y \neq x$. We introduce a new spherical integral with argument θ' and use again the continuity of J_N to show that:

$$\begin{aligned}
\frac{\mathbb{E}[\mathbb{1}_{X_N \in V_{\delta,y}^K} I_N(X_N, \theta)]}{\mathbb{E} I_N(X_N, \theta)} &= \frac{\mathbb{E}[\mathbb{1}_{X_N \in V_{\delta,y}^K} \frac{I_N(X_N,\theta')}{I_N(X_N,\theta')} I_N(X_N, \theta)]}{\mathbb{E} I_N(X_N, \theta)} \\
&= e^{-NJ(y,\theta')-N\underline{F}(\theta)+NJ(y,\theta)+N\varepsilon(\delta)} \mathbb{E}[1_{X_N \in V_{\delta,y}^K} I_N(X_N, \theta')] \\
&\leq e^{-NJ(y,\theta')-N\underline{F}(\theta)+NJ(y,\theta)+N\overline{F}(\theta')+N\varepsilon(\delta)},
\end{aligned}$$

where $\varepsilon(\delta) \to 0$ as $\delta \to 0$. We can conclude that

$$\begin{aligned}
\limsup_{\substack{N\to+\infty \\ \delta\to 0}} \frac{1}{N} \log \frac{\mathbb{E}[\mathbb{1}_{X_N \in V_{\delta,y}^K} I_N(X_N, \theta)]}{\mathbb{E} I_N(X_N, \theta)} &\leq -\sup_{\theta'} \{J(y, \theta') - \overline{F}(\theta')\} + J(y, \theta) - \underline{F}(\theta) \\
&= -\overline{I}(y) + J(y, \theta) - \underline{F}(\theta) \qquad (12)
\end{aligned}$$

By assumption, $\theta \notin \Theta_y$, and $\theta \in \mathcal{C}_\mu$ so that $\underline{F}(\theta) = \overline{F}(\theta)$ hence

$$-\overline{I}(y) + J(y, \theta) - \overline{F}(\theta) < 0$$

and the conclusion follows from (12). Therefore, coming back to (11), we obtain since $\theta_x \in \Theta_x$ and $I(x) = \bar{I}(x)$,

$$\liminf_{N\to+\infty} \frac{1}{N} \log \mathbb{P}\left(|\lambda_{X_N} - x| \leq \delta\right) \geq -I(x).$$

□

In a first step, we identify a subset defined in terms of the subdifferential sets of F at the points of non-differentiability where the large deviation lower bound holds. Let $\mathcal{D}$ be the set of $\theta \geq 0$ such that $\overline{F}$ is differentiable at θ.

Lemma 2.1. *The lower bound holds for any $x > 2$ such that $I(x) = \bar{I}(x) > 0$ and*

$$x \notin E := \bigcup_{\theta \in \mathcal{D}^c} \Big(\frac{1}{2\theta} + \partial\overline{F}(\theta)\Big), \qquad (13)$$

where $\partial\overline{F}(\theta)$ denotes the subdifferential of $\overline{F}$ at θ

Note that since $\overline{F}$ is a convex function, its subdifferentials are well defined.

Proof. Let $x > 2$ such that $I(x) = \bar{I}(x) > 0$ and $x \notin E$. Since $\overline{F}(\theta) \geq \theta^2$ for any $\theta \geq 0$ by (6) and $\overline{F}$ is continuous by Remark 1.7, we deduce from Lemma 1.5 that $\theta \mapsto J(x,\theta) - \overline{F}(\theta)$ is continuous and goes to $-\infty$ as θ goes to $+\infty$. Since $\mathcal{C}_\mu$ is closed, the supremum

$$\sup_{\theta \in \mathcal{C}_\mu} \left\{ J(\theta, x) - \overline{F}(\theta) \right\} \tag{14}$$

is achieved at some $\theta \in \mathcal{C}_\mu$. We will show that $\theta \in \mathcal{D}$.

As $\bar{I}(x) \neq 0$ we must have $\theta > \frac{1}{2} G_\sigma(x)$. Indeed, $\overline{F}(\theta) \geq \theta^2$ and $J(\theta, x) = \theta^2$ for $\theta \leq \frac{1}{2} G_\sigma(x)$ by Lemma 1.5 so that

$$\sup_{\theta \leq \frac{1}{2} G_\sigma(x)} \{ J(\theta, x) - \overline{F}(\theta) \} = 0.$$

Since $\bar{I}(x) = I(x)$, we deduce by Fermat's rule that θ is a critical point of $J(., x) - \overline{F}$ and therefore satisfies the condition:

$$0 \in \frac{\partial J}{\partial \theta}(\theta, x) - \partial \overline{F}(\theta) = x - \frac{1}{2\theta} - \partial \overline{F}(\theta)\,.$$

Since $x \notin E$, we deduce that $\overline{F}$ is differentiable at θ.

According to Theorem 1.12, to prove that the lower bound holds at x, it suffices to show that $\theta \notin \Theta_y$ for any $y \neq x$. Let us proceed by contradiction and assume that there exists $y \geq 2$ such that $\theta \in \Theta_y$. As $\overline{F}$ is differentiable at θ, it should be a critical point of both $J(y, .) - \overline{F}$ and $J(x, .) - \overline{F}$. Therefore, we should have

$$\frac{\partial}{\partial \theta} J(y, \theta) = \frac{\partial}{\partial \theta} J(x, \theta).$$

If $G_\sigma(y) < 2\theta$, then we obtain by Lemma 1.5 and the fact that $G_\sigma(x) < 2\theta$ that $x = y$. If $G_\sigma(y) \geq 2\theta$, then we have

$$x - \frac{1}{2\theta} = 2\theta.$$

On the other hand, $2\theta \leq G_\sigma(y) \leq 1$ and therefore we get the unique solution $2\theta = G_\sigma(x)$. As we assumed that $2\theta > G_\sigma(x)$, we get a contradiction and conclude that $\theta \notin \Theta_y$ for any $G_\sigma(y) \geq 2\theta$, which completes the proof. □

We are now ready to prove the following result:

Proposition 2.2. *Assume that there exists $\theta_0 > 0$ such that $[\theta_0, +\infty[\subset \mathcal{C}_\mu$ and such that $\overline{F}$ is differentiable on $(\theta_0, +\infty)$. There exists $x_\mu \geq 2$ such that for any $x \geq x_\mu$, $I(x) = \bar{I}(x)$ and the large deviation lower bound holds for any $x \geq x_\mu$ with rate function $I(x)$.*

Proof. On one hand,

$$\sup_{\theta \leq \theta_0} \{ J(\theta, x) - \overline{F}(\theta) \} \leq \theta_0 x + C,$$

where C is some positive constant. Since $\bar{I}(x) \geq x^2/4A - o(x^2)$ by (9), we deduce that there exists $x_\mu \geq 2$ such that for $x \geq x_\mu$, $\bar{I}(x) > 0$ and the supremum of $J(., x) - \overline{F}$ is achieved on $[\theta_0, +\infty)$, in particular on $\mathcal{C}_\mu$. Thus, for any $x \geq x_\mu$, $\bar{I}(x) = I(x) > 0$. In

view of Lemma 2.1, it remains to show that E, defined in (13), is a bounded set. From our assumption that $\overline{F}$ is differentiable on $(\theta_0, +\infty)$ and Lemma 1.8, we deduce that

$$\mathcal{D}^c \subset \Big[\frac{1}{2\sqrt{A-1}}, \theta_0\Big].$$

We observe that since $0 \leq \overline{F}(\theta) \leq A\theta^2$, we have for any $\zeta \in \partial\overline{F}(\theta)$,

$$\zeta\theta \leq \overline{F}(2\theta) - \overline{F}(\theta) \leq 4A\theta^2,$$

and thus $\zeta \leq 4A\theta$. Therefore, E is a bounded set, which ends the proof.

□

3. Asymptotics of the annealed spherical integral

In this section we prove Proposition 1.6. Taking the expectation first with respect to X_N, we have

$$\begin{aligned} F_N(\theta) &= \frac{1}{N}\log \mathbb{E}_{X_N}\mathbb{E}_e[\exp(N\theta\langle e, X_N e\rangle)] \\ &= \frac{1}{N}\log \mathbb{E}_e \exp\big(f(e)\big), \end{aligned}$$

where

$$f(e) = \sum_{i<j} L\big(2\sqrt{N}\theta e_i e_j\big) + \sum_i L(\sqrt{2N}\theta e_i^2).$$

In a first step, we will prove the following variational representation of the upper and lower limits of $F_N(\theta)$.

Lemma 3.1. *Let X_N be a Wigner matrix satisfying Assumptions 1.1 and 1.2. Then for any $\theta > 0$,*

$$\underline{F}(\theta) \leq \liminf_{N\to+\infty} F_N(\theta) \leq \limsup_{N\to+\infty} F_N(\theta) \leq \overline{F}(\theta)$$

with

$$\begin{aligned} \overline{F}(\theta) &= \limsup_{\substack{\delta\to 0, K\to+\infty \\ \delta K\to 0}} \sup_{\substack{c=c_1+c_2+c_3 \\ c_i\geq 0}} \limsup_{N\to+\infty} F^N_{c_1,c_2,c_3}(\delta, K), \\ \underline{F}(\theta) &= \sup_{\substack{c=c_1+c_2+c_3 \\ c_i\geq 0}} \liminf_{\substack{\delta\to 0, K\to+\infty \\ \delta K\to 0}} \liminf_{N\to+\infty} F^N_{c_1,c_2,c_3}(\delta, K), \end{aligned}$$

where

$$\begin{aligned} F^N_{c_1,c_2,c_3}(\delta, K) = \sup_{\substack{s_i\geq\sqrt{cK}N^{1/4} \\ |\sum s_i^2 - c_3 N|\leq\delta N}} \sup_{\substack{\sqrt{c\delta}\leq t_i N^{-1/4}\leq\sqrt{cK} \\ |\sum t_i^2 - Nc_2|\leq\delta N}} \Big\{ &\frac{\theta^2}{c^2}\big(c_1^2 + 2c_1c_2 + Bc_3^2\big) - \frac{1}{2}\big(c_2 + c_3\big) \\ &+ \frac{1}{N}\sum_{i,j} L\Big(\frac{2\theta s_i t_j}{\sqrt{N}c}\Big) + \frac{1}{2N}\sum_{i,j} L\Big(\frac{2\theta t_i t_j}{\sqrt{N}c}\Big) + \sup_{\substack{\nu\in\mathcal{P}(I_1) \\ \int x^2 d\nu_1(x) = c_1}} \big\{\Phi(\nu, s) - H(\nu|\gamma)\big\}\Big\}, \end{aligned}$$

and

$$\Phi(\nu, s) = \sum_{i=1}^k \int L\Big(\frac{2\theta x s_i}{\sqrt{N}c}\Big) d\nu(x),$$

with $I_1 = [-\sqrt{c\delta}N^{1/4}, \sqrt{c\delta}N^{1/4}]$. *Here, we have set* γ *to be the standard Gaussian law and*

$$H(\nu|\gamma) = \int \ln \frac{d\nu}{d\gamma}(x) d\nu(x)\,.$$

Proof. We use the representation of the law of the vector e uniformly distributed on the sphere as a renormalized Gaussian vector $g/\|g\|_2$ where g is a standard Gaussian vector in $\mathbb{R}^N$, to write

$$\mathbb{E}_e \exp\big(f(e)\big) = \mathbb{E}\left[\exp(\Sigma(g))\right],$$

where

$$\Sigma(g) = \sum_{i<j} L\Big(2\sqrt{N}\theta \frac{g_i g_j}{\sum_{i=1}^N g_i^2}\Big) + \sum_i L\Big(\sqrt{2N}\theta \frac{g_i^2}{\sum_{i=1}^N g_i^2}\Big).$$

To study the large deviation of $\Sigma(g)$, we split the entries of g into three possible regime: the regime where $g_i \ll N^{1/4}$, an intermediate regime where $g_i \simeq N^{1/4}$ and finally $g_i \gg N^{1/4}$. Fix some $K, \delta > 0$ and $0 < 4\delta < K^{-1}$. Let $c_1, c_2, c_3 > 0$ and $c = c_1 + c_2 + c_3$. We assume that $0 < K^{-1} \le c_1 \le c \le K$. Define I_1, I_2, I_3 as

$$\begin{aligned} I_1 &= [0, \sqrt{c\delta}N^{\frac{1}{4}}] \\ I_2 &= [\sqrt{c\delta}N^{\frac{1}{4}}, \sqrt{cK}N^{\frac{1}{4}}] \\ I_3 &= [\sqrt{cK}N^{\frac{1}{4}}, \sqrt{N(c+3\delta)}]. \end{aligned}$$

Let for $i = 1, 2, 3$, $J_i = \{j : |g_j| \in I_i\}$ and $\hat{c}_i^N = \sum_{j\in J_i} g_j^2/N$. We will fix the empirical variances $\hat{c}_i^N$ in a first step. We will compute the asymptotics of

$$F^N_{c_1,c_2,c_3}(\theta,\delta) = \mathbb{E}[\exp(\Sigma(g))\mathbb{1}_{\mathcal{A}^\delta_{c_1,c_2,c_3}}].$$

where

$$\mathcal{A}^\delta_{c_1,c_2,c_3} := \bigcap_{1\le i\le 3} \{|\hat{c}_i^N - c_i| \le \delta\}.$$

Let

$$\Sigma_c(g) = \sum_{i<j} L\Big(2\theta \frac{g_i g_j}{\sqrt{N}c}\Big) + \sum_i L\Big(\sqrt{2}\theta \frac{g_i^2}{\sqrt{N}c}\Big).$$

Using the fact the $L(\sqrt{\cdot})$ is Lipschitz, we prove in the next lemma that on the event $\mathcal{A}^\delta_{c_1,c_2,c_3}$, $\Sigma_c(g)$ is a good approximation of $\Sigma(g)$.

Lemma 3.2. *On the event* $\mathcal{A}^\delta_{c_1,c_2,c_3}$,

$$\Sigma(g) - \Sigma_c(g) = N o_{\delta K}(1), \quad \text{as } \delta K \to 0. \tag{15}$$

Moreover

$$|J_3| \le \frac{2\sqrt{N}}{K}, \qquad |J_2 \cup J_3| \le \frac{2\sqrt{N}}{\delta}.$$

Proof. Note that since μ is symmetric, $L(x) = L(|x|)$ and since we assumed $L(\sqrt{\cdot})$ Lipschitz, for any $x, y \in \mathbb{R}$, $|L(x) - L(y)| \le L|x^2 - y^2|$ for some finite constant L. Therefore,

$$\sum_{1\le i\neq j\le N} \Big|L\Big(2\sqrt{N}\theta\frac{g_ig_j}{\sum_{i=1}^N g_i^2}\Big) - L\Big(2\theta\frac{g_ig_j}{\sqrt{N}c}\Big)\Big| \le \frac{4L\theta^2}{N}\sum_{i,j} g_i^2 g_j^2 \left|\frac{1}{(\hat{c}_1^N+\hat{c}_2^N+\hat{c}_3^N)^2} - \frac{1}{c^2}\right|$$
$$\le CNL\theta^2(c+\hat{c}_1^N+\hat{c}_2^N+\hat{c}_3^N)\frac{\delta}{c^2} \le C'NL\theta^2\delta K,$$

where C, C' are numerical constants and we used $K^{-1} < c < K$, and $4\delta < K^{-1}$. We get a similar estimate for the diagonal terms. The estimates on $|J_3|$ and $|J_2|$ are straightforward consequences of Tchebychev's inequality.

□

We next fix the positions of the set of indices J_1, J_2, J_3. Using the invariance under permutation of the coordinates of the Gaussian measure, we can write

$$F^N_{c_1,c_2,c_3}(\theta,\delta) = \sum_{\substack{0\le k\le 2\sqrt{N\delta}\\ 0\le l\le 2\sqrt{NK}}} \binom{N}{k}\binom{N-k}{l} F^{k,l}_{c_1,c_2,c_3},$$

where

$$F^{k,l}_{c_1,c_2,c_3} = \mathbb{E}\Big[\exp(\Sigma_c(g))\mathbb{1}_{\mathcal{A}^\delta_{c_1,c_2,c_3}\cap\mathcal{I}_{k,l}}\Big]$$

and

$$\mathcal{I}_{k,l} = \big\{J_3 = \{1,\dots,k\}, J_2 = \{k+1,\dots,k+l\}, J_1 = \{k+l+1,\dots,N\}\big\}.$$

As the number of all the possible configurations of J_2 and J_3 are sub-exponential in N by Lemma 3.2, that is, for any $k \le 2\sqrt{N}/K$ and $l \le 2\sqrt{N}/\delta$,

$$\max\left(\binom{N}{k}, \binom{N-k}{l}\right) = e^{O(\frac{\sqrt{N}}{\delta}\log N)},$$

we are reduced to compute $F^{k,l}_{c_1,c_2,c_3}$ for fixed k, l. More precisely, we obtain the following result.

Lemma 3.3.

$$\log F^N_{c_1,c_2,c_3}(\theta,\delta) = \max_{\substack{k\le 2\sqrt{N}/K\\ l\le 2\sqrt{N}/\delta}} \log F^{k,l}_{c_1,c_2,c_3} + O\Big(\frac{\sqrt{N}}{\delta}\log N\Big),$$

To simplify the notations, we denote for any $a, b \in \{1,2,3\}$,

$$\forall x, y \in \mathbb{R}^N,\ \Sigma_{a,b}(x,y) = \frac{1}{2N}\sum_{i\in J_a, j\in J_b} L\left(2\theta\frac{x_iy_j}{\sqrt{N}c}\right),$$

if $a \neq b$, and

$$\forall x, y \in \mathbb{R}^N,\ \Sigma_{a,a}(x,y) = \frac{1}{N}\sum_{i\neq j\in J_a} L\left(2\theta\frac{x_iy_j}{\sqrt{N}c}\right) + \frac{1}{N}\sum_{i\in J_a} L\Big(\sqrt{2}\theta\frac{x_iy_i}{\sqrt{N}c}\Big),$$

where now $J_3 = \{1,\dots,k\}, J_2 = \{k+1,\dots,k+l\}, J_1 = \{k+l+1,\dots,N\}$.

Next, we single out the interaction terms which involve either the quadratic behavior of L at 0 or at $+\infty$.

Lemma 3.4. *On the event* $\mathcal{A}^{\delta}_{c_1,c_2,c_3}$,

$$\Sigma_{1,1}(g,g)+2\Sigma_{1,2}(g,g)+\Sigma_{3,3}(g,g)=\frac{\theta^2}{c^2}\big(c_1^2+2c_1c_2+Bc_3^2\big)+o_{\delta K}(1)+o_K(1),$$

as $\delta K\to 0$ and $K\to+\infty$.

Proof. Observe that for $i\in J_1$, $j\in J_1\cup J_2$, $|g_ig_j|\le\sqrt{NK\delta}c$. Since $L(x)\sim_0 x^2/2$, we get,

$$\Sigma_{1,1}(g,g)+2\Sigma_{1,2}(g,g)=\theta^2\frac{\hat c_1^2}{c^2}+2\theta^2\frac{\hat c_1\hat c_2}{c^2}+o_{\delta K}(1), \tag{16}$$

as $\delta K\to 0$. Since $g\in\mathcal{A}^{\delta}_{c_1,c_2,c_3}$, we have $|\hat c_i^2-c_i^2|=O(\delta c)$ for any $i\in\{1,2,3\}$. But $c\ge K^{-1}$, therefore

$$\Sigma_{1,1}(g,g)+2\Sigma_{1,2}(g,g)=\theta^2\frac{c_1^2}{c^2}+2\theta^2\frac{c_1c_2}{c^2}+o_{\delta K}(1).$$

For $i,j\in J_3$, $|g_ig_j|\ge K\sqrt{cN}$. Since $L(x)\sim_{+\infty}\frac{B}{2}x^2$, we deduce similarly that

$$\Sigma_{3,3}(g,g)=\Big(\frac{B}{2}+o_\delta(1)\Big)\frac{\theta^2}{N^2c^2}\Big(\sum_{i\in J_3}g_i^2\Big)^2=B\theta^2\frac{c_3^2}{c^2}+o_K(1), \tag{17}$$

as $K\to+\infty$, which gives the claim.

□

From the Lemmas 3.2 and 3.4, we have on the event $\mathcal{A}^{\delta}_{c_1,c_2,c_3}$,

$$\Sigma(g)=\frac{\theta^2}{c^2}\big(c_1^2+2c_1c_2+Bc_3^2\big)+\Sigma_{1,3}(g,g)+2\Sigma_{2,3}(g,g)+\Sigma_{2,2}(g,g)+o_{\delta K}(1)+o_K(1). \tag{18}$$

Lemma 3.5. *Let* $k,l\in\mathbb{N}$ *such that* $k+l\le N$.

$$\begin{aligned}
&\log\mathbb{E}\Big(\exp\Big\{N(\Sigma_{1,3}(g,g)+2\Sigma_{2,3}(g,g)+\Sigma_{2,2}(g,g))\Big\}\mathbb{1}_{\mathcal{A}^{\delta}_{c_1,c_2c_3}\cap\mathcal{I}_{k,l}}\Big)\\
&=-\frac{N}{2}(c_2+c_3)+o_\delta(1)+O\Big(\frac{\sqrt{N}}{\delta^2}\log(1/\delta)\Big)\\
&+\max_{|\sum_i t_i^2-c_2N|\le\delta N}\ \max_{|\sum_i s_i^2-c_3N|\le\delta N}\ \log\mathbb{E}\Big(\exp\Big\{N(\Sigma_{1,3}(g,s)+2\Sigma_{2,3}(t,s)+\Sigma_{2,2}(t,t))\Big\}\mathbb{1}_{\mathcal{A}^{\delta}_{c_1}}\Big),
\end{aligned}$$

where

$$\mathcal{A}^{\delta}_{c_1}=\Big\{|\sum_{i=1}^{N-k-l}g_i^2-Nc_1|\le N\delta,\ g_i\in I_1,\ \forall i\in\{1,\dots,N-k-l\}\Big\}.$$

Proof. Integrating on $g_i, i \le k+l$, we find

$$\begin{aligned}
&\mathbb{E}\Big(\exp\Big\{ N(\Sigma_{1,3}(g,g) + 2\Sigma_{2,3}(g,g) + \Sigma_{2,2}(g,g))\Big\} \mathbb{1}_{\mathcal{A}^\delta_{c_1,c_2c_3}\cap \mathcal{I}_{k,l}}\Big) \\
&\le \frac{1}{(2\pi)^{\frac{k+l}{2}}} \int \mathbb{1}_{\{|\frac{1}{N}\sum_{i=1}^{k+l} g_i^2 - (c_2+c_3)|\le 2\delta\}} e^{-\frac12 \sum_{i=1}^{k+l} g_i^2} \prod_{i=1}^{k+l} dg_i \\
&\times \max_{\substack{t\in I_2^l \\ |\sum_i t_i^2 - c_2N|\le \delta N}} \ \max_{\substack{s\in I_3^k \\ |\sum_i s_i^2 - c_3N|\le \delta N}} \mathbb{E}\Big(\exp\Big\{ N(\Sigma_{1,3}(g,s) + 2\Sigma_{2,3}(t,s) + \Sigma_{2,2}(t,t))\Big\} \mathbb{1}_{\mathcal{A}^\delta_{c_1}}\Big),
\end{aligned}$$

But,

$$\begin{aligned}
&\frac{1}{(2\pi)^{\frac{k+l}{2}}} \int \mathbb{1}_{\{|\frac{1}{N}\sum_{i=1}^{k+l} g_i^2 - (c_2+c_3)|\le 2\delta\}} e^{-\frac12 \sum_{i=1}^{k+l} g_i^2} \prod_{i=1}^{k+l} dg_i \\
&\qquad \le \frac{1}{(2\pi)^{\frac{k+l}{2}}} e^{-\frac12 (1-\delta)(c_2+c_3-2\delta)N} \int e^{-\frac{\delta}{2}\sum_{i=1}^{k+l} g_i^2} \prod_{i=1}^{k+l} dg_i \\
&\qquad \le e^{-\frac{N}{2}(c_2+c_3) + (O(\delta)+O(\delta K))N} \big(\sqrt{2\pi}\delta^{-1}\big)^{\frac{k+l}{2}},
\end{aligned}$$

where we used the fact that $c \le K$. By Lemma 3.2 we have $k+l = O(\sqrt{N}/\delta)$.Therefore,

$$\begin{aligned}
&\log \mathbb{E}\Big(\exp\Big\{ N(\Sigma_{1,3}(g,g) + 2\Sigma_{2,3}(g,g) + \Sigma_{2,2}(g,g))\Big\} \mathbb{1}_{\mathcal{A}^\delta_{c_1,c_2,c_3}\cap \mathcal{I}_{k,l}}\Big) \\
&\le -\frac{N}{2}(c_2+c_3) + O(\delta)N + O(\delta K)N + O\Big(\frac{\sqrt{N}}{\delta}\log(1/\delta)\Big) \\
&+ \max_{\substack{t_i\in I_2 \\ |\sum_i t_i^2 - c_2N|\le \delta N}} \ \max_{\substack{s_i\in I_3 \\ |\sum_i s_i^2 - c_3N|\le \delta N}} \log \mathbb{E}\Big(\exp\Big\{ N(\Sigma_{1,3}(g,s) + 2\Sigma_{2,3}(t,s) + \Sigma_{2,2}(t,t))\Big\} \mathbb{1}_{\mathcal{A}^\delta_{c_1}}\Big).
\end{aligned}$$

To get the converse bound, we take $t \in I_2^l, s \in I_3^k$ optimizing the above maximum where δ is replaced by $\delta/2$. We next localize the integral on the set $\mathcal{B}_\delta$ where $|g_i - t_i| \le \delta/2, N-k-l \le i \le N-k, |g_i - s_i| \le \delta/2, N-k+1 \le i \le N$. Observe that on $\mathcal{A}_{c_1} \times \mathcal{B}_\delta \subset \mathcal{A}^\delta_{c_1,c_2c_3} \cap \mathcal{I}_{k,l}$, because $L \circ \sqrt{.}$ is Lipschitz,

$$|\Sigma_{1,3}(g,s) + 2\Sigma_{2,3}(t,s) + \Sigma_{2,2}(t,t) - (\Sigma_{1,3}(g,g) + 2\Sigma_{2,3}(g,g) + \Sigma_{2,2}(g,g))| \le C\theta^2\delta$$

Hence

$$\begin{aligned}
&\mathbb{E}\Big(\exp\Big\{ N(\Sigma_{1,3}(g,g) + 2\Sigma_{2,3}(g,g) + \Sigma_{2,2}(g,g))\Big\} \mathbb{1}_{\mathcal{A}^\delta_{c_1,c_2c_3}\cap \mathcal{I}_{k,l}}\Big) \\
&\ge \frac{1}{(2\pi)^{\frac{k+l}{2}}} \prod_{1\le i\le k} \int_{-\delta/2}^{\delta/2} e^{-\frac12 (g-s_i)^2} dg \prod_{1\le i\le l} \int_{-\delta/2}^{\delta/2} e^{-\frac12 (g-t_i)^2} dg \\
&\times e^{-CN\theta^2\delta} \mathbb{E}\Big(\exp\Big\{ N(\Sigma_{1,3}(g,s) + 2\Sigma_{2,3}(t,s) + \Sigma_{2,2}(t,t))\Big\} \mathbb{1}_{\mathcal{A}^\delta_{c_1}}\Big), \\
&\ge e^{-\frac{N}{2}(c_2+c_3) + O(\delta)N} \big(\frac{\delta}{(2\pi)}\big)^{\frac{k+l}{2}} \mathbb{E}\Big(\exp\Big\{ N(\Sigma_{1,3}(g,s) + 2\Sigma_{2,3}(t,s) + \Sigma_{2,2}(t,t))\Big\} \mathbb{1}_{\mathcal{A}^\delta_{c_1}}\Big)
\end{aligned}$$

which completes the claim as $k+l = O(\sqrt{N}/\delta)$.

□

Hence, we are left to estimate

$$\Lambda_1^N = \mathbb{E}\big[\exp\big(N\Sigma_{1,3}(g,s)\big)\mathbb{1}_{\mathcal{A}_{c_1}^\delta}\big],$$

where $s \in I_3^k$ such that $|\sum_i s_i^2 \leq \delta N - c_3 N|$. Let $\hat{\mu}_N = \frac{1}{|J_1|}\sum_{i\in J_1}\delta_{g_i}$. We can write

$$\Sigma_{1,3}(g,s) = \frac{|J_1|}{N}\sum_{i=1}^k \int L\left(\frac{2\theta x s_i}{\sqrt{N}c}\right) d\hat{\mu}_N(x).$$

The first difficulty in estimating Λ_1^N lies in the fact that the function

$$x \mapsto \sum_{i=1}^k L\Big(\frac{2\theta x s_i}{\sqrt{N}c}\Big),$$

is not bounded so that Varadhan's lemma (see [9, Theorem 4.3.1]) cannot be applied directly. The second issue is that we need a large deviation estimate which is uniform in the choice of $s \in I_3^k$ such that $|\sum_{i=1}^k s_i^2 - Nc_3| \leq \delta N$. In the next lemma, we prove a uniform large deviation estimate of the type of Varadhan's lemma. The proof is postponed in the appendix 7.2.

Lemma 3.6. *Let $f : \mathbb{R} \to \mathbb{R}$ such that $f(0) = 0$ and $f(\sqrt{\cdot})$ is a L-Lipschitz function. Let M_N, m_N be sequences such that $M_N = o(\sqrt{N})$ and $m_N \sim N$. Let $g_1, \dots, g_{m_N}$ be independent Gaussian random variables conditioned to belong to $[-M_N, M_N]$. Let $\delta \in (0,1)$ and $c > 0$ such that $K^{-1} < c < K$ and $2\delta < K^{-1}$. Then,*

$$\Big|\frac{1}{N}\log \mathbb{E}e^{\sum_{i=1}^{m_N} f\left(\frac{g_i}{\sqrt{c}}\right)}\mathbb{1}_{|\sum_{i=1}^{m_N} g_i^2 - cN| \leq \delta N} - \sup_{\substack{\nu\in\mathcal{P}([-M_N,M_N])\\ \int x^2 d\nu = c}} \Big\{\int f\Big(\frac{x}{\sqrt{c}}\Big)d\nu(x) - H(\nu|\gamma)\Big\}\Big|$$
$$\leq g_{L,K}(N) + h_L(\delta K),$$

where $g_{L,K}(N) \to +\infty$ as $N \to +\infty$ and $h_L(x) \to 0$ as $x \to 0$.

Let $s \in I_3^l$ such that $|\sum_i s_i^2 - Nc_3| \leq \delta N$. We consider the function

$$f : x \mapsto \sum_{i=1}^k L\Big(\frac{2\theta x s_i\sqrt{c_1}}{\sqrt{N}c}\Big).$$

One can observe that $f(\sqrt{\cdot})$ is $4\theta^2 L$-Lipschitz. Using the fact that

$$\frac{1}{N}\log \mathbb{P}(g_i \in I_1, \forall i \leq N-k-l) = o_N(1),$$

and the previous lemma, we deduce that for any $c_1 \geq K^{-1}$,

$$\Big|\frac{1}{N}\log\Lambda_1^N - \sup_{\substack{\nu\in\mathcal{P}(I_1)\\ \int x^2 d\nu(x) = c_1}} \Big\{\sum_{i=1}^k \int L\Big(\frac{2\theta x s_i}{\sqrt{N}c}\Big)d\nu(x) - H(\nu|\gamma)\Big\}\Big| \leq g_K(N) + h_L(\delta K), \quad (19)$$

where $g_K(N) \to 0$ as $N \to +\infty$. Putting together (18) and (19), we obtain

$$\Big|\frac{1}{N}\log\mathbb{E}\Big(\exp(\Sigma(g))\mathbb{1}_{\mathcal{A}_{c_1,c_2,c_3}\cap\mathcal{I}_{k,l}}\Big) - \Psi_{k,l}(c_1,c_2,c_3)\Big| \leq g_{L,K}(N) + o_{\delta K}(1) + o_K(1),$$

where

$$\Psi_{k,l}(c_1,c_2,c_3) = Q(c_1,c_2,c_3)\Big) - \frac{1}{2}\big(c_2+c_3\big) + \max_{\substack{t_i\in I_2 \\ |\sum_i t_i^2 - c_2N|\le \delta N}} \max_{\substack{s_i\in I_3 \\ |\sum_i s_i^2 - c_3N|\le \delta N}} \Big\{\Sigma_{2,3}(t,s) + \Sigma_{2,2}(t,t) + \sup_{\substack{\nu\in\mathcal{P}(I_1) \\ \int x^2 d\nu_1(x) = c_1}} \{\Phi(\nu,s) - H(\nu|\gamma_1)\}\Big\},$$

with

$$Q(x,y,z) = \theta^2 \frac{x^2+2xy+Bz^2}{(x+y+z)^2},$$

and

$$\Phi(\nu,s) = \sum_{i=1}^k \int L\Big(\frac{2\theta x s_i}{\sqrt{N}c}\Big) d\nu(x).$$

By Lemma 3.3, we obtain as $\delta < K^{-1}$,

$$\Big|\frac{1}{N}\log F^N_{c_1,c_2,c_3}(\theta,\delta) - \max_{\substack{k\le 4\sqrt{N}/K \\ l\le 4\sqrt{N}/\delta}} \Psi_{k,l}(c_1,c_2,c_3)\Big| \le \tilde{g}_K(N) + o_{\delta K}(1) + o_K(1), \tag{20}$$

where $\tilde{g}_K(N) \to 0$ as $N \to +\infty$.

Let $c_1, c_2, c_3 \ge 0$ such that $c_1 > 0$. Let $c = c_1 + c_2 + c_3$. There exists $K > 0$ such that $K^{-1} \le c_1 \le c \le K$. We have for any $0 < 4\delta < K^{-1}$,

$$\liminf_{N\to+\infty} F_N(\theta) \ge \liminf_{N\to+\infty} \frac{1}{N}\log F^N_{c_1,c_2,c_3}(\theta,\delta).$$

Therefore,

$$\liminf_{N\to+\infty} F_N(\theta) \ge \liminf_{N\to+\infty} \max_{\substack{k\le 2\sqrt{N}/K \\ l\le 2\sqrt{N}/\delta}} \Psi_{k,l}(c_1,c_2,c_3) - o_{\delta K}(1) - o_K(1).$$

To complete the proof of the lower bound of Lemma 3.5, one can observe that

$$\Sigma_{2,2}(t,t) = \frac{1}{N}\sum_{i,j} L\Big(2\theta\frac{t_i t_j}{\sqrt{N}c}\Big) + o_N(1),$$

uniformly in $t_i \in I_2$ such that $|\sum_i t_i^2 - c_2 N| \le \delta N$. Indeed, the diagonal terms are negligible in this case since

$$\frac{1}{N}\sum_{t_i\in J_2} \frac{t_i^4}{Nc^2} \le c\delta^{-1}N^{-1/2}\frac{1}{N}\sum_{i\in J_2} t_i^2 = O(N^{-1/2}).$$

To conclude the proof of the upper bound, we will use the exponential tightness of $||g||^2$ and of $||g||_{J_1}^{-2}$. More precisely, we claim that

$$\lim_{K\to+\infty} \limsup_{N\to+\infty} \frac{1}{N}\mathbb{P}(||g||^2 \ge KN, ||g||_{J_1}^2 \le K^{-1}N) = -\infty. \tag{21}$$

Indeed, it is clear by Chernoff's inequality that

$$\mathbb{P}(||g||^2 \ge KN) \le Ce^{-CNK^2},$$

where C is a positive numerical constant. Whereas, using Lemma 3.2 and a union bound,

$$\mathbb{P}(||g||_{J_1}^2 \leq K^{-1}N) \leq \sum_{m \leq 2\sqrt{N}/\delta} \binom{N}{m} \mathbb{P}\big(\sum_{i=1}^{N-m} g_i^2 \leq K^{-1}N\big).$$

Let $\Lambda(\zeta) = \log \mathbb{E}e^{\theta g_1^2}$ for any $\zeta \in \mathbb{R}$. By Chernoff's inequality, we have for any $m \leq N$,

$$\mathbb{P}\big(\sum_{i=1}^{N-m} g_i^2 \leq K^{-1}N\big) \leq e^{-N\Lambda^*(K^{-1})},$$

where Λ^* is the Legendre transform of Λ. Since 0 is not in the support of the law g_1^2, we have $\Lambda^*(0) = +\infty$. Since Λ^* is lower semi-continuous we have $\Lambda^*(K^{-1}) \to +\infty$ as $K \downarrow +\infty$. Using the fact that for any $m \leq 2\sqrt{N}/\delta$,

$$\binom{N}{m} \leq e^{o(N)},$$

we get the claim (21). Using that $L(x) \leq Ax^2/2$, we have

$$\Sigma(g) \leq A\theta^2 N.$$

From the exponential tightness (21), we deduce that there exists $\delta > 0$ such that,

$$\mathbb{E}\exp(\Sigma(g))\mathbb{1}_{\{\frac{1}{N}||g||^2 \geq K, \frac{1}{N}||g||_{J_1}^2 \leq K^{-1}\}} \leq 1. \tag{22}$$

Let $\mathcal{E}_K = \{\frac{1}{N}||g||^2 \geq K, \frac{1}{N}||g||_{J_1}^2 \leq K^{-1}\}$. We have

$$0 \leq \limsup_{N\to+\infty} F_N(\theta) \leq \max\Big(\limsup_{N\to+\infty} \frac{1}{N}\log \mathbb{E}\big(e^{\Sigma(g)}\mathbb{1}_{\mathcal{E}_K}\big), \limsup_{N\to+\infty} \frac{1}{N}\log \mathbb{E}\big(e^{\Sigma(g)}\mathbb{1}_{\mathcal{E}_K^c}\big)\Big)$$

Since we took K so that (22) holds, we have

$$\limsup_{N\to+\infty} F_N(\theta) \leq \limsup_{N\to+\infty} \frac{1}{N}\log \mathbb{E}\big(e^{\Sigma(g)}\mathbb{1}_{\mathcal{E}_K}\big). \tag{23}$$

Let now $\mathcal{C}_\delta$ be a δ-net for the ℓ^∞-norm of the set

$$\{(c_1, c_2, c_3) \in \mathbb{R}_+^3 : c_1 + c_2 + c_3 \leq K, c_1 \geq K^{-1}\}.$$

As $|\mathcal{C}_\delta| = O(K/\delta)$, we have

$$\limsup_{N\to+\infty} F_N(\theta) \leq \limsup_{N\to+\infty} \max_{(c_1,c_2,c_3)\in\mathcal{C}_\delta} \frac{1}{N}\log F_{c_1,c_2,c_3}^N(\theta,\delta).$$

$$\limsup_{N\to+\infty} F_N(\theta) \leq \max_{(c_1,c_2,c_3)\in\mathcal{C}_\delta} \limsup_{N\to+\infty} \frac{1}{N}\log F_{c_1,c_2,c_3}^N(\theta,\delta).$$

From (20), we deduce

$$\limsup_{N\to+\infty} F_N(\theta) \leq \max_{\substack{c=c_1+c_2+c_3 \\ c_1\geq K^{-1}}} \limsup_{N\to+\infty} \max_{\substack{k\leq 2\sqrt{N}/K \\ l\leq 2\sqrt{N}/\delta}} \Psi_{k,l}(c_1,c_2,c_3) + o_{\delta K}(1) + o_K(1).$$

Taking now the limit as $\delta \to 0$ and $K \to +\infty$ such that $\delta K \to 0$, we obtain the upper bound of Lemma 3.1. □

We are now ready to give a proof of Proposition 1.6 . Building on Lemma 3.1, we show that we can optimize on the total norm c in order to simplify the variational problem, and replace the supremum over discrete measures $\frac{1}{N}\sum_i \delta_{t_i}$ by a supremum over all measures with a certain constraint of support and second moment.

Proof of Proposition 1.6 . We use the notation of Lemma 3.1. We make the following changes of variables:

$$\forall i = 1,2,3,\ \alpha_i = \frac{c_i}{c},\ c = c_1 + c_2 + c_3,$$

$$\nu = \nu_1 \circ (x \mapsto x/\sqrt{c})^{-1}.$$

For any $\nu_1 \in \mathcal{P}(I_1)$ such that $\int x^2 d\nu_1(x) = c_1$ we have,

$$H(\nu_1|\gamma) = H(\nu_1) + \frac{1}{2}c_1 + \frac{1}{2}\log(2\pi).$$

Moreover, the density of ν and ν_1 are linked by the relation

$$\forall x \in \mathbb{R},\ \frac{d\nu}{dx}(x) = \sqrt{c}\frac{d\nu_1}{dx}(\sqrt{c}x).$$

Therefore,

$$H(\nu) = \int \log \frac{d\nu_1}{dx} d\nu_1 = H(\nu_1) + \frac{1}{2}\log c. \tag{24}$$

We obtain

$$\overline{F}(\theta) = \limsup_{\substack{\delta\to 0, K\to+\infty \\ \delta K \to 0}} \sup_{\substack{\alpha_1+\alpha_2+\alpha_3=1 \\ \alpha_i \geq 0}} \sup_{c\geq 0} \limsup_{N\to+\infty} \mathcal{F}^N_{\alpha_1,\alpha_2,\alpha_3,c}(\delta, K),$$

and similarly for $\underline{F}(\theta)$, where

$$\mathcal{F}^N_{\alpha_1,\alpha_2,\alpha_3,c}(\delta, K) = \sup_{\substack{s_i\in I_3 \\ |\sum s_i^2 - N\alpha_3| \leq N\delta}} \sup_{\substack{t_i \in I_2 \\ |\sum t_i^2 - N\alpha_2| \leq N\delta}} \Big\{ \theta^2\big(\alpha_1^2 + 2\alpha_1\alpha_2 + B\alpha_3^2\big) - \frac{1}{2}c + \frac{1}{2}\log c$$
$$+ \frac{1}{N}\sum_{i,j} L\Big(\frac{2\theta s_i t_j}{\sqrt{N}}\Big) + \frac{1}{2N}\sum_{i,j} L\Big(\frac{2\theta t_i t_j}{\sqrt{N}}\Big) + \sup_{\substack{\nu\in\mathcal{P}(I_1) \\ \int x^2 d\nu_1(x) = \alpha_1}} \big\{\Phi(\nu, s) - H(\nu)\big\}\Big\} - \frac{1}{2}\log(2\pi),$$

and

$$\Phi(\nu, s) = \sum_{i=1}^k \int L\Big(\frac{2\theta x s_i}{\sqrt{N}}\Big) d\nu(x),$$

with $I_1 = [-\sqrt{\delta}N^{1/4}, \sqrt{\delta}N^{1/4}]$, $I_2 = [\sqrt{\delta}N^{1/4}, \sqrt{K}N^{1/4}]$ and $I_3 = [\sqrt{K}N^{1/4}, \sqrt{\alpha_3 N}]$. We see that we can optimize in c, and find that the maximum is achieved at $c = 1$ by concavity of the log.

□

4. The large deviations close to the bulk

We prove in this section Proposition 1.15. By Theorem 1.12, the large deviation lower bound holds at every $x > 2$ such that $I(x) = \bar{I}(x) \neq 0$ so that there exists $\theta \in \Theta_x$ which does not belong to any Θ_y for $y \neq x$. In the following lemma, we prove that if $\overline{F}(\theta) = \underline{F}(\theta) = \theta^2$ on a interval $(0, b)$ with $b > 1/2$, then the large deviation lower bound holds in a neighborhood of 2 and the rate function I is equal to the one of the GOE.

Lemma 4.1. *If for $\theta \in \left(0, \frac{1}{2\varepsilon}\right)$, for some $\varepsilon \in (0,1)$, $\overline{F}(\theta) = F(\theta) = \theta^2$, then for any $x \in [2, \varepsilon + \frac{1}{\varepsilon})$,*

$$\bar{I}(x) = I(x) = I_{GOE}(x).$$

As a consequence, $\forall x > 2, \bar{I}(x) > 0$. Moreover, for $x \in [2, \varepsilon + \frac{1}{\varepsilon})$ the optimizer in I is taken at $\theta_x = 1/2G(x)$ and $\theta_x \notin \Theta_y$ for all $y \neq x$.

Proof. As $\underline{F}(\theta) \geq \theta^2$ for any $\theta \geq 0$, we have that

$$\sup_{\theta \in [0, 1/2\varepsilon)} \left\{ J(\theta, x) - \theta^2 \right\} \leq I(x) \leq \bar{I}(x) \leq \sup_{\theta \geq 0} \left\{ J(\theta, x) - \theta^2 \right\}.$$

But if $x \in [2, \varepsilon + \frac{1}{\varepsilon})$,

$$I_{GOE}(x) = \sup_{\theta \geq 0} \left\{ J(\theta, x) - \theta^2 \right\},$$

is achieved at $\theta = 1/2G(x) \in (0, 1/2\varepsilon)$ since $G^{-1}(\varepsilon) = \varepsilon + 1/\varepsilon$. Therefore, if $x \in [2, 2\varepsilon + \frac{1}{2\varepsilon})$, then we obtain

$$I(x) = \bar{I}(x) = I_{GOE}(x).$$

In particular, $I(x) > 0$ for any $x \in (2, 2\varepsilon + 1/2\varepsilon)$. As I is convex by Proposition 1.11 and $I(2) = 0$, we conclude that $I(x) > 0$ for any $x > 2$. Moreover, the optimizer in $I(x)$ is taken at $\theta_x = 1/2G(x)$ for which $\underline{F}(\theta) = \overline{F}(\theta)$ is differentiable and clearly $\theta_x \notin \Theta_y$ for any $y \neq x$ as G is invertible on $[2, +\infty)$.

□

The result of Proposition 1.15 then follows from Lemma 4.1 and Lemma 1.8. We now study the convergence of the annealed spherical integrals for large values of θ, in which case we need to make additional assumptions on μ.

5. Case where ψ is an increasing function

In this section we make the additional assumption that ψ is non-decreasing.

Example 5.1 (Sparse Gaussian distribution)**.** *Let μ be the law of $\xi\Gamma$ where ξ is a Bernoulli variable of pararmeter $p \in (0,1)$ and Γ is a standard Gaussian random variable. In that case we have for any $x \in \mathbb{R}$,*

$$\psi(x) = \frac{\log[(1-p) + p\exp(x^2/2p)]}{x^2} = \int_0^1 \frac{t\exp((xt)^2/2p)}{(1-p) + p\exp((xt)^2/2p)} dt.$$

One can observe that this last expression is indeed increasing in x.

5.1. **Simplification of the variational problem.** We prove in this section that when ψ is increasing on $\mathbb{R}^+$, $\mathcal{C}_\mu = \mathbb{R}^+$ and we can simplify the limit $F(\theta)$ as follows.

Proposition 5.2. *For any $\theta \geq 0$, $\overline{F}(\theta) = \underline{F}(\theta) = F(\theta)$ where*

$$F(\theta) = \sup_{\alpha\in[0,1]} \sup_{\substack{\nu\in\mathcal{P}(\mathbb{R}) \\ \int x^2 d\nu(x)=\alpha}} \Big\{\theta^2\alpha^2 + B\theta^2(1-\alpha)^2 + \int L(2\theta\sqrt{1-\alpha}x)d\nu(x) - H(\nu) - \frac{1}{2}\log(2\pi) - \frac{1}{2}\Big\},$$

Proof. Recall that we set for any $\delta, K > 0$, and $\alpha_1 + \alpha_2 + \alpha_3 = 1$,

$$\begin{aligned}\mathcal{F}^N_{\alpha_1,\alpha_2,\alpha_3}(\delta,K) = &\sup_{\substack{s\in I_2^l, k\geq 1 \\ |\sum t_i^2 - N\alpha_2|\leq \delta N}} \sup_{\substack{s\in I_3^k, k\geq 1 \\ |\sum s_i^2 - N\alpha_3|\leq \delta N}} \Big\{\theta^2\big(\alpha_1^2 + 2\alpha_1\alpha_2 + B\alpha_3^2\big) \\ &+ \frac{1}{N}\sum_{j=1}^{l}\sum_{i=1}^{k} L\Big(\frac{2\theta s_i t_j}{\sqrt{N}}\Big) + \frac{1}{2N}\sum_{i,j=1}^{l} L\Big(\frac{2\theta t_i t_j}{\sqrt{N}}\Big) \\ &+ \sup_{\substack{\nu_1\in\mathcal{P}(I_1) \\ \int x^2 d\nu_1(x)=\alpha_1}} \Big\{\sum_{i=1}^{k}\int L\Big(\frac{2\theta s_i x}{\sqrt{N}}\Big)d\nu_1(x) - H(\nu_1)\Big\} - \frac{1}{2}\log(2\pi) - \frac{1}{2}\Big\},\end{aligned}$$

where $I_1 = \{x : |x| \leq \sqrt{\delta}N^{1/4}\}$, $I_2 = \{x : \sqrt{\delta}N^{1/4} \leq |x| \leq \sqrt{K}N^{1/4}\}$, $I_3 = \{\sqrt{K}N^{1/4} \leq |x| \leq \sqrt{N\alpha_3}\}$. Since ψ is non-decreasing, we have on one hand for any $s \in I_3^k$ such that $|\sum_i s_i^2 - \alpha_3 N| \leq N\delta$, $\psi\Big(\frac{2\theta s_i x}{\sqrt{N}}\Big) \leq \psi\Big(2\theta\sqrt{(\alpha_3+\delta)}x\Big)$ so that

$$\begin{aligned}\sum_{i=1}^{k}\int L\Big(\frac{2\theta s_i x}{\sqrt{N}}\Big)d\nu_1(x) &= \frac{4\theta^2}{N}\sum_{i=1}^{k} s_i^2 \int x^2\psi\Big(\frac{2\theta s_i x}{\sqrt{N}}\Big)d\nu_1(x) \\ &\leq \int L\Big(2\theta\sqrt{\alpha_3+\delta}x\Big)d\nu_1(x) = \int L\Big(2\theta\sqrt{\alpha_3}x\Big)d\nu_1(x) + o_\delta(1)\end{aligned}$$

where we finally use that $L\circ\sqrt{\cdot}$ is Lipschitz. On the other hand, since $L(x) \leq Bx^2/2$ for any $x \geq 0$,

$$\frac{1}{N}\sum_{j=1}^{l}\sum_{i=1}^{k} L\Big(\frac{2\theta s_i t_j}{\sqrt{N}}\Big)d\nu_2(x) + \frac{1}{2N}\sum_{i,j=1}^{l} L\Big(\frac{2\theta t_i t_j}{\sqrt{N}}\Big) \leq \theta^2 B(\alpha_2^2 + 2\alpha_3\alpha_2) + o_\delta(1).$$

Therefore, we have the upper bound,

$$\begin{aligned}\overline{F}(\theta) \leq \sup_{\alpha_1+\alpha_2+\alpha_3=1} \sup_{\substack{\nu_1\in\mathcal{P}(\mathbb{R}) \\ \int x^2 d\nu_1(x)=\alpha_1}} &\Big\{\theta^2\big(\alpha_1^2 + 2\alpha_1\alpha_2\big) + \theta^2 B(\alpha_3+\alpha_2)^2 \\ &+ \int L(2\theta\sqrt{\alpha_3}x)d\nu_1(x) - H(\nu_1) - \frac{1}{2}\log(2\pi) - \frac{1}{2}\Big\}.\end{aligned}$$

We can further simplify this optimization problem by showing that the assumption on the monotonicity of ψ entails that we can take $\alpha_2 = 0$. Indeed, note that $\psi(0) = 1/2$.

Therefore, for any $x \geq 0$, $\psi(x) \geq 1/2$. Hence, we deduce that

$$2\theta^2\alpha_1\alpha_2 + \int L(2\theta\sqrt{\alpha_3}x)d\nu_1(x) = 2\theta^2\alpha_1\alpha_2 + 4\theta^2\alpha_3 \int x^2\psi(2\theta\sqrt{\alpha_3}x)d\nu_1(x)$$
$$\leq 4\theta^2(\alpha_2+\alpha_3)\int x^2\psi(2\theta\sqrt{\alpha_2+\alpha_3}x)d\nu_1(x).$$

Thus, with the change of variables $\alpha_3+\alpha_2 \to \alpha_3$, we deduce

$$\overline{F}(\theta) \leq \sup_{\alpha_1+\alpha_3=1} \sup_{\substack{\nu_1\in\mathcal{P}(\mathbb{R}) \\ \int x^2 d\nu_1(x)=\alpha_1}} \Big\{\theta^2\alpha_1^2 + \theta^2 B\alpha_3^2 + \int L(2\theta\sqrt{\alpha_3}x)d\nu_1(x) - H(\nu_1) - \frac{1}{2}\log(2\pi) - \frac{1}{2}\Big\}.$$

To prove that $\underline{F}(\theta)$ is bounded from below by the same quantity, we fix $\alpha_1, \alpha_2, \alpha_3$ such that $\alpha_1+\alpha_2+\alpha_3 = 1, \alpha_2 = 0$, and $\nu \in \mathcal{P}(\mathbb{R})$ such that $\int x^2 d\nu(x) = \alpha_1$. We take in the definition of $\mathcal{F}^N_{\alpha_1,\alpha_2,\alpha_3}(K,\delta)$, $k=1$, $s=\sqrt{\alpha_3 N}$, and $\nu_1 \in \mathcal{P}(I_1)$ defined by

$$\nu_1 = \frac{1}{\nu(I_1)}\nu(.\cap I_1)\circ h^{-1}_{\sqrt{\lambda}},$$

where $h_{\sqrt{\lambda}} : x \mapsto \sqrt{\lambda}x$, and λ is such that $\int x^2 d\nu_1(x) = \alpha_1$, that is,

$$\frac{1}{\lambda} = \frac{1}{\nu(I_1)\alpha_1}\int_{I_1} x^2 d\nu(x).$$

We have,

$$\mathcal{F}^N_{\alpha_1,\alpha_2,\alpha_3}(K,\delta) \geq \theta^2\big(\alpha_1^2 + B\alpha_3\big)$$
$$+ \int L(2\theta\sqrt{\alpha_3}x)d\nu_1(x) - H(\nu_1) - \frac{1}{2}\log(2\pi) - \frac{1}{2}.$$

We deduce by monotone convergence and using the fact that $L(\sqrt{.})$ is Lipschitz that

$$\underline{F}(\theta) \geq \theta^2\big(\alpha_1^2 + B\alpha_3\big) + \int L(2\theta\sqrt{\alpha_3}x)d\nu(x) - H(\nu) - \frac{1}{2}\log(2\pi) - \frac{1}{2}.$$

□

We next compute the supremum over ν in the definition of F in Proposition 5.2. To this end, we denote by $G : [B/2, +\infty) \to \mathbb{R}\cup\{+\infty\}$ the function given by

$$\forall \zeta \in [B/2, +\infty),\ G(\zeta) = \log \int \exp(L(x) - \zeta x^2)dx\,. \tag{25}$$

Lemma 5.3. *Let*

$$l = -\lim_{\zeta \to B/2} G'(\zeta) \in (0, +\infty]. \tag{26}$$

For any $C \in (0, l)$, there exists a unique $\zeta \in (B/2, +\infty)$ solution to the equation

$$G'(\zeta) = -C\,.$$

It is denoted by ζ_C. For $C \geq l$, we set $\zeta_C = B/2$. Then,

$$\sup_{\substack{\nu\in\mathcal{P}(\mathbb{R})\\ \int x^2 d\nu(x)=C}} \Big[\int L(x)d\nu(x) - H(\nu)\Big] = \sup_{\substack{\nu\in\mathcal{P}(\mathbb{R})\\ \int x^2 d\nu(x)\leq C}} \Big[\frac{BC}{2} + \int \Big(L(x) - \frac{B}{2}x^2\Big)d\nu(x) - H(\nu)\Big]$$
$$= C\zeta_C + G(\zeta_C)$$

Proof. Define the function

$$\forall \nu \in \mathcal{P}(\mathbb{R}),\ E(\nu) = H(\nu) + \int \Big(\frac{B}{2}x^2 - L(x)\Big)d\nu(x).$$

We first show that

$$\inf_{\substack{\nu\in\mathcal{P}(\mathbb{R})\\ \int x^2 d\nu(x)=C}} E(\nu) = \inf_{\substack{\nu\in\mathcal{P}(\mathbb{R})\\ \int x^2 d\nu(x)\leq C}} E(\nu). \tag{27}$$

To prove this equality, we will show that for any $\nu \in \mathcal{P}(\mathbb{R})$ such that $\int x^2 d\nu(x) \leq C$, there exists ν_ε such that $\int x^2 d\nu_\varepsilon(x) = C$, and

$$\lim_{\varepsilon\to 0} \nu_\varepsilon = \nu, \quad \lim_{\varepsilon\to 0} E(\nu_\varepsilon) = E(\nu).$$

We set $\nu_\varepsilon = (1-\varepsilon)\nu + \varepsilon\gamma_\varepsilon$ where γ_ε is a Gaussian measure of variance 1 and mean m_ε, defined by,

$$m_\varepsilon = \sqrt{\frac{C - (1-\varepsilon)D}{\varepsilon} - 1}$$

With this choice of m_ε, one can check that $\int x^2 d\nu_\varepsilon(x) = C$. Moreover,

$$\nu_\varepsilon \underset{\varepsilon\to 0^+}{\longrightarrow} \nu,$$

for the weak topology. As $\int x^2 d\nu_\varepsilon(x) \leq C$, we deduce that for any continuous function $f : \mathbb{R} \to \mathbb{R}$ such that $f = o_{\pm\infty}(x^2)$,

$$\int f d\nu_\varepsilon \underset{\varepsilon\to 0^+}{\longrightarrow} \int f d\nu.$$

Since $\frac{B}{2}x^2 - L(x) = o_{\pm\infty}(x^2)$, we have

$$\lim_{\varepsilon\to 0^+} \int \Big(\frac{B}{2}x^2 - L(x)\Big)d\nu_\varepsilon(x) = \int \Big(\frac{B}{2}x^2 - L(x)\Big)d\nu(x). \tag{28}$$

Besides, as H is convex,

$$H(\nu_\varepsilon) \leq (1-\varepsilon)H(\nu) + \varepsilon H(\gamma_\varepsilon).$$

But,

$$H(\gamma_\varepsilon) = H(\gamma),$$

where γ is a standard Gaussian distribution. Therefore,

$$\limsup_{\varepsilon\to 0^+} H(\nu_\varepsilon) \leq H(\nu).$$

As H is lower semi-continuous with respect to the weak topology, we can conclude together with (28) that,

$$\lim_{\varepsilon\to 0^+} E(\nu_\varepsilon) = E(\nu),$$

which ends the proof of the claim (48).

Observe that E is a lower semi-continuous function for the weak topology since H is lower semi-continuous and $x \mapsto Bx^2/2 - L(x)$ is non-negative and continuous. Moreover, the set

$$\big\{\nu \in \mathcal{P}(\mathbb{R}) : \int x^2 d\nu(x) \leq C\big\},$$

is a compact subset. Thus, the supremum of E over the set above is achieved. We will identify the maximizer. For any $\zeta \in [B/2, +\infty)$ such that $G(\zeta) < +\infty$, we let ν_ζ be the probability measure given by

$$d\nu_\zeta = \frac{\exp(L(x) - \zeta x^2)}{\int \exp(L(y) - \zeta y^2) dy} dx.$$

We will show that

$$\inf_{\substack{\nu \in \mathcal{P}(\mathbb{R}) \\ \int x^2 d\nu(x) \leq C}} E(\nu) = E(\nu_{\zeta_C}). \tag{29}$$

Let μ be a probability measure such that $H(\mu) < +\infty$ and $\int x^2 d\mu(x) \leq C$. As $H(\mu) < +\infty$, we can write,

$$\mu = (1 + \varphi) d\nu_{\zeta_C},$$

where φ is some measurable function such that $\varphi \geq -1$ ν_{ζ_C}-a.s. and $\int \varphi d\nu_{\zeta_C} = 0$. We have,

$$\begin{aligned} E(\mu) = E(\nu_{\zeta_C}) &+ \int \Big(\frac{B}{2} x^2 - L(x)\Big) \varphi(x) d\nu_{\zeta_C}(x) \\ &+ \int (1 + \varphi) \log(1 + \varphi) d\nu_{\zeta_C} + \int \log \frac{d\nu_{\zeta_C}}{dx} \varphi d\nu_{\zeta_C}. \end{aligned}$$

By convexity of $x \mapsto x \log x$, we have

$$\int (1 + \varphi) \log(1 + \varphi) d\nu_{\zeta_C} \geq (1 + \int \varphi d\nu_{\zeta_C}) \log(1 + \int \varphi d\nu_{\zeta_C}) = 0.$$

Therefore,

$$E(\mu) \;\; \geq \;\; E(\nu_{\zeta_C}) + \int \left(\frac{B}{2} x^2 - L(x) + \log \frac{d\nu_{\zeta_C}}{dx}\right) \varphi(x) d\nu_{\zeta_C}(x) + \Big(\frac{B}{2} - \zeta_C\Big) \int x^2 \varphi(x) d\nu_{\zeta_C}(x),$$

where we used again that $\int \varphi d\nu_{\zeta_C} = 0$. But,

$$\Big(\frac{B}{2} - \zeta_C\Big) \int x^2 \varphi(x) d\nu_{\zeta_C}(x) = \Big(\frac{B}{2} - \zeta_C\Big)\Big(\int x^2 d\mu(x) - \int x^2 d\nu_{\zeta_C}(x)\Big).$$

If $C < l$, then $\int x^2 d\nu_{\zeta_C}(x) = C$. Since $\zeta_C \leq B/2$ and $\int x^2 d\mu(x) \leq C$ we get

$$E(\mu) \geq E(\nu_{\zeta_C}).$$

If $C \geq l$, then $\zeta_C = B/2$, and we also get $E(\mu) \geq E(\nu_{\zeta_C})$. This shows that ν_{ζ_C} achieves the infimum in (29), and ends the proof of Lemma 5.3. □

5.2. **Differentiability of the limit of the annealed spherical integral.** This section is devoted to the proof of the following proposition.

Proposition 5.4. *F is continuously differentiable on $(1/\sqrt{B-1}, +\infty)$ except possibly at the point θ_0 such that*

$$\theta_0 = \inf\big\{\theta : F(\theta) > \theta^2\big\}.$$

Moreover, for any $\theta \leq 1/2\sqrt{B-1}$,

$$F(\theta) = \theta^2. \tag{30}$$

The second part of the claim of the above proposition (30) is due to Proposition 1.8 and the fact that $A = B$. From now on, we assume that $\theta^2(B-1) > 1$ and wish to prove the first part of Proposition 5.4. We define for any $\alpha \in [0,1]$, and $\nu \in \mathcal{P}(\mathbb{R})$,

$$H_\theta(\alpha,\nu) = \theta^2\alpha^2 + \theta^2 B(1-\alpha)^2 + \int L(2\theta\sqrt{1-\alpha}x)d\nu(x) - H(\nu) - \frac{1}{2}\log(2\pi) - \frac{1}{2}. \tag{31}$$

By Proposition 5.2, we have

$$F(\theta) = \sup_{(\alpha,\nu)\in S} H_\theta(\alpha,\nu), \tag{32}$$

where

$$S = \{(\alpha,\nu) \in [0,1]\times\mathcal{P}(\mathbb{R}) : \int x^2 d\nu(x) = \alpha\}.$$

We first show that we can can restrict the parameter α to the set $[0,\frac{1}{2}]\cup\{1\}$, as described in the following lemma.

Lemma 5.5. *If $\theta^2(B-1) \geq 1$, then*

$$F(\theta) = \max\Big(\sup_{\substack{(\alpha,\nu)\in S\\ \alpha\leq\frac{1}{2}}} H_\theta(\alpha,\nu), \theta^2\Big).$$

Proof. Making the change of variable which consists in replacing ν by the push-forward of ν by the map $x \mapsto x/\sqrt{\alpha}$, we get

$$\sup_{(\alpha,\nu)\in S} H_\theta(\alpha,\nu) = \sup_{\substack{\alpha\in(0,1]\\ \int x^2 d\nu = 1}} \{\tilde{H}_\theta(\alpha,\nu) - H(\nu)\},$$

where for any $\alpha \in (0,1]$, and $\nu \in \mathcal{P}(\mathbb{R})$,

$$\tilde{H}_\theta(\alpha,\nu) = \theta^2\alpha^2 + \theta^2 B(1-\alpha)^2 + \frac{1}{2}\log\alpha + \int L(2\theta\sqrt{(1-\alpha)\alpha}x)d\nu(x) - \frac{1}{2}\log(2\pi) - \frac{1}{2}.$$

We claim that for any $\nu \in \mathcal{P}(\mathbb{R})$ such that $\int x^2 d\nu(x) = 1$,

$$\max_{\alpha\in[1/2,1]} \tilde{H}_\theta(\alpha,\nu) \leq \max\Big(\tilde{H}_\theta\Big(\frac{1}{2},\nu\Big), \tilde{H}_\theta(1,\nu)\Big). \tag{33}$$

Indeed, first notice that since ψ is increasing, for all $\alpha \in [0,1]$ we have,

$$\begin{aligned}\int L(2\theta\sqrt{(1-\alpha)\alpha}x)d\nu(x) &= 4\theta^2\alpha(1-\alpha)\int x^2\psi(2\theta\sqrt{\alpha(1-\alpha)}x)d\nu(x)\\ &\leq 4\theta^2\alpha(1-\alpha)\int x^2\psi(\theta x)d\nu(x)\,.\end{aligned}$$

Denote by $m = 2\int x^2\psi(\theta x)d\nu(x)$. As ψ takes its values in $[1/2, B/2]$ we have that $m \in [1, B]$. For any $\alpha \in (0, 1]$,

$$\tilde{H}_\theta(\alpha, \nu) \leq \theta^2\alpha^2 + \theta^2 B(1-\alpha)^2 + \frac{1}{2}\log\alpha + 2\theta^2\alpha(1-\alpha)m - \frac{1}{2}\log(2\pi) - \frac{1}{2} =: f_{\theta,m}(\alpha)\,.$$

We find that

$$f'_{\theta,m}\Big(\frac{1}{2}\Big) = \theta^2(1-B) + 1, \quad f''_{\theta,m}(\alpha) = 2\theta^2(B+1-2m) - \frac{1}{2\alpha^2}.$$

Since $f''_{\theta,m}$ is increasing and $f''_{\theta,m}(0) = -\infty$, we deduce that $f'_{\theta,m}$ is either decreasing or decreasing and then increasing. Since $f'_{\theta,m}(1/2) \leq 0$, we conclude that $f_{\theta,m}$ is either decreasing or decreasing and then increasing on $[1/2, 1]$. Therefore,

$$\max_{\alpha\in[\frac{1}{2},1]} f_{\theta,m}(\alpha) = \max\Big(f_{\theta,m}\Big(\frac{1}{2}\Big), f_{\theta,m}(1)\Big)\,,$$

which yields the claim (33) since $f_{\theta,m}(\alpha) = \tilde{H}_\theta(\alpha, \nu)$ at the two points $\alpha = 1/2$ and 1. To conclude the proof we observe that since $\tilde{H}_\theta(1, \nu) = \theta^2$ for any $\nu \in \mathcal{P}(\mathbb{R})$, we have

$$\sup_{\int x^2 d\nu(x)=1} \{\tilde{H}_\theta(1, \nu) - H(\nu)\} = \theta^2 + \frac{1}{2} + \frac{1}{2}\log(2\pi) - \inf_{\int x^2 d\nu(x)=1} H(\nu) = \theta^2.$$

□

Due to Lemma 5.3, we can further simplify the optimization problem defining $F(\theta)$ in (32) by optimizing on $\nu \in \mathcal{P}(\mathbb{R})$ such that $\int x^2 d\nu(x) = \alpha$, given $\alpha \in (0, 1)$.

Corollary 5.6. *Let R be the function*

$$R : C \in (0, +\infty) \mapsto C\zeta_C + G(\zeta_C),$$

where ζ_C is defined as in Lemma 5.3. Denote for any $\alpha \in (0, 1)$,

$$K_\theta(\alpha) = \theta^2(\alpha^2 + B(1-\alpha)^2) + R(4\theta^2\alpha(1-\alpha)) - \frac{1}{2}\log(1-\alpha) - \log(2\theta) - \frac{1}{2}\log(2\pi) - \frac{1}{2},$$

and $K_\theta(1) = \theta^2$. Then, for any $\theta \geq 1/\sqrt{B-1}$,

$$F(\theta) = \sup_{\alpha\in(0,1]} K_\theta(\alpha)\,.$$

Proof. When $\alpha < 1$, we make the following change of variables which consists in replacing ν by its pushforward by $x \mapsto 2\theta\sqrt{1-\alpha}x$. Using (24), we find that

$$H(\nu) = \int \log\frac{d\nu}{dx} d\nu = H(\nu_1) - \frac{1}{2}\log(1-\alpha) - \log(2\theta)$$

and $\int x^2 d\nu(x) = 4\alpha(1-\alpha)\theta^2$. Thus,

$$F(\theta) = \max\Big(\theta^2, \sup_{(\alpha,\nu)\in S'} K_\theta(\alpha, \nu)\Big), \tag{34}$$

where

$$S' = \Big\{(\alpha, \nu) \in (0, 1) \times \mathcal{P}(\mathbb{R}) : \int x^2 d\nu(x) = 4\alpha(1-\alpha)\theta^2\Big\},$$

and for any $\alpha \in (0,1), \nu \in \mathcal{P}(\mathbb{R})$,

$$K_\theta(\alpha,\nu) = \theta^2(\alpha^2 + B(1-\alpha)^2) + \int L(x)d\nu(x) - H(\nu) - \frac{1}{2}\log(1-\alpha) - \log(2\theta) - \frac{1}{2}\log(2\pi) - \frac{1}{2}.$$

By Lemma 5.3, we obtain for any $\alpha \in (0,1)$,

$$\sup_{\int x^2 d\nu(x) = 4\alpha(1-\alpha)\theta^2} K_\theta(\alpha,\nu) = \theta^2(\alpha^2 + B(1-\alpha)^2) + 4\theta^2\alpha(1-\alpha)\zeta_{\alpha,\theta} + G(\zeta_{\alpha,\theta}) - \frac{1}{2}\log(1-\alpha) - \log(2\theta) - \frac{1}{2}\log(2\pi) - \frac{1}{2},$$

where $\zeta_{\alpha,\theta} = \zeta_{4\theta^2\alpha(1-\alpha)}$. Hence, if we set, for $\alpha \in (0,1)$,

$$K_\theta(\alpha) = \theta^2(\alpha^2 + B(1-\alpha)^2) + R(4\theta^2\alpha(1-\alpha)) - \frac{1}{2}\log(1-\alpha) - \log(2\theta) - \frac{1}{2}\log(2\pi) - \frac{1}{2},$$

we deduce from (34) that

$$F(\theta) = \max\Big(\theta^2, \sup_{\alpha\in(0,1)} K_\theta(\alpha)\Big). \tag{35}$$

To study the maximum of K_θ, we will need the following result on the limit of R at 0, which will allow us to compute the limit of K_θ at 1.

Lemma 5.7. *When $C \to 0^+$,*

$$R(C) = \frac{1}{2} + \frac{1}{2}\log(2\pi C) + o(1).$$

Proof. For $C < l$, we have

$$G'(\zeta_C) = -C. \tag{36}$$

Since we have,

$$\lim_{C\to+\infty} G'(C) = 0,$$

and G' is invertible as $G''(x) > 0$ on $(-\infty, l)$, we get

$$\lim_{C\to 0} \zeta_C = +\infty.$$

From the inequalities, $0 \le L(x) \le Bx^2/2$, we deduce that by the definition (25) of G we have the bounds

$$\frac{1}{2}\log\frac{\pi}{\zeta} \le G(\zeta) \le \frac{1}{2}\log\frac{\pi}{\zeta - \frac{B}{2}},$$

which yields,

$$G(\zeta) \sim_{+\infty} \frac{1}{2}\log\frac{\pi}{\zeta}. \tag{37}$$

On the other hand, inserting these bounds in the numerator and the denominator of the derivative, we obtain

$$\sqrt{\frac{\zeta - \frac{B}{2}}{\zeta}}\frac{1}{2\zeta} \le -G'(\zeta) \le \sqrt{\frac{\zeta}{\zeta - \frac{B}{2}}}\frac{1}{2(\zeta - \frac{B}{2})}.$$

We deduce, since $\zeta_C \to +\infty$ as $C \to 0$,

$$G'(\zeta_C) = -\frac{1}{2\zeta_C} + o\Big(\frac{1}{\zeta_C}\Big).$$

Therefore, we get from the definition of ζ_C (36) that

$$\zeta_C \sim_0 \frac{1}{2C}.$$

Using (37), we can conclude that

$$R(C) = \frac{1}{2} + o(1) + \frac{1}{2}\log\frac{\pi}{\frac{1}{2C} + o\Big(\frac{1}{2C}\Big)} = \frac{1}{2} + \frac{1}{2}\log(2\pi C) + o(1).$$

□

From the previous Lemma 5.7, we deduce that

$$\lim_{\alpha\to 1} K_\theta(\alpha) = \theta^2,$$

so that we can continuously extend K_θ to 1. Therefore, we can write,

$$F(\theta) = \sup_{\alpha\in(0,1]} K_\theta(\alpha).$$

□

We now study K_θ and show that it is continuously differentiable on $(0,1)$. This amounts to prove that R is continuously differentiable on $(0,1)$. On $(0,l)$, it is clear that R is continuously differentiable due to the implicit function theorem. Indeed, ζ_C is by definition the unique solution of the equation

$$G'(\zeta) = -C,$$

and G is strictly convex. On $(l,+\infty)$, R is an affine function, therefore it is sufficient to prove that

$$\lim_{C\to l^-} R'(C) = \frac{B}{2}. \tag{38}$$

We have for any $C < l$,

$$R'(C) = C\partial\zeta_C + \zeta_C + \partial\zeta_C G'(\zeta_C) = \zeta_C,$$

which gives (38). We deduce that K_θ is continuously differentiable on $(0,1)$ and

$$\forall\alpha\in(0,1),\ K'_\theta(\alpha) = 2\theta^2(\alpha + B(\alpha-1)) + 4\theta^2\zeta_{\alpha,\theta}(1-2\alpha) + \frac{1}{2(1-\alpha)}.$$

From Lemma 5.7, we get,

$$\lim_{\alpha\to 0} K_\theta(\alpha) = -\infty.$$

Thus, the supremum of K_θ on $(0,1]$ is achieved either at 1 or on $(0,1)$. From Lemma 5.5, we have

$$F(\theta) = \max\big(K_\theta(1), \sup_{\alpha\le\frac{1}{2}} K_\theta(\alpha)\big).$$

Let us assume that the maximum of K_θ is achieved on $(0,1)$. We deduce that the maximum of K_θ is achieved on $(0,\frac{1}{2}]$ at a critical point since K_θ is differentiable. The critical points α of K_θ satisfy the equation,

$$2\theta^2(\alpha + B(\alpha-1)) + 4\theta^2\zeta_{\alpha,\theta}(1-2\alpha) + \frac{1}{2(1-\alpha)} = 0, \tag{39}$$

As $\theta^2(B-1) > 1$, $1/2$ does not satisfy the above equation so that the critical points of K_θ are the $\alpha \neq 1/2$, such that

$$\zeta_{\alpha,\theta} = \frac{2\theta^2(\alpha + B(\alpha-1)) + \frac{1}{2(1-\alpha)}}{4\theta^2(2\alpha-1)} := \varphi(\alpha). \tag{40}$$

We have,

$$\varphi(\alpha) = \frac{q(\alpha)}{8\theta^2(1-\alpha)(2\alpha-1)},$$

with

$$q(\alpha) = -4\theta^2(1+B)\alpha^2 + 4\theta^2(1+2B)\alpha - 4\theta^2 B + 1.$$

We find,

$$\varphi(\alpha) \geq \frac{B}{2} \iff \begin{cases} P(\alpha) \geq 0 & \text{if } \alpha \geq \frac{1}{2} \\ P(\alpha) \leq 0 & \text{if } \alpha \leq \frac{1}{2}, \end{cases}$$

with $P(\alpha) = 4\theta^2(B-1)\alpha^2 - 4\theta^2(B-1)\alpha + 1$. As $\zeta_{\alpha,\theta} \geq B/2$, we obtain that the maximum of K_θ is achieved at $\alpha \in (0,1/2)$ such that $P(\alpha) \leq 0$. The roots of P are

$$\alpha_\pm = \frac{1 \pm \sqrt{1 - [\theta^2(B-1)]^{-1}}}{2}. \tag{41}$$

Thus, the maximum of K_θ is achieved on $[\alpha_-, 1/2]$. We will show that K_θ is strictly concave on $(0,\frac{1}{2})$. Note that,

$$4\theta^2\alpha(1-\alpha) \geq l \iff \alpha \in [\beta_-, \beta_+],$$

with

$$\beta_\pm = \frac{1 \pm \sqrt{1 - l\theta^{-2}}}{2}.$$

For any $\alpha \in (\beta_-, \frac{1}{2})$, we must have $\zeta_C = B/2$ and therefore

$$K_\theta(\alpha) = \theta^2(\alpha^2 + B(1-\alpha)^2) + 2B\theta^2\alpha(1-\alpha) + G\Big(\frac{B}{2}\Big) - \frac{1}{2}\log(1-\alpha) + C_\theta,$$

where C_θ is some constant depending on θ. Thus, for $\alpha \in (\beta_-, \frac{1}{2})$,

$$K_\theta''(\alpha) = 2\theta^2(1-B) + \frac{1}{2(1-\alpha)^2} < 2\theta^2(1-B) + 2 < 0.$$

For any $\alpha \in (0, \beta_-)$, we have

$$\begin{aligned} K_\theta(\alpha) = {} & \theta^2(\alpha^2 + B(1-\alpha)^2) + 4\theta^2\alpha(1-\alpha)\zeta_{\alpha,\theta} + G(\zeta_{\alpha,\theta}) \\ & - \frac{1}{2}\log(1-\alpha) + C_\theta, \end{aligned}$$

where $\zeta_{\alpha,\theta}$ is such that

$$G'(\zeta_{\alpha,\theta}) = -4\theta^2\alpha(1-\alpha).$$

As G is strictly convex, we deduce by the implicit function theorem that $\alpha \in (0,\beta_-) \mapsto \zeta_{\alpha,\theta}$ is differentiable, and we have

$$\partial_\alpha \zeta_{\alpha,\theta} G''(\zeta_{\alpha,\theta}) = -4\theta^2(1-2\alpha).$$

We get that $\partial_\alpha \zeta_{\alpha,\theta} < 0$, for any $\alpha \in (0,\beta_-)$. Therefore, for $\alpha \in (0,\beta_-)$, we obtain

$$K''_\theta(\alpha) = 2\theta^2(B+1) - 8\theta^2\zeta_{\alpha,\theta} + 4\theta^2\partial_\alpha\zeta_{\alpha,\theta}(1-2\alpha) + \frac{1}{2(1-\alpha)^2}.$$

Using that $\zeta_{\alpha,\theta} > B/2$ and that $\partial_\alpha\zeta_{\alpha,\theta} < 0$ for $\alpha \in (0,\beta_-)$, we deduce,

$$\begin{aligned}\forall \alpha \in (0,\beta_-),\ K''_\theta(\alpha) &\le 2\theta^2(B+1) - 4\theta^2 B + \frac{1}{2(1-\alpha)^2} \\ &< 2\theta^2(1-B) + 2 \le 0.\end{aligned}$$

Thus, K'_θ is decreasing on $(0,\beta_-)$ and $(\beta_-,\frac{1}{2})$. Since K'_θ is continuous, we deduce that K'_θ is decreasing on $(0,\frac{1}{2})$ and K_θ is strictly concave on $(0,\frac{1}{2})$. Therefore, the maximum is achieved at the unique critical point of K_θ on $(0,\frac{1}{2})$ which we denote by α_θ. We distinguish two cases.

1^{st} *case:* $l \le \frac{1}{B-1}$. We have,

$$\beta_- \le \alpha_- \le \alpha_+ \le \beta_+.$$

We know that on one hand $P(\alpha_-) = 0$, so that

$$\varphi(\alpha_-) = \frac{B}{2}.$$

On the other hand $\zeta_{\alpha_-,\theta} = B/2$ since $\alpha_- \in [\beta_-,\beta_+]$. We deduce by (40) that α_- is a critical point of K_θ which lies in $(0,\frac{1}{2})$. Therefore

$$\alpha_\theta = \alpha_-.$$

2^{nd} *case:* $l > \frac{1}{B-1}$. We have,

$$\alpha_- < \beta_- < \beta_+ < \alpha_+.$$

Note that $0 \le \alpha_- < \frac{1}{2} < \alpha_+ \le 1$. Since $\varphi(\alpha) \ne B/2$ for any $\alpha \in [\beta_-,\beta_+]^c$, we deduce that $\alpha_\theta \in [\alpha_-,\beta_-)$, and in particular $K''_\theta(\alpha_\theta) < 0$. We deduce by the implicit function theorem that $\theta \mapsto \alpha_\theta$ is C^1, and therefore $\theta \mapsto K_\theta(\alpha_\theta)$ is continuously differentiable on $(1/\sqrt{B-1},+\infty)$. In conclusion, we have shown that for any $\theta^2(B-1) \ge 1$, if $l \le \frac{1}{B-1}$,

$$F(\theta) = \max\big(\theta^2, K_\theta(\alpha_-)\big),$$

where α_- is defined in (41), whereas if $l \ge \frac{1}{B-1}$,

$$F(\theta) = \max\big(\theta^2, K_\theta(\alpha_\theta)\big),$$

where α_θ is the unique solution in $(0,\beta_-)$ such that

$$G'(\zeta_{\alpha,\theta}) = -4\theta^2\alpha(1-\alpha).$$

To conclude that F is continuously differentiable on $(1/\sqrt{B-1}, +\infty)$ except at most at one point, we show that there exists θ_0 such that

$$\forall \theta \le \theta_0,\ F(\theta) = \theta^2, \text{ and } \forall \theta > \theta_0,\ F(\theta) > \theta^2.$$

Since $F(\theta) \ge \theta^2$ for any $\theta \ge 0$, it suffices to prove that $\theta \mapsto F(\theta) - \theta^2$ is non-decreasing. Recall that

$$F(\theta) = \lim_{N \to +\infty} F_N(\theta),$$

where

$$F_N(\theta) = \frac{1}{N} \log \mathbb{E}_e \exp\Big(\sum_i L\big(\sqrt{2N}\theta e_i^2\big) + \sum_{i<j} L\big(2\sqrt{N}\theta e_i e_j\big)\Big),$$

and e is uniformly sampled on $\mathbb{S}^{N-1}$. Therefore,

$$F_N(\theta) - \theta^2 = \frac{1}{N} \log \mathbb{E}_e \exp\Big(\sum_i 2N\theta^2 \Big(\psi\big(\sqrt{2N}\theta e_i^2\big) - \frac{1}{2}\Big) e_i^4 + \sum_{i<j} 4N\theta^2 \Big(\psi\big(2\sqrt{N}\theta e_i e_j\big) - \frac{1}{2}\Big) e_i^2 e_j^2 \Big).$$

As ψ is increasing and $\psi(0) = 1/2$, $\theta \mapsto F_N(\theta) - \theta^2$ is non-decreasing, and therefore $\theta \mapsto F(\theta) - \theta^2$ is non-decreasing as well.

For the sake of completeness, we show the following Proposition which indicates that it is unlikely we could prove the large deviation principle for all values of x by following our strategy because F is in general not differentiable everywhere.

Proposition 5.8. *Assume $\theta_0 = \inf\{\theta \in \mathbb{R}^+ : F(\theta) > \theta^2\} > 1/\sqrt{B-1}$. Then, F is not differentiable at θ_0.*

Proof. Let $\theta > \theta_0$. We know from Lemma 5.5 that

$$F(\theta) = \max_{\substack{\int x^2 d\nu(x) \le \alpha \\ \alpha \le \frac{1}{2}}} H_\theta(\alpha, \nu),$$

where for any $\alpha \in (0,1)$ and $\nu \in \mathcal{P}(\mathbb{R})$,

$$H_\theta(\alpha, \nu) = \theta^2 \alpha^2 + \theta^2 B (1-\alpha)^2 + \int L(2\theta\sqrt{1-\alpha}x) d\nu(x) - H(\nu) - \frac{1}{2}\log(2\pi) - \frac{1}{2}.$$

Since $\theta_0 \ge 1/\sqrt{B-1}$, we know from the proof of Proposition 5.4 that there exists $\alpha_0 \le 1/2$ and $\nu_0 \in \mathcal{P}(\mathbb{R})$ such that

$$H_{\theta_0}(\alpha_0, \nu_0) = F(\theta_0).$$

Define $g(\theta) = H_\theta(\alpha_0, \nu_0)$ for any $\theta \ge \theta_0$. Let F'_+ denote the right derivatives of F. We have as $F \ge g$ and $F(\theta_0) = g(\theta_0)$,

$$F'_+(\theta_0) \ge g'(\theta_0).$$

We find

$$g'(\theta_0) = 2\theta_0(\alpha_0^2 + B(1-\alpha_0)^2) + 2\sqrt{1-\alpha_0} \int x L'(2\theta_0\sqrt{1-\alpha_0}x) d\nu_0(x).$$

Since ψ is increasing, we have $xL'(x) \geq 2L(x)$, and $L(x) \geq x^2/2$ for any $x \geq 0$. Therefore, $xL'(x) \geq x^2$ and we get

$$\begin{aligned} g'(\theta_0) &\geq 2\theta_0(\alpha_0^2 + B(1-\alpha_0)^2) + 4\theta_0\alpha_0(1-\alpha_0) \\ &\geq 2\theta_0 + 2\theta_0(1-\alpha)^2(B-1), \end{aligned}$$

which shows that $g'(\theta_0) > 2\theta_0$ and therefore $F'_+(\theta) > 2\theta_0$. It yields that F is not differentiable at θ_0.

□

5.3. **Proof of Proposition 1.13.** By Proposition 5.4, we know that F is differentiable on $(1/\sqrt{B-1}, +\infty)$ except possibly at θ_0. Using Proposition 2.2 we deduce that there exists x_μ such that the lower large deviation lower bound holds with rate function $I(x) = \overline{I}(x)$ for any $x \geq x_\mu$.

6. The case $B < A$

We consider the case where the following assumption holds.

Assumption 6.1. *B exists and is strictly smaller than A. Moreover, we assume that ψ achieves its maximum A at a unique point m^* such that $\psi''(m_*) < 0$.*

The first condition includes in particular the case where the law of the entries have a compact support (since in this case $B = 0$) and we believe the second condition is true quite generically, as we check in the following example.

Example 6.1. *Let*

$$\mu = \frac{p}{2}(\delta_{-1/\sqrt{p}} + \delta_{1/\sqrt{p}}) + (1-p)\delta_0, \qquad \psi(x) = \frac{1}{x^2}\log(p(\cosh(\frac{x}{\sqrt{p}}) - 1) + 1)\,.$$

Then, we show that for $p < 1/3$, μ satisfies Assumption 6.1 (but for $p > 1/3$ μ has a sharp-subGaussian tail). We have

$$\forall x \geq 0,\ \psi'(x) = \frac{L'(x)}{x^2} - \frac{2L(x)}{x^3},\ \psi''(x) = \frac{L''(x)}{x^2} - \frac{4L'(x)}{x^3} + \frac{6L(x)}{x^4}.$$

We claim that $h : x \mapsto xL'(x) - 2L(x)$ is increasing and then decreasing on $\mathbb{R}_+$. Indeed,

$$\forall x \geq 0,\ h'(x) = xL''(x) - L'(x),\ h''(x) = xL^{(3)}(x),$$

and we have,

$$L(x) = \log\left(p\cosh\left(\frac{x}{\sqrt{p}}\right) + 1 - p\right),\ L'(x) = \frac{\sqrt{p}\sinh\left(\frac{x}{\sqrt{p}}\right)}{p\cosh\left(\frac{x}{\sqrt{p}}\right) + 1 - p}.$$

$$L''(x) = \frac{p + (1-p)\cosh\left(\frac{x}{\sqrt{p}}\right)}{(p\cosh\left(\frac{x}{\sqrt{p}}\right) + 1 - p)^2},\ L^{(3)}(x) = \frac{\frac{(1-p)^2}{\sqrt{p}} - 2p\sqrt{p} - \sqrt{p}(1-p)\cosh\left(\frac{x}{\sqrt{p}}\right)}{(p\cosh\left(\frac{x}{\sqrt{p}}\right) + 1 - p)^3}\sinh\left(\frac{x}{\sqrt{p}}\right).$$

We have, for $p > p_* = 1/3$

$$\frac{(1-p)^2}{\sqrt{p}} - 2p\sqrt{p} \geq \sqrt{p}(1-p).$$

Therefore, $L^{(3)}$ is negative and therefore h' is decreasing. Since $h'(0) = 0$, we deduce that h' is negative and ψ is decreasing. If $p > p_*$, we have that h'' is positive and then negative. Therefore, h' is increasing on $[0, x_0]$ and then decreasing on $[x_0, +\infty)$, with $x_0 = \sqrt{p}\cosh^{-1}(\frac{1-2p-p^2}{p(1-p)})$. But,

$$h'(0) = 0, \ \lim_{x\to+\infty} h'(x) = -\frac{1}{\sqrt{p}},$$

as $L'(x) \sim_{+\infty} 1/\sqrt{p}$ and $L''(x) \sim_{+\infty} 2(1-1/p)e^{-x/\sqrt{p}}$. Therefore, there exists $m_* > x_0$ such that h' is positive on $(0, m_*)$ and negative on $(m_*, +\infty)$. We deduce that ψ is increasing on $(0, m_*)$ and decreasing on $(m_*, +\infty)$ so that ψ achieves its unique maximum at m_*. Moreover,

$$\begin{aligned}\psi'(m_*) = 0, \psi''(m_*) = 0 &\iff m_*L'(m_*) = 2L(m_*), \ m_*^2L''(m_*) = 4m_*L'(m_*) - 6L(m_*)\\ &\iff L'(m_*) = m_*L''(m_*), \ m_*L'(m_*) = 2L(m_*).\end{aligned}$$

As $m_* > x_0$, we have that $h'(m_*) < 0$ and therefore $\psi''(m_*) < 0$. Note that $x_0/\sqrt{p}$ goes to infinity when p goes to zero and that

Studying the variational problem arising from the limit of the annealed spherical integral $\overline{F}(\theta)$ and $\underline{F}(\theta)$ defined in Proposition 1.6, we will show that for θ large enough we can give an explicit formula as stated in the following proposition.

Proposition 6.2. *There exists $\theta_0 > 1/\sqrt{A-1}$ such that for any $\theta \geq \theta_0$, $\overline{F}(\theta) = \underline{F}(\theta) = F(\theta)$ where*

$$F(\theta) = \sup_{\alpha\in(0,1]} V(\alpha),$$

with

$$\forall \alpha > 0, \ V(\alpha) = \theta^2(A-1)\alpha^2 + \theta^2 + \frac{1}{2}\log(1-\alpha).$$

More explicitly,

$$F(\theta) = \frac{\theta^2}{4}(A-1)\Big(1 + \sqrt{1 - \frac{1}{\theta^2(A-1)}}\Big)^2 + \theta^2 + \frac{1}{2}\log\Big(1 - \sqrt{1 - \frac{1}{\theta^2(A-1)}}\Big) - \frac{1}{2}\log 2.$$

We prove this proposition by first showing that $\underline{F}(\theta) \geq F(\theta)$ for all θ and then that, for large θ, $\overline{F}(\theta) \leq F(\theta)$.

6.1. **Proof of the lower bound.** Recall that by Proposition 1.6, we have the following formulation of the limit $\underline{F}(\theta)$.

$$\underline{F}(\theta) = \sup_{\substack{\alpha_1+\alpha_2+\alpha_3=1\\ \alpha_i\geq 0}} \liminf_{\substack{\delta\to 0, K\to+\infty\\ \delta K\to 0}} \limsup_{N\to+\infty} \mathcal{F}^N_{\alpha_1,\alpha_2,\alpha_3}(\delta, K)\,,$$

where

$$\mathcal{F}^N_{\alpha_1,\alpha_2,\alpha_3}(\delta,K)=\theta^2\big(\alpha_1^2+2\alpha_1\alpha_2+B\alpha_3^2\big)$$
$$+\sup_{\substack{t_i\in I_2,i\le l\\ |\sum_i t_i^2-N\alpha_2|\le\delta N}}\ \sup_{\substack{s_i\in I_3,i\le k\\ |\sum_i s_i^2-N\alpha_3|\le\delta N}}\Big\{\frac{1}{N}\sum_{i=1}^k\sum_{j=1}^l L\Big(\frac{2\theta s_it_j}{\sqrt{N}}\Big)+\frac{1}{2N}\sum_{i,j=1}^l L\Big(\frac{2\theta t_it_j}{\sqrt{N}}\Big)$$
$$+\sup_{\substack{\nu_1\in\mathcal{P}(I_1)\\ \int x^2d\nu_1(x)=\alpha_1}}\Big\{\sum_{i=1}^k\int L\Big(\frac{2\theta s_ix}{\sqrt{N}}\Big)d\nu_1(x)-H(\nu_1)\Big\}-\frac{1}{2}\log(2\pi)-\frac{1}{2}\Big\},$$

Our goal is to show that we can take $\alpha_3=0$ and in the supremum defining $\mathcal{F}^N_{\alpha_1,\alpha_2,\alpha_3}(\delta,K)$ we can take all the t_i's equal. In fact we first prove the lower bound:

Lemma 6.3. *For any $\theta\ge 0$,*

$$\underline{F}(\theta)\ge\sup_{\alpha\in(0,1]}V(\alpha),$$

where V is defined in Proposition 6.2.

Proof. Indeed, if we take $\alpha_3=0$ and $t_j=N^{1/4}\sqrt{\frac{m^*}{2\theta}}, 1\le j\le l$, $\alpha_2\in[lm^*/2\theta\sqrt{N}-\delta, lm^*/2\theta\sqrt{N}+\delta]$, $\alpha_1=1-\alpha_2$, ν_1 to be the Gaussian law restricted to I_1 with variance α_1, then we get the lower bound

$$\mathcal{F}^N_{\alpha_1,\alpha_2,0}(\delta,K)\ge\theta^2(\alpha_1^2+2\alpha_1\alpha_2+\alpha_2^2A)+\frac{1}{2}\log\alpha_1=V(\alpha_2)\,.$$

Hence, to derive the lower bound it is enough to remark that we can achieve any possible value of α_2 in $[0,1]$ as some large N limit of $l_Nm^*/2\theta\sqrt{N}$ for some sequence of integer numbers l_N, which is obvious. □

6.2. **Proof of the upper bound.** The rest of this section is devoted to prove that the previous lower bound is sharp when θ is big enough. To this end, recall that by Proposition 1.6, we have the following formulation of the limit $\overline{F}(\theta)$.

$$\overline{F}(\theta)=\sup_{\substack{\alpha_1+\alpha_2+\alpha_3=1\\ \alpha_i\ge0}}\ \limsup_{\substack{\delta\to0,K\to+\infty\\ \delta K\to0}}\ \limsup_{N\to+\infty}\mathcal{F}^N_{\alpha_1,\alpha_2,\alpha_3}(\delta,K)\,.$$

We first reformulate the supremum in $\mathcal{F}^N_{\alpha_1,\alpha_2,\alpha_3}(\delta,K)$ by denoting for $t\in I_2^l$ so that $|\sum t_i^2-N\alpha_2|\le\delta N$,

$$\mu_2=\frac{1}{\alpha_2N}\sum_{i=1}^l t_i^2\delta_{\frac{\sqrt{2\theta}t_i}{N^{1/4}}}\,.$$

μ_2 is a positive measure on $S_2=\{x:\sqrt{2\delta\theta}\le|x|\le\sqrt{2K\theta}\}$ whose mass belongs to $[1-\frac{\delta}{\alpha_2},1+\frac{\delta}{\alpha_2}]$. We also denote by $S_3=\{x:\sqrt{K}\le|x|\le N^{1/4}\sqrt{\alpha_3}\}$. Then it is not hard to see that for any $\theta\ge0$,

$$\overline{F}(\theta)\le\hat{F}(\theta),\tag{42}$$

where $\hat{F}(\theta)$ is defined by

$$\hat{F}(\theta)=\sup_{\substack{\alpha_1+\alpha_2+\alpha_3=1\\ \alpha_i\ge0}}\ \limsup_{\substack{\delta\to0,K\to+\infty\\ \delta K\to0}}\ \limsup_{N\to+\infty}\ \sup_{\mu_2\in\mathcal{P}(S_2)}\ \sup_{s\in S_3}\mathcal{G}^N_{\alpha_1,\alpha_2,\alpha_3}(\delta,K,s,\mu_2)$$

if

$$\mathcal{G}^N_{\alpha_1,\alpha_2,\alpha_3}(\delta,K) = \theta^2\big(\alpha_1^2+2\alpha_1\alpha_2+B\alpha_3^2\big) + 4\theta^2\alpha_3\alpha_2\int\psi\big(\sqrt{2\theta}sx\big)d\mu_2(x) + 2\theta^2\alpha_2^2\int\psi(xy)d\mu_2(x)d\mu_2(y) + \sup_{\substack{\nu_1\in\mathcal{P}(I_1)\\ \int x^2d\nu_1(x)=\alpha_1}}\Big\{4\theta^2\alpha_3\int x^2\psi\Big(\frac{2\theta sx}{N^{\frac14}}\Big)d\nu_1(x)-H(\nu_1)\Big\}-\frac12\log(2\pi)-\frac12.$$

Indeed, the upper bound proceeds in two steps: first we take the supremum over all measures μ_2 on S_2 with mass in $[1-\frac{\delta}{\alpha_2},1+\frac{\delta}{\alpha_2}]$, and then restrict ourselves to probability measures as δ goes to zero (since ψ is bounded). Then, we observe that for any $\mu_2\in\mathcal{P}(S_2)$, $\nu_1\in\mathcal{P}(I_1)$, and $s\in S_3^k$ such that $|\sum_i s_i^2-\alpha_3\sqrt{N}|\le\delta\sqrt{N}$,

$$\frac{\alpha_2}{\sqrt{N}}\sum_{i=1}^k s_i^2\int\psi\big(\sqrt{2\theta}s_ix\big)d\mu_2(x)+\frac{1}{\sqrt{N}}\sum_{i=1}^k s_i^2\int x^2\psi\Big(\frac{2\theta s_ix}{N^{\frac14}}\Big)d\nu_1(x) \le \alpha_3\int\psi\big(\sqrt{2\theta}sx\big)d\mu_2(x)+\int x^2\psi\Big(\frac{2\theta sx}{N^{1/4}}\Big)d\nu_1(x)+o_\delta(1),$$

where s is a maximizer of the function

$$s\in S_3\mapsto\alpha_2\int\psi\big(\sqrt{2\theta}sx\big)d\mu_2(x)+\int x^2\psi\Big(\frac{2\theta sx}{N^{1/4}}\Big)d\nu_1(x),$$

which ends the proof of the claim (42). We will see that under our assumptions that $B<A$ and that the maximum of ψ is uniquely achieved at m^* such that $\psi''(m^*)<0$, the upper bound $\hat F(\theta)$ is sharp when θ is large.

The starting point of our analysis of the variational problem defining $\hat F(\theta)$ in the regime where θ is large is the fact that $\underline{F}(\theta)$ and $\hat F(\theta)$ behave like $A\theta^2$. More precisely, we know from (10) that there exists $\theta_0>0$ (depending on A) such that for all $\theta\ge\theta_0$,

$$\hat F(\theta)\ge\underline{F}(\theta)\ge A\theta^2-\kappa\log\theta, \tag{43}$$

where $\kappa>0$ is a numerical constant.

As a consequence, we can localize the suprema over $(\alpha_1,\alpha_2,\alpha_3)$ and μ_2 in the definitions of $\hat F(\theta)$ in some subset of the constraint set, denoted by $\mathcal{S}$, and defined as follow,

$$\mathcal{S}=\Big\{(\underline{\alpha},\mu_2)\in[0,1]^3\times\mathcal{P}(S_2):\alpha_1+\alpha_2+\alpha_3=1\Big\}.$$

Lemma 6.4. *There exists a constant $\theta_0>0$ depending on A such that for any $\theta\ge\theta_0$, the suprema defining $\hat F(\theta)$ can be restricted to the set $\mathcal{A}_\theta\times\mathcal{B}_\theta\subset\mathcal{S}$ defined by,*

$$\underline{\alpha}\in\mathcal{A}_\theta\iff\alpha_2\ge1-\frac{C\sqrt{\log\theta}}{\theta},\ \alpha_1\le\frac{C\sqrt{\log\theta}}{\theta},\ \alpha_3\le\frac{C\sqrt{\log\theta}}{\theta},$$

and

$$\mu_2\in\mathcal{B}_\theta\iff\int\Big(\frac{A}{2}-\psi(xy)\Big)d\mu_2(x)d\mu_2(y)\le\frac{C\log\theta}{\theta^2}.$$

where C is a some positive constant depending also on A.

Proof. From (43) we deduce that we can restrict the suprema in the definitions of $\hat{F}(\theta)$ to the parameters $\underline{\alpha}, s, \nu_1, \mu_2$ with $\alpha_1+\alpha_2+\alpha_3=1$, $s\in S_3$, $\int x^2 d\mu_1(x)=\alpha_1$ such that,

$$(A-1)(\alpha_1^2+2\alpha_1\alpha_2)+(A-B)\alpha_3^2+4\alpha_2\alpha_3\int\Big(\frac{A}{2}-\psi(2\theta sy)\Big)d\mu_2(y)$$
$$+2\alpha_2^2\int\Big(\frac{A}{2}-\psi(2\theta xy)\Big)d\mu_2(y)d\mu_2(x)+4\alpha_3\int y^2\Big(\frac{A}{2}-\psi(2\theta sy)\Big)d\nu_1(y)$$
$$+\frac{1}{\theta^2}\big(H(\nu_1)+\log\sqrt{2\pi}\big)\leq\frac{2\kappa\log\theta}{\theta^2}.$$

But

$$4\alpha_3\int y^2\Big(\frac{A}{2}-\psi(2\theta sy)\Big)d\nu_1(y)+\frac{1}{\theta^2}\big(H(\nu_1)+\log\sqrt{2\pi}\big)\geq\frac{1}{2}\log\frac{1}{\alpha_1}\geq 0.$$

Therefore,

$$(A-1)(\alpha_1^2+2\alpha_1\alpha_2)+(A-B)\alpha_3^2+4\alpha_2\alpha_3\int y^2\Big(\frac{A}{2}-\psi(2\theta sy)\Big)d\mu_2(y)$$
$$+2\alpha_2^2\int\Big(\frac{A}{2}-\psi(xy)\Big)d\mu_2(y)d\mu_2(x)\leq\frac{2\kappa\log\theta}{\theta^2}.$$

Since each term is non-negative, they are all bounded by $2\kappa\log\theta/\theta^2$. Note that this already yields with $C=2\kappa/\min\{(A-B),A-1\}$,

$$\alpha_1^2\leq\frac{C\log\theta}{\theta^2},\ \alpha_3^2\leq\frac{C\log\theta}{\theta^2},\ \alpha_1\alpha_2\leq\frac{C\log\theta}{\theta^2}. \tag{44}$$

The two first estimates imply since $\alpha_2=1-\alpha_1-\alpha_3$,

$$\alpha_2\geq 1-2\frac{\sqrt{C\log\theta}}{\theta}.$$

□

Next, note that because ψ is bounded continuous, the function $\mathcal{G}^N_{\alpha_1,\alpha_2,\alpha_3}(\delta,K,s,.)$ we are optimizing over μ_2 is bounded continuous in μ_2 and therefore it achieves its maximal value. We denote by μ_2 such an optimizer. In the next lemma we prove that the optimizers of $\mathcal{G}^N_{\alpha_1,\alpha_2,\alpha_3}(\delta,K,s,.)$ are concentrated around $\sqrt{m_*}$ if ψ takes its maximum value at m^* only.

Lemma 6.5. *Assume that ψ achieves its maximum value at m^* only and that it is strictly concave in an open neighborhood of this point. Let μ_2 be an optimizer There exists $\varepsilon_0>0$ such that for any $\mu_2\in\mathcal{B}_\theta$,*

$$\forall 0<\varepsilon<\varepsilon_0,\ \mu_2\big(|x-\sqrt{m^*}|\geq\varepsilon\big)\leq\frac{C\sqrt{\log\theta}}{\theta\varepsilon},$$

where C is a positive constant depending on ψ.

Proof. Let $\mu_2\in\mathcal{B}_\theta$. By Lemma 6.4 we have,

$$\int\Big(\frac{A}{2}-\psi(xy)\Big)d\mu_2(x)d\mu_2(y)\leq\frac{C\log\theta}{\theta^2}.$$

Since ψ is strictly concave in a neighborhood of m_*, and m_* is its unique maximizer, we deduce that there exists $\eta_0 > 0$ such that for all $0 < \eta < \eta_0$,

$$\forall |x - m_*| \in [\sqrt{\eta}, \sqrt{\eta_0}], \quad \frac{A}{2} - \psi(x) \geq \eta/c,$$

for some constant $c > 0$. As ψ is analytic, it admits a finite number of local maximum. Therefore, we can find $\eta_0 > 0$ such that for all $0 < \eta < \eta_0$,

$$\forall |x - m_*| \geq \sqrt{\eta}, \quad \frac{A}{2} - \psi(x) \geq \eta/c,$$

Since $\frac{A}{2} - \psi$ is non-negative, we obtain

$$\forall \eta < \eta_0, \ \mu_2^{\otimes 2}\big(|xy - m^*| \geq \sqrt{\eta}\big) \leq \frac{C' \log \theta}{\eta \theta^2},$$

where $C' \geq 1$ is a constant depending on ψ. For $\varepsilon < \sqrt{m_*}/2$, we have

$$\mu_2([0, \sqrt{m_*} - \varepsilon])^2 \leq \mu_2^{\otimes 2}\big(xy \leq m_* - \sqrt{m_*}\varepsilon\big).$$

Therefore, for ε small enough,

$$\mu_2([0, \sqrt{m_*} - \varepsilon])^2 \leq \frac{C' \log \theta}{m_* \theta^2 \varepsilon^2}.$$

On the other hand,

$$\mu_2\big([\sqrt{m_*} + \varepsilon, +\infty)\big)^2 \leq \mu_2^{\otimes 2}\big(xy \geq m_* + 2\sqrt{m_*}\varepsilon\big).$$

Using a union bound, we obtain the claim. □

Using Lemma 6.5, we will show that the optimization problem over μ_2 is asymptotically solved by $\delta_{\sqrt{m_*}}$, with an error which vanishes when K, the lower boundary point of S_3, goes to $+\infty$.

Lemma 6.6. *There exists θ_0 depending on ψ such that for $\theta \geq \theta_0$, such that for any $\underline{\alpha} \in \mathcal{A}_\theta$ and $s \in S_3$, if the measure μ_2 that realizes the maximum:*

$$\sup_{\mu_2 \in \mathcal{P}(S_2)} \Big\{ 2\alpha_3 \int \psi\big(\sqrt{2\theta} s x\big) d\mu_2(x) + \alpha_2 \int \psi(xy) d\mu_2(x) d\mu_2(y) \Big\}$$

is in $\mathcal{B}_\theta$, then

$$\sup_{\mu_2 \in \mathcal{B}_\theta} \Big\{ 2\alpha_3 \int \psi\big(\sqrt{2\theta} s x\big) d\mu_2(x) + \alpha_2 \int \psi(xy) d\mu_2(x) d\mu_2(y) \Big\} = \frac{A\alpha_2}{2} + \frac{B\alpha_3}{2} + o_K(1).$$

Proof. Letting $\overline{\psi}(x) = \psi(x) - \frac{B}{2}$, it is equivalent to show that :

$$\sup_{\mu_2 \in \mathcal{B}_\theta} \Big\{ 2\alpha_3 \int \overline{\psi}\big(\sqrt{2\theta} s x\big) d\mu_2(x) + \alpha_2 \int \overline{\psi}(xy) d\mu_2(x) d\mu_2(y) \Big\} = \frac{(A - B)\alpha_2}{2} + o_K(1).$$

Let us fix $\theta \geq \theta_0$ where θ_0 is given by Lemma 6.4. Observe that since ψ is bounded continuous,

$$Z : \mu \in \mathcal{P}(S_2) \mapsto 2\alpha_3 \int \overline{\psi}\big(\sqrt{2\theta} s x\big) d\mu(x) + \alpha_2 \int \overline{\psi}(xy) d\mu(x) d\mu(y)$$

achieves its maximum value in the closed set $\mathcal{B}_\theta$. Let μ_2 be an optimizer, and therefore a critical point of this function. Writing that $Z(\mu_2) \geq Z(\mu_2 + \varepsilon\nu)$ for all measures ν on S_2 such that $\mu_2 + \varepsilon\nu$ is a probability measure for small ϵ, we deduce that there exists a constant $C > 0$ such that,

$$\forall x \in S_2,\ \alpha_3 \overline{\psi}(\sqrt{\theta} s x) + \alpha_2 \int \overline{\psi}(xy) d\mu_2(y) \leq C, \tag{45}$$

with equality μ_2-almost surely. Using Lemma 6.5, we get for any ε small enough,

$$\int \overline{\psi}(xy) d\mu_2(y) = \overline{\psi}(\sqrt{m_*}x) + \int_{[\sqrt{m_*}-\varepsilon, \sqrt{m_*}+\varepsilon]} \big(\overline{\psi}(xy) - \overline{\psi}(\sqrt{m_*}x)\big) d\mu_2(x) + O\Big(\frac{\sqrt{\log\theta}}{\theta\varepsilon}\Big).$$

Where we notice that our $O\Big(\frac{\sqrt{\log\theta}}{\theta\varepsilon}\Big)$ is a function that does not depend on δ, K or N. As L is the log-Laplace transform of a sub-Gaussian distribution, we have that $|L'|$ is bounded. In particular, $|\psi'|$ is bounded and thus ψ is Lipschitz. Therefore, for any $x \leq M$,

$$\int \overline{\psi}(xy) d\mu_2(y) = \overline{\psi}(\sqrt{m^*}x) + O\Big(\varepsilon M + \frac{\sqrt{\log\theta}}{\theta\varepsilon}\Big).$$

Again, the $O\Big(\varepsilon M + \frac{\sqrt{\log\theta}}{\theta\varepsilon}\Big)$ does not depend on δ, K or N. We choose $\varepsilon = \theta^{-1/2}$ and $M = \theta^{1/4}$ so that the two error term above goes to zero when θ goes to ∞, so that we have for any $x \geq 0$,

$$\int \overline{\psi}(xy) d\mu_2(y) = \overline{\psi}(\sqrt{m_*}x) + o_\theta(1). \tag{46}$$

In particular,

$$\alpha_3 \overline{\psi}(\sqrt{\theta} s x) + \alpha_2 \int \overline{\psi}(xy) d\mu_2(y) = \overline{\psi}(\sqrt{m_*}x) + o_\theta(1).$$

Taking $x = \sqrt{m_*}$ in (45), we get

$$C \geq \frac{A - B}{2} + o_\theta(1), \tag{47}$$

since $s \geq K$ and $1 - \alpha_2 \leq O(\frac{\sqrt{\log\theta}}{\theta})$. The terms $o_\theta(1)$ above do not depend on K, δ or N. We claim that there exists θ_0 such that for any $\theta \geq \theta_0$,

$$\mu_2([0, \sqrt{m_*}/2]) = 0.$$

Indeed, if $x \leq \sqrt{m_*}/2$, we have by (46) and the fact that α_2 goes to 1 as θ goes to infinity,

$$\alpha_3 \overline{\psi}(\sqrt{\theta} s x) + \alpha_2 \int \overline{\psi}(xy) d\mu(y) \leq \sup_{t \leq \sqrt{m_*}/2} \overline{\psi}(\sqrt{m_*}t) + o_\theta(1),$$

with $\sup_{t \leq \sqrt{m_*}/2} \overline{\psi}(\sqrt{m_*}t) < (A - B)/2$ since the maximum of ψ is uniquely achieved at m_*. From (47) and the fact that equality in (45) holds μ_2-a.s, we deduce that for θ large

enough (and not depending on δ, K or N) $[0, \sqrt{m_*}/2] \cap \mathrm{supp}(\mu_2) = \emptyset$. Therefore,

$$2\alpha_3 \int \overline{\psi}\big(\sqrt{2\theta} s x\big) d\mu_2(x) + \alpha_2 \int \overline{\psi}(xy) d\mu_2(x) d\mu_2(y) \leq \frac{(A-B)\alpha_2}{2} + 2 \sup_{y \geq K \frac{\sqrt{m_* \theta}}{2}} \overline{\psi}(y)$$
$$= \frac{(A-B)\alpha_2}{2} + o_K(1).$$

Thus,

$$\sup_{\mu_2 \in \mathcal{B}_\theta} \Big\{ 2\alpha_3 \int \overline{\psi}\big(\sqrt{2\theta} s x\big) d\mu_2(x) + \alpha_2 \int \overline{\psi}(xy) d\mu_2(x) d\mu_2(y) \Big\} \leq \frac{(A-B)\alpha_2}{2} + o_K(1).$$

The reverse inequality is achieved by taking $\mu_2 = \delta_{\sqrt{m_*}}$, which completes the proof. □

We deduce that f

Proposition 6.7. *There exists θ_0 depending on ψ such that for any $\theta \geq \theta_0$, $\overline{F}(\theta) \leq \hat{F}(\theta)$, where*

$$\hat{F}(\theta) = \sup_{(\underline{\alpha}, s, \nu) \in \mathcal{S}'} \mathcal{F}(\underline{\alpha}, s, \nu),$$

with

$$\mathcal{F}(\underline{\alpha}, s, \nu) = \theta^2 \big(\alpha_1^2 + 2\alpha_1 \alpha_2\big) + \theta^2 A \alpha_2^2 + \theta^2 B(\alpha_3^2 + 2\alpha_3 \alpha_2)$$
$$+ 4\theta^2 \alpha_3 \int x^2 \psi(2\theta s \sqrt{\alpha_3} x) d\nu(x) - H(\nu) - \frac{1}{2} \log(2\pi) - \frac{1}{2},$$

and

$$\mathcal{S}' = \Big\{ (\underline{\alpha}, s, \nu) \in [0,1]^3 \times [0,1] \times \mathcal{P}(\mathbb{R}) : \alpha_1 + \alpha_2 + \alpha_3 = 1, \int x^2 d\nu(x) = \alpha_1 \Big\}.$$

Proof. By Lemmas 6.4 and 6.6, we know that

$$\hat{F}(\theta) = \sup_{\underline{\alpha} \in \mathcal{A}_\theta} \limsup_{\substack{\delta \to 0, K \to +\infty \\ \delta K \to 0}} \limsup_{N \to +\infty} \hat{\mathcal{F}}^N_{\alpha_1, \alpha_2, \alpha_3}(\delta, K),$$

where

$$\hat{\mathcal{F}}^N_{\alpha_1, \alpha_2, \alpha_3}(\delta, K) = \sup_{s \in S_3} \sup_{\substack{\nu_1 \in \mathcal{P}(I_1) \\ \int x^2 d\nu_1(x) = \alpha_1}} \Big\{ 4\theta^2 \alpha_3 \int x^2 \psi\Big(\frac{2\theta s x}{N^{\frac{1}{4}}}\Big) d\nu_1(x) - H(\nu_1) \Big\}$$
$$+ \theta^2 \big(\alpha_1^2 + 2\alpha_1 \alpha_2\big) + A\alpha_2^2 + B(\alpha_3^2 + 2\alpha_2 \alpha_3) - \frac{1}{2} \log(2\pi) - \frac{1}{2},$$

$S_3 = [K, N^{1/4} \sqrt{\alpha_3}]$. Using the change of variable $s \mapsto s N^{-1/4}$ we have the upper bound,

$$\hat{F}(\theta) \leq \sup_{(\underline{\alpha}, s, \nu) \in \mathcal{S}'} \mathcal{F}(\underline{\alpha}, s, \nu).$$

□

We finally prove that the supremum is taken at $\alpha_3 = 0$.

Proposition 6.8. *There exists θ_0 depending on A such that for any $\theta \geq \theta_0$,*

$$\sup_{(\underline{\alpha}, s, \nu) \in \mathcal{S}'} \mathcal{F}(\underline{\alpha}, s, \nu) = \sup_{(\alpha_1, \alpha_2, 0, s, \nu) \in \mathcal{S}'} \mathcal{F}((\alpha_1, \alpha_2, 0), s, \nu).$$

Proof. We claim that for any $((\alpha_1,\alpha_2,\alpha_3),s,\nu)\in\mathcal{S}'$ such that $\alpha_2\geq\frac{A-1}{2A-B-1}$, we have

$$\mathcal{F}((\alpha_1,\alpha_2,\alpha_3),s,\nu)\leq\sup_{\nu\in\mathcal{P}(\mathbb{R})}\mathcal{F}((\alpha_1,\alpha_2+\alpha_3,0),\nu). \tag{48}$$

Note that

$$\sup_{\nu\in\mathcal{P}(\mathbb{R})}\mathcal{F}((\alpha_1,\alpha_2+\alpha_3,0),\nu)=\theta^2(\alpha_1+2\alpha_1\alpha_2+2\alpha_1\alpha_3)+\theta^2A(\alpha_2+\alpha_3)^2+\frac{1}{2}\log\alpha_1.$$

Now, for any $((\alpha_1,\alpha_2,\alpha_3),s,\nu)\in\mathcal{S}'$, using the fact that $\psi(x)\leq A/2$ for any $x\in\mathbb{R}$, we have

$$\mathcal{F}((\alpha_1,\alpha_2,\alpha_3),s,\nu)\leq\theta^2(\alpha_1+2\alpha_1\alpha_2)+\theta^2A\alpha_2^2+2\theta^2A\alpha_1\alpha_3+\theta^2B(\alpha_3^2+2\alpha_2\alpha_3)+\frac{1}{2}\log\alpha_1.$$

Therefore, it suffices to prove that for α_2 sufficiently near 1:

$$(A-B)(2\alpha_2\alpha_3+\alpha_3^2)\geq 2(A-1)\alpha_1\alpha_3$$

that is,

$$2(A-1)\alpha_1\leq(A-B)(\alpha_3+2\alpha_2).$$

A sufficient condition for this inequality to be true is that $(A-1)(1-\alpha_2)\leq(A-B)\alpha_2$, which ends the proof of the claim (48). By Lemma 6.4, we know that for $\theta\geq\theta_0$,

$$\sup_{(\underline{\alpha},s,\nu)\in S}\mathcal{F}(\underline{\alpha},s,\nu)=\sup_{\substack{(\underline{\alpha},s,\nu)\in S\\ \alpha_1,\alpha_3\leq C\sqrt{\log\theta}/\theta}}\mathcal{F}(\underline{\alpha},s,\nu).$$

For θ such that

$$1-2C\frac{\sqrt{\log\theta}}{\theta}\geq\frac{A-1}{2A-B-1},$$

we obtain from (48) that

$$\sup_{\substack{(\underline{\alpha},s,\nu)\in S\\ \alpha_1,\alpha_3\leq C\sqrt{\log\theta}/\theta}}\mathcal{F}(\underline{\alpha},s,\nu)\leq\sup_{((\alpha_1,\alpha_2,0),s,\nu)\in S}\mathcal{F}((\alpha_1,\alpha_2,0),s,\nu).$$

We deduce that for θ large enough,

$$\sup_{(\underline{\alpha},s,\nu)\in S}\mathcal{F}(\underline{\alpha},s,\nu)\leq\sup_{((\alpha_1,\alpha_2,0),s,\nu)\in S}\mathcal{F}(\alpha_1,\alpha_2,0,s,\nu),$$

which ends the proof. □

We can now conclude from the last two Propositions 6.7 and 6.8, that for $\theta\geq\theta_0$

$$\overline{F}(\theta)\leq\sup_{((\alpha_1,\alpha_2,0),s,\nu)\in S}\mathcal{F}(\alpha_1,\alpha_2,0,s,\nu)=\sup_{\alpha\in[0,1)}V(\alpha)$$

where we optimized over ν (at the centered Gaussian law with covariance α_1). This completes the proof of the proof of Proposition 6.2 with Lemma 6.3.

7. Appendix

7.1. Concentration for Wigner matrices with sub-Gaussian log-concave entries.

Proposition 7.1. [14, 1] *Let μ be a symmetric probability measure on $\mathbb{R}$ which has log-concave tails in the sense that $t \mapsto \mu(x : |x| \geq t)$ is concave, and which is sub-Gaussian in the sense that* (1) *holds. Let X_N be a symmetric random matrix of size N such that $(X_{i,j})_{i\leq j}$ are independent random variables. Assume $\sqrt{N}X_{i,j}$ and $\sqrt{N/2}X_{i,i}$ have law μ for any $i \neq j$. There exists a numerical constant $\kappa > 0$ such that for any convex 1-Lipschitz function $f : \mathbb{R} \to \mathbb{R}$, and $t \geq 0$,*

$$\mathbb{P}\Big(\big|\frac{1}{n}\mathrm{Tr}f(X_N) - \frac{1}{n}\mathbb{E}\mathrm{Tr}f(X_N)\big| > t\Big) \leq 2e^{-\frac{\kappa}{A}N^2t^2}. \tag{49}$$

Moreover, for any $t > 0$,

$$\mathbb{P}\big(|\lambda_{X_N} - \mathbb{E}\lambda_{X_N}| > t\big) \leq 2e^{-\frac{\kappa}{A}Nt^2}. \tag{50}$$

One can take $\kappa = 1/8\beta^2$ with $\beta = 1680e$.

From these concentration inequalities, one can deduce that a Wigner matrix with entries having sub-Gaussian and log-concave laws satisfy Assumptions 1.2.

Corollary 7.2. *Assume μ satisfies the assumptions of Proposition 7.1 and has variance 1. Then the matrix X_N satisfies the Assumptions 1.2.*

Proof. Using the concentration inequality (50) and the convergence in expectation of the spectral radius of X_N (see [1, Theorem 2.1.22, Exercise 2.1.7]), we obtain that the spectral radius of X_N is exponentially tight at the scale N in the sense of (3).

Let $\mathcal{K} \subset \mathbb{R}$ be a compact subset of $\mathbb{R}$ and denote by $|\mathcal{K}|$ its diameter. Let $\mathcal{F}_{\mathrm{Lip},\mathcal{K}}$ be the set of 1-Lipschitz functions with support in $\mathcal{K}$. From the concentration inequality (49), we can deduce by arguing as in the proof of [14, Theorem 1.3] that for any $\delta > 0$,

$$\mathbb{P}\Big(\sup_{f\in\mathcal{F}_{\mathrm{Lip},\mathcal{K}}}\big|\frac{1}{N}\mathrm{Tr}f(X_N) - \mathbb{E}\frac{1}{N}\mathrm{Tr}f(X_N)\big| > \delta\Big) \leq C\frac{|\mathcal{K}|}{\delta}\exp\Big(-C\frac{\delta^2N^2}{|\mathcal{K}|^2}\Big). \tag{51}$$

Let $d_{\mathcal{K}}$ be defined by

$$\forall \nu,\mu \in \mathcal{P}(\mathbb{R}),\ d_{\mathcal{K}}(\mu,\nu) = \sup_{f\in\mathcal{F}_{\mathrm{Lip},\mathcal{K}}}\Big|\int f d\nu - \int f d\mu\Big|,$$

and let d denote the bounded-Lipschitz distance defined by

$$\forall \nu,\mu \in \mathcal{P}(\mathbb{R}),\ d(\mu,\nu) = \sup_{f\in\mathcal{F}_{LU}}\Big|\int f d\nu - \int f d\mu\Big|,$$

where $\mathcal{F}_{LU}$ is the set of 1-Lipschitz functions uniformly bounded by 1. Let fix $\mathcal{K} = [-2,2]$. Since $\sigma(\mathcal{K}) = 1$, we have

$$d(\mu_{X_N},\sigma) \leq d_{\mathcal{K}}(\mu_{X_N},\sigma) + \mu_{X_N}(\mathcal{K}^c).$$

But $\mu_{X_N}(\mathcal{K}^c) \leq d_{\mathcal{K}}(\mu_{X_N},\sigma)$ by choosing $f = \mathbb{1}_{\mathcal{K}}$. Therefore,

$$d(\mu_{X_N},\sigma) \leq 2d_{\mathcal{K}}(\mu_{X_N},\sigma).$$

From (51), we have that for any $\delta > 0$,

$$\mathbb{P}\big(d(\mu_{X_N}, \sigma) > \delta + 2d_{\mathcal{K}}(\mathbb{E}\mu_{X_N}, \sigma)\big) \leq \frac{4C}{\delta} e^{-C\frac{\delta^2 N^2}{16}}. \tag{52}$$

But, we know by [3, Theorem 4.1] that,

$$d_{KS}(\mathbb{E}\mu_{X_N}, \sigma) = O(N^{-1/4}),$$

where d_{KS} denotes the Kolmogorov-Smirnov distance, defined by

$$\forall \nu, \mu \in \mathcal{P}(\mathbb{R}),\ d_{KS}(\mu, \nu) = \sup\big\{\big|\mu\big((-\infty, x]\big) - \nu\big((-\infty, x]\big)\big|, x \in \mathbb{R}\big\}.$$

It is a standard fact that the Kolmogorov-Smirnov distance controls the integrals with respect to functions with bounded variations, that is, for any $\nu, \mu \in \mathcal{P}(\mathbb{R})$,

$$d_{KS}(\mu, \nu) = \sup\big\{ \int f d\mu - \int f d\nu : ||f||_{BV} \leq 1\big\},$$

with $||f||_{BV} = \sup \sum_{i=1}^{m} |f(x_{i+1}) - f(x_i)|$, where the supremum holds over all families $(x_i)_{1 \leq i \leq m+1}$ with $m \in \mathbb{N}$ such that $x_1 < x_2 < \ldots < x_{m+1}$. Since for any 1-Lipschitz function f supported on $\mathcal{K}$, $||f||_{BV} = O(1)$, we deduce that

$$d_{\mathcal{K}}(\mathbb{E}\mu_{X_N}, \sigma) = O(N^{-1/4}).$$

From (52), we deduce that

$$\lim_{N \to +\infty} \frac{1}{N} \log \mathbb{P}\big(d(\mu_{X_N}, \sigma) > N^{-\kappa}\big) = -\infty,$$

for any $\kappa < 1/4$. □

We now give a proof of Proposition 7.1. It will be a direct consequence of Klein's lemma (see [1, Lemma 4.4.12]) and the following concentration of convex Lipschitz functions under μ^n.

Proposition 7.3. *Let μ be a symmetric probability measure on $\mathbb{R}$ which has log-concave tails in the sense that $t \mapsto \mu(x : |x| \geq t)$ is concave, and which is sub-Gaussian in the sense that* (1) *holds. For any lower-bounded convex 1-Lipschitz function $f : \mathbb{R} \to \mathbb{R}$ such that $\int f d\mu^n = 0$ and any $t > 0$,*

$$\mu^n\big(x : |f(x)| > t\big) \leq 2e^{-\frac{t^2}{4\beta^2 A}},$$

where β is numerical constant. One can take $\beta = 1680e$.

Proof. By [17, Corollary 2.2], we know that there exists a numerical constant β such that μ^n satisfies a convex infimum convolution inequality with cost function $\Lambda^*(./\beta)$, where Λ^* is the Legendre transfom of Λ defined by,

$$\forall \theta \in \mathbb{R}^n,\ \Lambda(\theta) = \log \int e^{\langle \theta, x \rangle} d\mu^n(x).$$

Moreover, β can be taken to be $1680e$. More precisely, for any convex lower-bounded function $f : \mathbb{R} \to \mathbb{R}$,

$$\Big(\int e^{f \square \Lambda^*(./\beta)} d\mu^n \Big) \Big(\int e^{-f} d\mu^n \Big) \leq 1, \tag{53}$$

where $\square$ denotes the infimum convolution operator, defined by

$$f\square\Lambda^*(./\beta)(x) = \inf_{y\in\mathbb{R}^n}\big\{f(y)+\Lambda^*\Big(\frac{y-x}{\beta}\Big)\big\}.$$

Since μ is sub-Gaussian in the sense of (1), for any $x\in\mathbb{R}^n$,

$$\Lambda^*(x)\geq \frac{1}{2A}||x||^2,$$

where $||\ ||$ denotes the Euclidean norm in $\mathbb{R}^n$. Therefore,

$$f\square\Lambda^*(./\beta)(x)\geq \inf_{y\in\mathbb{R}^n}\big\{f(y)+\frac{1}{2\beta^2A}||y-x||^2\big\}.$$

Assume f is L-Lipschitz for some $L>0$. Reproducing the arguments of [15, section 1.9, p19] we have for any $x\in\mathbb{R}^n$,

$$\begin{aligned} f\square\Lambda^*(./\beta)(x) &\geq f(x)+\inf_{y\in\mathbb{R}^n}\big\{-L||y-x||+\frac{1}{2\beta^2A}||y-x||^2\big\}\\ &\geq f(x)-\frac{1}{2}\beta^2AL^2. \end{aligned}$$

Thus, by (53) we deduce that

$$\Big(\int e^f d\mu^n\Big)\Big(\int e^{-f}d\mu^n\Big)\leq e^{\frac{1}{2}\beta^2AL^2}. \tag{54}$$

Assume now that f is 1-Lipschitz and $\int f d\mu^n=0$. Using Jensen's inequality, we get for any $\lambda>0$,

$$\int f d\mu^n\leq e^{\frac{1}{2}\beta^2A\lambda^2}.$$

Using Chernoff inequality we obtain that for any $t>0$,

$$\mu^n\big(x: f(x)\geq t\big)\leq e^{-\frac{t^2}{4\beta^2A}}.$$

Using the symmetry in (54) between f and $-f$, we get similarly that for any $t>0$,

$$\mu^n\big(x: f(x)\leq -t\big)\leq e^{-\frac{t^2}{4\beta^2A}},$$

which gives the claim. $\square$

7.2. **A Uniform Varadhan's lemma.**

Lemma 7.4. *Let $f:\mathbb{R}\to\mathbb{R}$ such that $f(0)=0$ and $f(\sqrt{.})$ is L-Lipschitz for some $L>0$. Let M_N, m_N be sequences such that $M_N=o(\sqrt{N})$ and $m_N\sim N$. Let $g_1,\dots,g_{m_N}$ be independent Gaussian random variables conditioned to belong to $[-M_N,M_N]$. Let $\delta\in(0,1)$ and $c>0$ such that $K^{-1}<c<K$ and $2\delta<K^{-1}$. Then,*

$$\begin{aligned}\Big|\frac{1}{N}\log\mathbb{E}e^{\sum_{i=1}^{m_N}f\left(\frac{g_i}{\sqrt{c}}\right)}\mathbb{1}_{|\sum_{i=1}^{m_N}g_i^2-cN|\leq\delta N} - \sup_{\substack{\nu\in\mathcal{P}([-M_N,M_N])\\ \int x^2d\nu=c}}\Big\{\int f\Big(\frac{x}{\sqrt{c}}\Big)d\nu(x)-H(\nu|\gamma)\Big\}\Big|\\ \leq g_{L,K}(N)+h_L(\delta K),\end{aligned}$$

where $g_{L,K}(N)\to+\infty$ as $N\to+\infty$ and $h_L(x)\to 0$ as $x\to 0$.

Proof. Let $\varepsilon = 1/N$ and l_0 be the smallest integer such that $(1+\varepsilon)^{-l_0} \leq \varepsilon$. Define

$$I_{l_0} = [-(1+\varepsilon)^{-l_0}, (1+\varepsilon)^{-l_0}], \text{ and } B_{l_0} = \{i : g_i \in I_{l_0}\}.$$

For any $k > -l_0$, we set

$$I_k = \{x \in \mathbb{R} : (1+\varepsilon)^{k-1} \leq |x| \leq (1+\varepsilon)^k\} \text{ and } B_k = \big\{i : g_i \in I_k\big\}. \tag{55}$$

Let $\mu_k = |B_k|/m_N$. Let k_0 be the smallest integer such that $(1+\varepsilon)^{k_0} \geq M_N$. Since $g_i \in [-M_N, M_N]$ for all i, we obtain that for any $k > k_0$, $B_k = \emptyset$.

Lemma 7.5. *On the event* $\big\{|\sum_{i=1}^{m_N} g_i^2 - cN| \leq \delta N\big\}$,

$$\Big|\frac{1}{m_N}\sum_{i=1}^{m_N} f\Big(\frac{g_i}{\sqrt{c}}\Big) - \sum_{k=-l_0}^{k_0} \mu_k f\Big(\frac{(1+\varepsilon)^k}{\sqrt{c}}\Big)\Big| \leq g_{L,K}(N),$$

where $g_{L,K}$ *is a function such that* $g_{L,K}(x) \to 0$ *as* $x \to +\infty$.

Proof. As $f(\sqrt{.})$ is L-Lipschitz, we have

$$\begin{aligned}
\Big|\frac{1}{m_N}\sum_{i=1}^{m_N} f\Big(\frac{g_i}{\sqrt{c}}\Big) - \sum_{k=-l_0}^{k_0} \mu_k f\Big(\frac{(1+\varepsilon)^k}{\sqrt{c}}\Big)\Big| &\leq \frac{1}{m_N}\sum_{k=-l_0}^{k_0}\sum_{i\in B_k}\Big|f\Big(\frac{g_i}{\sqrt{c}}\Big) - f\Big(\frac{(1+\varepsilon)^k}{\sqrt{c}}\Big)\Big| \\
&\leq \frac{L}{c}\sum_{k=-l_0+1}^{k_0} \mu_k (1+\varepsilon)^{2k}\big(1-(1+\varepsilon)^{-2}\big) \\
&\quad + \frac{L}{c}\mu_{-l_0}(1+\varepsilon)^{-2l_0}.
\end{aligned}$$

Using the fact that $(1+\varepsilon)^{-l_0} \leq \varepsilon$, we deduce

$$\Big|\frac{1}{m_N}\sum_{i=1}^{m_N} f\Big(\frac{g_i}{\sqrt{c}}\Big) - \sum_{k=-l_0}^{k_0} \mu_k f\Big(\frac{(1+\varepsilon)^k}{\sqrt{c}}\Big)\Big| \leq \frac{3\varepsilon L}{c}\Big(\sum_{k=-l_0+1}^{k_0} \mu_k (1+\varepsilon)^{2k} + 1\Big)$$

But, on the other hand

$$\frac{1}{m_N}\sum_{i=1}^{m_N} g_i^2 \geq \sum_{k=-l_0+1}^{k_0} \mu_k (1+\varepsilon)^{2(k-1)}.$$

Thus,

$$\Big|\frac{1}{m_N}\sum_{i=1}^{m_N} f\Big(\frac{g_i}{\sqrt{c}}\Big) - \sum_{k=-l_0}^{k_0} \mu_k f\Big(\frac{(1+\varepsilon)^k}{\sqrt{c}}\Big)\Big| \leq \frac{3\varepsilon L}{c}\Big((1+\varepsilon)^2\big(\frac{1}{m_N}\sum_{i=1}^{m_N} g_i^2\big) + 1\Big),$$

which, as $m_N \sim N$ and $K^{-1} < c < K$, gives the claim. □

Let $I = \{-l_0, \dots, k_0\}$ and $\mathcal{L}_N$ be the set,

$$\mathcal{L}_N = \Big\{y \in \mathbb{R}_+^I : \sum_{k\in I} y_k = 1,\ \forall k \in I, m_N y_k \in \mathbb{N}\Big\}.$$

We know from [9, Lemma 2.1.6], that for any $y \in \mathcal{L}_N$,

$$(m_N+1)^{-n} e^{-m_N H(y|\gamma_{m_N})} \leq \mathbb{P}(\mu_k = y_k,\ \forall k \in I) \leq e^{-m_N H(y|\gamma_{m_N})}, \tag{56}$$

where $n = |I| = l_0 + k_0 + 1$,

$$H(y|\gamma_{m_N}) = \sum_{k\in I} y_k \log \frac{y_k}{\gamma_{m_N}(k)}, \quad \text{with } \gamma_{m_N}(k) = \frac{\gamma(I_k)}{\gamma([-M_N, M_N])},$$

and $I_k = [(1+\varepsilon)^{k-1}, (1+\varepsilon)^k]$.

Let $\mu = (\mu_k)_{k\in I}$. There exists a function h_K (depending on c but not δ) such that for

$$\begin{aligned}\big\{c - \delta - h_K(\varepsilon) \le \sum_{k=-l_0}^{k_0} (1+\varepsilon)^{2k}\mu_k \le c + \delta\big\} \qquad (57)\\ \subset \big\{|\sum_{i=1}^{m_N} g_i^2 - cN| \le \delta N\big\} \subset \{c - \delta \le \sum_{k=-l_0}^{k_0} (1+\varepsilon)^{2k}\mu_k \le c + \delta + h_K(\varepsilon)\},\end{aligned}$$

and $h_K(\varepsilon) \to 0$ as $\varepsilon \to 0$. We introduce the sets

$$\mathcal{A}_+ = \big\{y \in \mathcal{L}_N : c - \delta \le \sum_{k=-l_0}^{k_0} (1+\varepsilon)^{2k} y_k \le c + \delta + h_K(\varepsilon)\big\},$$

and

$$\mathcal{A}_- = \big\{y \in \mathcal{L}_N : c - \delta - h_K(\varepsilon) \le \sum_{k=-l_0}^{k_0} (1+\varepsilon)^{2k} y_k \le c + \delta\big\}.$$

Using (56), we have for the upper bound,

$$\mathbb{E} e^{m_N \sum_{k=-l_0}^{k_0} \mu_k f\left(\frac{(1+\varepsilon)^k}{\sqrt{c}}\right)} \mathbb{1}_{\mu\in\mathcal{A}_+} \le \sum_{y\in\mathcal{A}_+} e^{m_N \sum_{k=-l_0}^{k_0} y_k f\left(\frac{(1+\varepsilon)^k}{\sqrt{c}}\right)} e^{-m_N H(y|\gamma_{m_N})}, \qquad (58)$$

whereas for the lower bound,

$$\mathbb{E} e^{m_N \sum_{k=-l_0}^{k_0} \mu_k f\left(\frac{(1+\varepsilon)^k}{\sqrt{c}}\right)} \mathbb{1}_{\mu\in\mathcal{A}_-} \ge (m_N+1)^{-n} \sum_{y\in\mathcal{A}_-} e^{m_N \sum_{k=-l_0}^{k_0} y_k f\left(\frac{(1+\varepsilon)^k}{\sqrt{c}}\right)} e^{-m_N H(y|\gamma_{m_N})}. \qquad (59)$$

Let $y \in \mathcal{A}_+$ and define $\nu \in \mathcal{P}(\mathbb{R})$ by $d\nu(x) = \varphi(x)d\gamma(x)$, where

$$\varphi(x) = \sum_{k=-l_0}^{k_0} \mathbb{1}_{x\in I_k} \frac{y_k}{\gamma(I_k)}.$$

With this notation, we have

$$H(y|\gamma_{m_N}) = H(\nu|\gamma) - \log \gamma([-M_N, M_N]).$$

With the same argument as in Lemma 7.5, we also have as $y \in \mathcal{A}_+$,

$$\Big|\sum_{k=-l_0}^{k_0} y_k f\Big(\frac{(1+\varepsilon)^k}{\sqrt{c}}\Big) - \int f\Big(\frac{x}{\sqrt{c}}\Big) d\nu(x)\Big| \le g_{L,K}(N), \qquad (60)$$

and

$$c - \delta - m_K(N) \le \int x^2 d\nu(x) \le c + \delta + m_K(N),$$

where $g_{L,K}(x)$ and $m_K(x) \to 0$ as $x \to +\infty$. From (58) and Lemma 7.5, we deduce that

$$\frac{1}{N}\log \mathbb{E} e^{\sum_{i=1}^{m_N} f\left(\frac{g_i}{\sqrt{c}}\right)} \mathbb{1}_{|\sum_{i=1}^{m_N} g_i^2 - cN| \leq \delta N} \leq \frac{m_N}{N} \sup_{\substack{\nu \in \mathcal{P}([-M_N, M_N] \\ |\int x^2 d\nu(x) - c| \leq \delta + m_K(N)}} \Big\{ \int f\Big(\frac{x}{\sqrt{c}}\Big) d\nu(x) - H(\nu|\gamma) \Big\} + \frac{1}{N}\log|\mathcal{L}_N| + r_{L,K}(N),$$

where $r_{L,K}(N) \to 0$ as $N \to +\infty$. But, $|\mathcal{L}_N| \leq m_N^n$, and $n = |I| \leq C \log N$, where $C > 0$ is a numerical constant, so that

$$\frac{1}{N}\log|\mathcal{L}_N| \leq C\frac{(\log N)^2}{N}.$$

To complete the proof of the upper bound, we show the following result.

Lemma 7.6. *Let $K, L, \delta > 0$ such that $\delta < 2K^{-1}$. There exists a function $s_{L,K}$ depending on K and L such that for any function $f : \mathbb{R} \to \mathbb{R}$ such that $f(0) = 0$ and $f(\sqrt{\cdot})$ is L-Lipschitz, and any $K^{-1} < c < K$,*

$$\sup_{|\int x^2 d\nu(x) - c| \leq \delta} \Big\{ \int f\Big(\frac{x}{\sqrt{c}}\Big) d\nu(x) - H(\nu|\gamma) \Big\} \leq \sup_{\substack{\nu \in \mathcal{P}([-M_N, M_N] \\ \int x^2 d\nu(x) = c}} \Big\{ \int f\Big(\frac{x}{\sqrt{c}}\Big) d\nu(x) - H(\nu|\gamma) \Big\} + s_L(\delta K),$$

where $s_L(x) \to 0$ as $x \to 0$.

Proof. Let $\nu \in \mathcal{P}([-M_N, M_N])$ such that $|\int x^2 d\nu(x) - c| \leq \delta$. Assume first $\int x^2 d\nu(x) \geq c$. Define $\tilde{\nu} \in \mathcal{P}([-M_N, M_N])$ such that for any bounded continuous function $\varphi : [-M_N, M_N] \to \mathbb{R}$,

$$\int \varphi\Big(\frac{x}{\sqrt{\lambda}}\Big) d\nu(x) = \int \varphi d\tilde{\nu},$$

where $\lambda \geq 1$ is such that $\int x^2 d\tilde{\nu}(x) = c$. Using the fact that $f(\sqrt{\cdot})$ is L-Lipschitz, and that $|\int x^2 d\nu(x) - c| \leq \delta$, we obtain on one hand,

$$\Big| \int f\Big(\frac{x}{\sqrt{\lambda c}}\Big) d\nu(x) - \int f\Big(\frac{x}{\sqrt{c}}\Big) d\nu(x) \Big| \leq \frac{L}{c}\delta \leq L\delta K.$$

On the other hand, we have as $\int x^2 d\tilde{\nu}(x) = c$,

$$H(\tilde{\nu}|\gamma) = H(\tilde{\nu}) + \frac{1}{2}c + \frac{1}{2}\log(2\pi). \tag{61}$$

Since $H(\tilde{\nu}) = H(\nu) + \frac{1}{2}\log\lambda$, we have that

$$H(\tilde{\nu}|\gamma) = H(\nu) + \frac{1}{2}\log\lambda + \frac{1}{2}c + \frac{1}{2}\log(2\pi).$$

But $c \leq \int x^2 d\nu(x) \leq c + \delta$ and $\lambda - 1 \leq \delta K$. Therefore

$$H(\tilde{\nu}|\gamma) \leq H(\nu|\gamma) + \frac{1}{2}\delta + \frac{L}{2}\delta K,$$

which ends the proof in the case where $\int x^2 d\nu(x) \geq c$.

Assume now that $c-\delta \leq \int x^2 d\nu(x) \leq c$. For N large enough, we can find $t \in [0, M_N]$ such that

$$\int x^2 d\big((1-\delta)\nu + \delta \mathbb{1}_{[t,t+1]}dx\big)(x) = c.$$

Let $\tilde{\nu} = (1-\delta)\nu + \delta \mathbb{1}_{[t,t+1]}dx \in \mathcal{P}([-M_N, M_N])$. Observe that the above condition implies that

$$\delta \int_{[t,t+1]} x^2 dx \leq c - (c-\delta)(1-\delta) \leq 2K\delta.$$

Using the fact that $f(x) \leq Lx^2$ for any $x \geq 0$, we deduce that,

$$\Big| \int f\Big(\frac{x}{\sqrt{c}}\Big) d\tilde{\nu}(x) - \int f\Big(\frac{x}{\sqrt{c}}\Big) d\nu(x) \Big| \leq 3L\delta K.$$

On the other hand, by convexity of the entropy and the fact that $H(\mathbb{1}_{[t,t+1]}dx) = 1$, we have,

$$H(\tilde{\nu}) \leq (1-\delta)H(\nu) + \delta.$$

Using (61) and $c - \delta \leq \int x^2 d\nu(x) \leq c$, we deduce that

$$H(\tilde{\nu}|\gamma) \leq H(\nu|\gamma) - \delta H(\nu) + \frac{3\delta}{2}.$$

But,

$$H(\nu) \geq -\frac{1}{2}c - \frac{1}{2}\log(2\pi),$$

so that we can conclude that,

$$H(\tilde{\nu}|\gamma) \leq H(\nu|\gamma) + O(\delta K),$$

which ends the proof.

□

For the lower bound, fix ν a probability measure on $[-M_N, M_N]$ such that $\nu \ll \gamma$. We set $\varepsilon = \varepsilon_N$ such that $M_N^2/m_N\varepsilon_N \to 0$, and we define I_k and B_k as in (55). Define, for $k \in \{-l_0+1, \dots, k_0\}$,

$$y_k = \frac{1}{m_N} \lfloor m_N \nu(I_k) \rfloor,$$

and $y_{-l_0} = 1 - \sum_{k=-l_0+1}^{k_0} y_k$. We claim that for N large enough and independent of ν,

$$\int x^2 d\nu(x) = c \Longrightarrow y \in \mathcal{A}_-.$$

Indeed, one can check that on one hand

$$\sum_{k=-l_0}^{k_0} y_k (1+\varepsilon)^{2k} \leq (1+\varepsilon)^2 \int x^2 d\nu(x) + \varepsilon^2,$$

and on the other hand

$$\sum_{k=-l_0}^{k_0} y_k (1+\varepsilon)^{2k} \geq \int x^2 d\nu(x) - \frac{(1+\varepsilon)^2 M_N^2}{m_N \varepsilon}.$$

We obtain from (59),

$$\log \mathbb{E}e^{m_N \sum_{k=-l_0}^{k_0} \mu_k f\left(\frac{(1+\varepsilon)^k}{\sqrt{c}}\right)} \mathbb{1}_{\mu \in \mathcal{A}_-} \geq (m_N+1)^{-n} e^{m_N \sum_{k=-l_0}^{k_0} y_k f\left(\frac{(1+\varepsilon)^k}{\sqrt{c}}\right)} e^{-m_N H(y|\gamma_{m_N})}. \quad (62)$$

In the next lemma we compare $H(y|\gamma_{m_N})$ and $H(\nu|\gamma)$.

Lemma 7.7.

$$H(y|\gamma_{m_N}) \leq H(\nu|\gamma) + o_N(1).$$

Proof. By definition we have,

$$H(y|\gamma_{m_N}) = \sum_{k=-l_0}^{k_0} y_k \log \frac{y_k}{\gamma(I_k)} + \log \gamma([-M_N, M_N]). \quad (63)$$

Let $f(x) = x \log x$ for $x > 0$ and $f(0) = 0$. We claim that

$$\forall 0 \leq x < y, \ f(x) \leq f(y) + (y - x). \quad (64)$$

Indeed, either $x > e^{-1}$ and $f(x) \leq f(y)$ since f is increasing on $[e^{-1}, +\infty)$. Either $x < e^{-1}$ and by convexity,

$$f(x) \leq f(y) + f'(x)(x - y).$$

Since $|f'(x)| \leq 1$ we get the claim. Note that we have for any $k > -l_0$,

$$\nu(I_k) - \frac{1}{m_N} < y_k \leq \nu(I_k),$$

and

$$\nu(I_{-l_0}) \leq y_{-l_0} < \nu(I_{-l_0}) + \frac{k_0 + l_0}{m_N}.$$

Thus we deduce from (64) that

$$\begin{aligned} H(y|\gamma_{m_N}) &\leq \sum_{k=-l_0}^{k_0} \nu(I_k) \log \frac{\nu(I_k)}{\gamma(I_k)} + \sum_{k=-l_0+1}^{k_0} \gamma(I_k) \frac{1}{\gamma(I_k) m_N} + \gamma(I_{l_0}) \frac{k_0 + l_0}{\gamma(I_{-l_0}) m_N} + o_N(1) \\ &\leq \sum_{k=-l_0}^{k_0} \nu(I_k) \log \frac{\nu(I_k)}{\gamma(I_k)} + \frac{2(k_0 + l_0)}{m_N} + o_N(1). \end{aligned}$$

We have $k_0 = O(\log(M_N)/\varepsilon_N)$ and $l_0 = O(\log(1/\varepsilon_N)/\varepsilon_N)$. Since $M_N^2/m_N\varepsilon_N \to 0$, we get

$$H(y|\gamma_{m_N}) \leq \sum_{k=-l_0}^{k_0} \nu(I_k) \log \frac{\nu(I_k)}{\gamma(I_k)} + \frac{2(k_0 + l_0)}{m_N} + o_N(1).$$

Since $f : x \mapsto x \log x$ is a convex function, we have by Jensen's inequality

$$\sum_{k=-l_0}^{k_0} \nu(I_k) \log \frac{\nu(I_k)}{\gamma(I_k)} = \sum_{k=-l_0}^{k_0} \gamma(I_k) f\Big(\frac{1}{\gamma(I_k)} \int_{I_k} \frac{d\nu}{d\gamma} d\gamma\Big) \leq \sum_{k=-l_0}^{k_0} \int_{I_k} \frac{d\nu}{d\gamma} \log \frac{d\nu}{d\gamma} d\gamma,$$

which ends the proof. □

Next, we claim that we can compare $\int f(x/\sqrt{c}) d\nu(x)$ and $\sum_{k=-l_0} y_k f((1+\varepsilon)^k/\sqrt{c})$.

Lemma 7.8.

$$\Big|\int f\Big(\frac{x}{\sqrt{c}}\Big)d\nu(x) - \sum_{k=-l_0}^{k_0} y_k f\Big(\frac{(1+\varepsilon)^k}{\sqrt{c}}\Big)\Big| \le g_{L,K}(N),$$

where $g_{L,K}(N) \to 0$ *as* $N \to +\infty$.

Proof. As $f(\sqrt{.})$ is L-Lipschitz, we have on one hand using the same argument as in the proof of Lemma 7.5,

$$\Big|\int f\Big(\frac{x}{\sqrt{c}}\Big)d\nu(x) - \sum_{k=-l_0}^{k_0} \nu(I_k) f\Big(\frac{(1+\varepsilon)^k}{\sqrt{c}}\Big)\Big| \le \frac{3L\varepsilon}{c}\Big(\int x^2 d\nu(x) + 1\Big).$$

Therefore,

$$\Big|\int f\Big(\frac{x}{\sqrt{c}}\Big)d\nu(x) - \sum_{k=-l_0}^{k_0} \nu(I_k) f\Big(\frac{(1+\varepsilon)^k}{\sqrt{c}}\Big)\Big| \le h_{L,K}(N), \tag{65}$$

where $h_{L,K}(N) \to 0$ as $N \to +\infty$. On the other hand, as $|\nu(I_k) - y_k| \le 1/m_N$ for any $k > -l_0$ and $|\nu(I_{-l_0}) - y_{-l_0}| \le (k_0 + l_0)/m_N$, we get

$$\sum_{k=-l_0}^{k_0} \Big|(y_k - \nu(I_k)) f\Big(\frac{(1+\varepsilon)^k}{\sqrt{c}}\Big)\Big| \le \frac{L}{cm_N} \sum_{k=-l_0+1}^{k_0} (1+\varepsilon)^{2k} + \frac{L(k_0+l_0)}{cm_N}(1+\varepsilon)^{-2l_0},$$

where we used the fact that $f(0) = 0$ and $f(\sqrt{.})$ is L-Lipschitz. There exists a numerical constant $\kappa > 0$ such that

$$\sum_{k=-l_0}^{k_0} |y_k - \nu(I_k)| f\Big(\frac{(1+\varepsilon)^k}{\sqrt{c}}\Big) \le \frac{\kappa L}{cm_N}\Big(\frac{M_N^2}{\varepsilon_N} + (k_0+l_0)\varepsilon_N^2\Big),$$

As $M_N^2/m_N \to 0$ and $k_0 = O(\log(M_N)/\varepsilon_N)$ and $l_0 = O(\log(1/\varepsilon_N)/\varepsilon_N)$, we deduce

$$\frac{(k_0+l_0)}{m_N}\varepsilon_N = o_N(1).$$

Since we choose ε_N such that $M_N^2/m_N\varepsilon_N \to 0$, we can conclude that

$$\sum_{k=-l_0}^{k_0} \Big|(y_k - \nu(I_k)) f\Big(\frac{(1+\varepsilon)^k}{\sqrt{c}}\Big)\Big| \le g_{L,K}(N),$$

where $g_{L,K}(N) \to 0$ as $N \to +\infty$. Combining the above estimate with (60), we get the claim. □

Coming back to (62), using the results of Lemmas 7.7 and 7.8, we deduce

$$\mathbb{E}e^{m_N \sum_{k=-l_0}^{k_0} \mu_k f\left(\frac{(1+\varepsilon)^k}{\sqrt{c}}\right)} \mathbb{1}_{\mu\in\mathcal{A}_-} \ge (m_N+1)^{-n} e^{m_N \int f(\frac{x}{\sqrt{c}})d\nu(x) - m_N g_{L,\delta}(N)} e^{-m_N(H(\nu|\gamma)+o_N(1))},$$

which gives at the logarithmic scale,

$$\frac{1}{N}\log \mathbb{E}e^{m_N \sum_{k=-l_0}^{k_0} \mu_k f\left(\frac{(1+\varepsilon)^k}{\sqrt{c}}\right)} \mathbb{1}_{\mu\in\mathcal{A}_-} \ge \frac{m_N}{N}\Big(\int f\Big(\frac{x}{\sqrt{c}}\Big)d\nu(x) - H(\nu|\gamma)\Big) - \tilde{g}_{L,K}(N),$$

where $\tilde{g}_{L,K}(N) \to 0$ as $N \to +\infty$. As this inequality is true for any $\nu \in \mathcal{P}([-M_N, M_N])$, such that $\int x^2 d\nu = c$,

$$\frac{1}{N} \log \mathbb{E} e^{m_N \sum_{k=-l_0}^{k_0} \mu_k f\left(\frac{(1+\varepsilon)^k}{\sqrt{c}}\right)} \mathbb{1}_{\mu \in \mathcal{A}_-} \geq \frac{m_N}{N} \sup_{\substack{\nu \in \mathcal{P}([-M_N, M_N]) \\ \int x^2 d\nu = c}} \Big(\int f\Big(\frac{x}{\sqrt{c}}\Big) d\nu(x) - H(\nu|\gamma) \Big) - \tilde{g}_{L,K}(N).$$

□

References

[1] G. Anderson, A. Guionnet, and O. Zeitouni. *An introduction to random matrices*, volume 118 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010.

[2] F. Augeri. Large deviations principle for the largest eigenvalue of Wigner matrices without Gaussian tails. *Electron. J. Probab.*, 21:Paper No. 32, 49, 2016.

[3] Z. D. Bai. Convergence rate of expected spectral distributions of large random matrices. II. Sample covariance matrices. *Ann. Probab.*, 21(2):649–672, 1993.

[4] Z. D. Bai and Y. Q. Yin. Necessary and sufficient conditions for almost sure convergence of the largest eigenvalue of a Wigner matrix. *Ann. Probab.*, 16(4):1729–1741, 1988.

[5] G. Ben Arous, A. Dembo, and A. Guionnet. Aging of spherical spin glasses. *Probab. Theory Related Fields*, 120(1):1–67, 2001.

[6] G. Ben Arous and A. Guionnet. Large deviations for Wigner's law and Voiculescu's non-commutative entropy. *Probab. Theory Related Fields*, 108(4):517–542, 1997.

[7] P. Bianchi, M. Debbah, M. Maida, and J. Najim. Performance of statistical tests for single-source detection using random matrix theory. *IEEE Trans. Inform. Theory*, 57(4):2400–2419, 2011.

[8] C. Bordenave and P. Caputo. A large deviation principle for Wigner matrices without Gaussian tails. *Ann. Probab.*, 42(6):2454–2496, 2014.

[9] A. Dembo and O. Zeitouni. *Large deviations techniques and applications*, volume 38 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2010. Corrected reprint of the second (1998) edition.

[10] A. Fey, R. van der Hofstad, and M. Klok. Large deviations for eigenvalues of sample covariance matrices, with applications to mobile communication systems. *Adv. in Appl. Probab.*, 40(4):1048–1071, 2008.

[11] Z. Fűredi and J. Komlós. The eigenvalues of random symmetric matrices. *Combinatorica*, 1(3):233–241, 1981.

[12] A. Guionnet and J. Husson. Large deviations for the largest eigenvalue of Rademacher matrices. *To appear in Annals of Probability*, arXiv:1810.0118, 2019.

[13] A. Guionnet and M. Maïda. A Fourier view on the R-transform and related asymptotics of spherical integrals. *J. Funct. Anal.*, 222(2):435–490, 2005.

[14] A. Guionnet and O. Zeitouni. Concentration of the spectral measure for large matrices. *Electron. Comm. Probab.*, 5:119–136, 2000.

[15] M. Ledoux. *The concentration of measure phenomenon*, volume 89 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2001.

[16] M. Maïda. Large deviations for the largest eigenvalue of rank one deformations of Gaussian ensembles. *Electron. J. Probab.*, 12:1131–1150, 2007.

[17] Marta Strzelecka, MichałStrzelecki, and Tomasz Tkocz. On the convex infimum convolution inequality with optimal cost function. *ALEA Lat. Am. J. Probab. Math. Stat.*, 14(2):903–915, 2017.

[18] D. Voiculescu. The analogues of entropy and of Fisher's information measure in free probability theory. V. Noncommutative Hilbert transforms. *Invent. Math.*, 132(1):189–227, 1998.

[19] E. P. Wigner. On the distribution of the roots of certain symmetric matrices. *Annals Math.*, 67:325–327, 1958.

(Fanny Augeri) Weizmann Institute of Science
E-mail address: fanny.augeri@weizmann.ac.il

(Alice Guionnet) Université de Lyon, ENSL, CNRS, France
E-mail address: Alice.Guionnet@umpa.ens-lyon.fr

(Jonathan Husson) Université de Lyon, ENSL, CNRS, France
E-mail address: Jonathan.Husson@umpa.ens-lyon.fr