

Sanov-type large deviations in Schatten classes

Zakhar Kabluchko^a, Joscha Prochno^b and Christoph Thäle^c

^a*Institut für Mathematische Stochastik, Westfälische Wilhelms-Universität Münster, Germany. E-mail: zakhar.kabluchko@uni-muenster.de*

^b*Institut für Mathematik und Wissenschaftliches Rechnen, Karl-Franzens-Universität Graz, Austria. E-mail: j.prochno@hull.ac.uk*

^c*Fakultät für Mathematik, Ruhr-Universität Bochum, Germany. E-mail: christoph.thaele@rub.de*

Received 16 August 2018; revised 17 January 2019; accepted 28 March 2019

Abstract. Denote by $\lambda_1(A), \dots, \lambda_n(A)$ the eigenvalues of an $(n \times n)$ -matrix A . Let Z_n be an $(n \times n)$ -matrix chosen uniformly at random from the matrix analogue to the classical ℓ_p^n -ball, defined as the set of all self-adjoint $(n \times n)$ -matrices satisfying $\sum_{k=1}^n |\lambda_k(A)|^p \leq 1$. We prove a large deviations principle for the (random) spectral measure of the matrix $n^{1/p} Z_n$. As a consequence, we obtain that the spectral measure of $n^{1/p} Z_n$ converges weakly almost surely to a non-random limiting measure given by the Ullman distribution, as $n \rightarrow \infty$. The corresponding results for random matrices in Schatten trace classes, where eigenvalues are replaced by the singular values, are also presented.

Résumé. Notons $\lambda_1(A), \dots, \lambda_n(A)$ les valeurs propres d'une matrice A de taille $n \times n$. Soit Z_n une matrice $n \times n$ choisie aléatoirement et uniformément dans l'équivalent matriciel de la boule unité de l'ensemble classique ℓ_p^n , défini comme l'ensemble des matrices $n \times n$ auto-adjointes satisfaisant $\sum_{k=1}^n |\lambda_k(A)|^p \leq 1$. Nous prouvons un principe de grandes déviations pour la mesure spectrale aléatoire de la matrice $n^{1/p} Z_n$. Comme conséquence, nous obtenons que la mesure spectrale de $n^{1/p} Z_n$ converge faiblement presque sûrement vers une mesure déterministe limite décrite par la loi d'Ullman lorsque n tend vers l'infini. Nous présentons également les résultats correspondants pour les matrices aléatoires dans les classes de trace de Schatten, où les valeurs propres sont remplacées par les valeurs singulières.

MSC: Primary 47B10; 60B20; 60F10; secondary 46B07; 47B10; 52A21; 52A23

Keywords: Asymptotic geometric analysis; Coulomb gas; Eigenvalues; Gaussian ensembles; High dimensional convexity; Large deviations principles; Matrix unit balls; Random matrix theory; Schatten classes; Singular values

1. Introduction and main results

1.1. General introduction

The systematic study of high-dimensional convex bodies is one of the central aspects in theory of Asymptotic Geometric Analysis. Probably the most prominent example are the unit balls of the classical ℓ_p^n -spaces. Their geometry as well as their analytic and probabilistic aspects are well understood by now and we refer the reader to the research monographs and surveys [5, 8, 13, 14] and the references therein.

In the local theory of Banach spaces and today in Asymptotic Geometric Analysis there has ever since been a particular interest and focus on the non-commutative settings. These are represented on the one hand by the Schatten trace classes S_p ($0 < p \leq +\infty$), consisting of all compact linear operators on a Hilbert space for which the sequence of their singular values belongs to the sequence space ℓ_p^n , and on the other hand by their self-adjoint subclasses, where the sequence of eigenvalues values belongs to ℓ_p^n , which in dimension n are formed by the so-called classical matrix ensembles $\mathcal{H}_n(\mathbb{R})$ (Gaussian orthogonal ensemble = GOE), $\mathcal{H}_n(\mathbb{C})$ (Gaussian unitary ensemble = GUE) and $\mathcal{H}_n(\mathbb{H})$ (Gaussian symplectic ensemble = GSE). For example, it was Gordon and Lewis [12] who obtained that the space S_1 does not have local unconditional structure, Tomczak-Jaegermann [34] demonstrated that this space (naturally identified with the projective tensor product $\ell_2 \otimes_\pi \ell_2$) has Rademacher cotype 2, Szarek and Tomczak-Jaegermann [33] provided bounds for the volume ratio of S_1^n , and König, Meyer and Pajor [27] proved the boundedness of the isotropic constants of S_p^n ($1 \leq p \leq +\infty$).

More recently, Guédon and Paouris [15] have established concentration of mass properties for the unit balls of Schatten p -classes and classical matrix ensembles, Barthe and Cordero-Erausquin [6] studied variance estimates, Chávez-Domínguez and Kutzarova determined the Gelfand widths of certain identity mappings between finite-dimensional trace classes \mathcal{S}_p , Radke and Vritsiou [29] and Vritsiou [36] proved the thin-shell conjecture and the variance conjecture for the operator norm, respectively, Hinrichs, Prochno and Vybíral [19] computed the entropy numbers for natural embeddings of \mathcal{S}_p^n in all possible regimes, and Kabluchko, Prochno and Thäle [20,21] obtained the precise asymptotic volumes of the unit balls in classical matrix ensembles and Schatten classes, studied volumes of intersections (in the spirit of Schechtman and Schmuckenschläger [31]) and determined the exact asymptotic volume ratios for \mathcal{S}_p^n ($0 < p \leq +\infty$). It can be seen from all this work referenced above that while those matrix spaces often show a certain similarity to the commutative setting of classical ℓ_p^n -spaces, there is a considerable difference in the behavior of certain quantities related to the geometry of Banach spaces. In fact, often other methods and tools are needed and proofs can be considerably more involved.

It was only recently that the probabilistic concept of a large deviations principle (LDP) was considered in Asymptotic Geometric Analysis by Gantert, Kim and Ramanan [11]. Contrary to a central limit theorem the LDP allows one to access non-universal features and unveil properties that distinguish between different convex bodies. In the setting of finite-dimensional ℓ_p^n -spaces, the authors proved an LDP for 1-dimensional projections of random vectors drawn uniformly from the unit ball \mathbb{B}_p^n of ℓ_p^n , demonstrating stark changes in large deviation behavior as the parameter p varies. This result was extended by Alonso-Gutiérrez, Prochno and Thäle [1] to a higher-dimensional setting in the case where projections to random subspaces are considered, showing that the Euclidean norm of the projection of a random vector uniformly distributed in \mathbb{B}_p^n onto a random subspace satisfies an LDP (see also [2] and [22] for complementing results and other distributions). In his recent PhD thesis, Kim [24] was able to extend further the results from [1] and [11] to more general classes of random vectors satisfying an asymptotic thin-shell-type condition in the spirit of [4] (see [24, Assumption 5.1.2]). Among others, this condition is satisfied by random vectors chosen uniformly at random from a (generalized) Orlicz ball. This body of research is complemented by [25], in which Kim and Ramanan obtained a so-called Sanov-type large deviations principle for the empirical measure of an $n^{1/p}$ multiple of the coordinates of a random point drawn from an ℓ_p^n -sphere with respect to the cone or surface measure. The rate function is essentially shown to be the so-called relative entropy perturbed by some p th moment penalty (see [25, Equation (9)]). More precisely (also to allow comparison with the rate function we obtain for non-commutative ℓ_p^n -spaces), they showed that if $X^{(n,p)} = (X_1^{(n,p)}, \dots, X_n^{(n,p)})$ is a random vector distributed according to the cone measure on an ℓ_p^n -sphere ($1 \leq p \leq +\infty$), then the corresponding sequence

$$\frac{1}{n} \sum_{i=1}^n \delta_{n^{1/p} X_i^{(n,p)}}$$

of empirical measures satisfies an LDP (on the space of probability measures on \mathbb{R} equipped with the q -Wasserstein topology for $q < p$) with good rate function given by

$$\mathcal{J}(v) = \begin{cases} H(v \parallel \rho_p) + \frac{1}{p}(1 - m_p(v)), & m_p(v) \leq 1, \\ +\infty, & m_p(v) > 1, \end{cases}$$

where v is a probability measure on \mathbb{R} , $m_p(v) = \int_{\mathbb{R}} |x|^p v(dx)$, ρ_p is the p -generalized Gaussian distribution on \mathbb{R} whose density is given by

$$\frac{\rho_p(dx)}{dx} = \frac{e^{-|x|^p/p}}{2p^{1/p}\Gamma\left(1 + \frac{1}{p}\right)}, \quad (1 \leq p < +\infty), \quad \frac{\rho_{\infty}(dx)}{dx} = \frac{1}{2}\mathbb{1}_{[-1,1]}(x), \quad (p = +\infty),$$

and

$$H(v \parallel \mu) = \begin{cases} \int_{\mathbb{R}} \log f(x) v(dx), & v \ll \mu, f = \frac{dv}{d\mu}, \\ +\infty, & \text{otherwise} \end{cases}$$

is the relative entropy of v with respect to the probability measure μ .

The purpose of the present paper is to leave the geometric setting of classical ℓ_p^n -spaces and study principles of large deviations in the non-commutative framework of self-adjoint and non self-adjoint Schatten p -classes. A large deviations principle for the law of the spectral measure of a Gaussian Wigner matrix has already been obtained by Ben Arous and Guionnet [7, Theorem 1.1 and Theorem 1.3]. A more general large deviations theorem for random measures (including the

case of the empirical eigenvalue distribution of Wishart matrices) has been obtained by Hiai and Petz in [16, Theorem 1] (see also [18]), which followed their preceding ideas from [17], where the empirical eigenvalue distribution of suitably distributed random unitary matrices was shown to satisfy a large deviations principle as the matrix size goes to infinity. In the same spirit we shall prove in this paper Sanov-type large deviations principles for the spectral measure of $n^{1/p}$ multiples of random matrices chosen uniformly (or with respect to the cone measure on the boundary) from the unit balls of self-adjoint and non self-adjoint Schatten p -classes where $0 < p \leq +\infty$ (see Theorems 1.1 and 1.3 for the self-adjoint case as well as Theorem 1.5 for the non self-adjoint case). In the proofs, we roughly follow a classical strategy in large deviations theory (see, e.g., [3, 7, 16, 17]). However, in our case we need to control simultaneously the deviations of the empirical measures *and* their p th moments towards arbitrary small balls in the product topology of the weak topology on the space of probability measures and the standard topology on \mathbb{R} in the self-adjoint or \mathbb{R}_+ in the non self-adjoint set-up, respectively, and then prove exponential tightness. We shall also use a probabilistic representation for random points in the unit balls of classical matrix ensembles obtained recently in [21] (see (4) and (6)) and a non self-adjoint counterpart (see Proposition 8.2). As we shall see, the good rate function governing the LDPs is essentially given by the logarithmic energy (which is remarkably the same as the negative of Voiculescu's free entropy introduced in [35]) and, which is quite interesting, a perturbation by a constant, which is strongly connected to the famous Ullman distribution. In fact, this constant already appeared in our recent works [20, 21], where the precise asymptotic volume of unit balls in classical matrix ensembles and Schatten trace classes were computed using ideas from the theory of logarithmic potentials with external fields. As a consequence of our LDPs, we obtain a law of large numbers and show that the spectral measure converges weakly almost surely to the Ullman distribution, as the dimension tends to infinity (see Corollary 1.4).

1.2. Results for the self-adjoint case

In order to present our main results for the self-adjoint case in detail, let us introduce some notation. Consider $\beta \in \{1, 2, 4\}$ and let $\mathcal{H}_n(\mathbb{F}_\beta)$ be the collection of all self-adjoint $(n \times n)$ -matrices with entries from the skew field \mathbb{F}_β , where $\mathbb{F}_1 = \mathbb{R}$, $\mathbb{F}_2 = \mathbb{C}$ or $\mathbb{F}_4 = \mathbb{H}$, the set of Hamiltonian quaternions. The standard Gaussian distribution on $\mathcal{H}_n(\mathbb{F}_\beta)$ is known as the GOE (Gaussian orthogonal ensemble) if $\beta = 1$, the GUE (Gaussian unitary ensemble) if $\beta = 2$, and the GSE (Gaussian symplectic ensemble) if $\beta = 4$. By $\lambda_1(A), \dots, \lambda_n(A)$ we denote the (real) eigenvalues of a matrix A from $\mathcal{H}_n(\mathbb{F}_\beta)$ and consider the following Schatten-type unit ball, which can be regarded as a matrix analogue to the classical ℓ_p^n -balls and is defined as

$$\mathbb{B}_{p,\beta}^n := \left\{ A \in \mathcal{H}_n(\mathbb{F}_\beta) : \sum_{j=1}^n |\lambda_j(A)|^p \leq 1 \right\}, \quad \beta \in \{1, 2, 4\} \text{ and } 0 < p \leq +\infty.$$

If $p = +\infty$, then the sum above is replaced by $\max\{|\lambda_j(A)| : j = 1, \dots, n\}$. The boundary of the matrix ball $\mathbb{B}_{p,\beta}^n$ is denoted by

$$\mathbb{S}_{p,\beta}^{n-1} := \partial \mathbb{B}_{p,\beta}^n = \left\{ A \in \mathcal{H}_n(\mathbb{F}_\beta) : \sum_{j=1}^n |\lambda_j(A)|^p = 1 \right\}$$

with the same convention if $p = +\infty$. The space $\mathcal{H}_n(\mathbb{F}_\beta)$ admits a natural scalar product $\langle A, B \rangle = \operatorname{Re} \operatorname{Tr}(AB^*)$ so that it becomes a Euclidean space. The corresponding Riemannian volume on $\mathcal{H}_n(\mathbb{F}_\beta)$ is denoted by $\operatorname{vol}_{\beta,n}$ and this measure coincides with the suitably normalized $(\frac{\beta n(n-1)}{2} + \beta n)$ -dimensional Hausdorff measure on $\mathcal{H}_n(\mathbb{F}_\beta)$. We can therefore define the uniform distribution on $\mathbb{B}_{p,\beta}^n$. The cone probability measure on $\mathbb{S}_{p,\beta}^{n-1}$ is defined as follows: the cone measure of a Borel set $K \subseteq \mathbb{S}_{p,\beta}^{n-1}$ is

$$\frac{\operatorname{vol}_{\beta,n}(\bigcup_{\lambda \in [0,1]} \lambda K)}{\operatorname{vol}_{\beta,n}(\mathbb{B}_{p,\beta}^n)}.$$

The main result of this manuscript for the self-adjoint case is the following Sanov-type large deviations principle for random matrices distributed according to the uniform distribution on $\mathbb{B}_{p,\beta}^n$ or the cone measure on $\mathbb{S}_{p,\beta}^{n-1}$. We denote by $\mathcal{M}(\mathbb{R})$ the space of Borel probability measures on \mathbb{R} equipped with the topology of weak convergence. On this topological space we consider the Borel σ -algebra, denoted by $\mathcal{B}(\mathcal{M}(\mathbb{R}))$.

Theorem 1.1. Fix $0 < p < +\infty$ and $\beta \in \{1, 2, 4\}$. For every $n \in \mathbb{N}$, let Z_n be a random matrix chosen according to the uniform distribution on $\mathbb{B}_{p,\beta}^n$ or the cone measure on its boundary $\mathbb{S}_{p,\beta}^{n-1}$. Then the sequence of random probability

measures

$$\mu_n := \frac{1}{n} \sum_{i=1}^n \delta_{n^{1/p} \lambda_i(Z_n)}, \quad n \in \mathbb{N},$$

satisfies an LDP on $\mathcal{M}(\mathbb{R})$ with speed n^2 and good rate function $\mathcal{I} : \mathcal{M}(\mathbb{R}) \rightarrow [0, +\infty]$ defined by

$$\mathcal{I}(\mu) := \begin{cases} -\frac{\beta}{2} \int_{\mathbb{R}} \int_{\mathbb{R}} \log |x - y| \mu(dx) \mu(dy) + \frac{\beta}{2p} \log \left(\frac{\sqrt{\pi} p \Gamma(\frac{p}{2})}{2^p \sqrt{e} \Gamma(\frac{p+1}{2})} \right), & \int_{\mathbb{R}} |x|^p \mu(dx) \leq 1, \\ +\infty, & \int_{\mathbb{R}} |x|^p \mu(dx) > 1. \end{cases} \quad (1)$$

Remark 1.2.

- (i) As we shall see later, the distribution of the eigenvalues of a point chosen uniformly at random in $\mathbb{B}_{p,\beta}^n$ can be related to the 2-dimensional Coulomb gas (whose density is given in Equation (5) below) of n particles at inverse temperature $\beta > 0$ in an external potential $V : t \mapsto |t|^p$ acting on each particle.
- (ii) It can be seen from Equation (7) below that the additive constant to the logarithmic energy is closely linked to the limit of the free energy whose precise value follows from results of potential theory.

In the next theorem, we consider the case $p = +\infty$.

Theorem 1.3. Fix $\beta \in \{1, 2, 4\}$. For every $n \in \mathbb{N}$, let Z_n be a random matrix chosen according to the uniform distribution on $\mathbb{B}_{\infty,\beta}^n$ or the cone measure on its boundary $\mathbb{S}_{\infty,\beta}^{n-1}$. Then the sequence of random probability measures

$$\mu_n := \frac{1}{n} \sum_{i=1}^n \delta_{\lambda_i(Z_n)}, \quad n \in \mathbb{N},$$

satisfies an LDP on $\mathcal{M}(\mathbb{R})$ with speed n^2 and good rate function $\mathcal{I} : \mathcal{M}(\mathbb{R}) \rightarrow [0, +\infty]$ defined by

$$\mathcal{I}(\mu) = \begin{cases} -\frac{\beta}{2} \int_{\mathbb{R}} \int_{\mathbb{R}} \log |x - y| \mu(dx) \mu(dy) - \frac{\beta}{2} \log 2, & \mu([-1, 1]) = 1, \\ +\infty, & \mu([-1, 1]) < 1. \end{cases}$$

As a corollary, we can derive a law of large numbers for μ_n and show that (weakly almost surely) the sequence of empirical measures converges to a non-random limiting distribution given by the Ullman measure (for $p < +\infty$) or arcsine measure (for $p = +\infty$). We denote the weak convergence of probability measures by \xrightarrow{w} .

Corollary 1.4. Fix $0 < p < +\infty$ and $\beta \in \{1, 2, 4\}$. Let $\mu_{\infty}^{(p)}$ be the probability measure on the interval $[-b_p, b_p]$ with the Ullman density $x \mapsto h_p(x/b_p)/b_p$, where

$$h_p(x) := \frac{p}{\pi} \left(\int_{|x|}^1 \frac{t^{p-1}}{\sqrt{t^2 - x^2}} dt \right) \mathbb{1}_{\{|y| \leq 1\}}(x), \quad b_p := \left(\frac{p \sqrt{\pi} \Gamma(\frac{p}{2})}{\Gamma(\frac{p+1}{2})} \right)^{1/p}. \quad (2)$$

(The normalization is chosen so that the p th moment of $\mu_{\infty}^{(p)}$ equals 1.) Then the random measures μ_n defined in Theorem 1.1 satisfy

$$\mathbb{P}[\mu_n \xrightarrow[n \rightarrow \infty]{w} \mu_{\infty}^{(p)}] = 1.$$

The result also holds in the case $p = +\infty$ with $\mu_{\infty}^{(\infty)}$ being the arcsine distribution with Lebesgue density $\frac{1}{\pi} (1 - t^2)^{-1/2}$, $t \in (-1, 1)$.

For example, for $p = 2$ and $p = 1$ the limiting spectral density takes the following form, see [18, pp. 195–196],

$$\frac{\mu_{\infty}^{(1)}(dx)}{dx} = \frac{1}{\pi^2} \log \frac{\pi + \sqrt{\pi^2 - x^2}}{|x|} \mathbb{1}_{(-\pi, \pi)}(x), \quad \frac{\mu_{\infty}^{(2)}(dx)}{dx} = \frac{1}{2\pi} \sqrt{4 - x^2} \mathbb{1}_{(-2, 2)}(x),$$

see Figure 1.

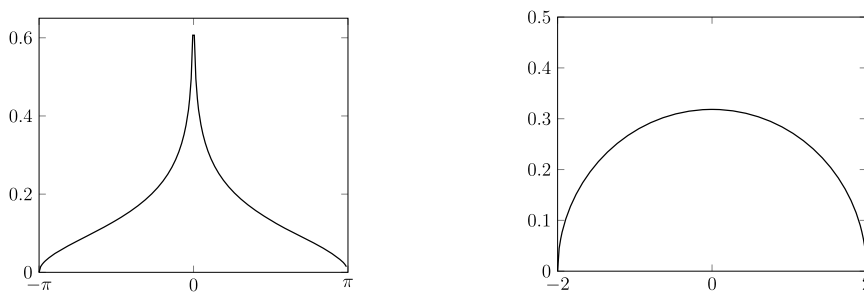


Fig. 1. Plots of the densities $\frac{\mu_\infty^{(1)}(dx)}{dx}$ (left panel) and $\frac{\mu_\infty^{(2)}(dx)}{dx}$ (right panel).

1.3. Results for the non self-adjoint case

After having discussed our main results for the self-adjoint case, we turn now to the non self-adjoint case, where the eigenvalues are replaced by the singular values. For an $(n \times n)$ -matrix $A \in \text{Mat}_n(\mathbb{F}_\beta)$ with entries from the skew field \mathbb{F}_β with $\beta \in \{1, 2, 4\}$ we denote by $s_1(A), \dots, s_n(A)$ the singular values of A . If $\beta = 1$ or $\beta = 2$ these are the eigenvalues of $\sqrt{AA^*}$, while in the Hamiltonian case $\beta = 4$ we refer to [3, Corollary E.13] for a formal definition. For $0 < p \leq +\infty$ the Schatten p -ball is defined as

$$\mathbb{SB}_p^n := \left\{ A \in \text{Mat}_n(\mathbb{F}_\beta) : \sum_{j=1}^n |s_j(A)|^p \leq 1 \right\}$$

with the convention that the sum is replaced by $\max\{|s_j| : j = 1, \dots, n\}$ in the case that $p = +\infty$. As in the self-adjoint case, $\text{Mat}_n(\mathbb{F}_\beta)$ can be supplied with the structure of a Euclidean space in such a way that the βn^2 -dimensional Hausdorff measure restricted to \mathbb{SB}_p^n is finite and can thus be normalized to a probability measure. Moreover, one can also define the cone probability measure on the boundary $\partial \mathbb{SB}_p^n$ of \mathbb{SB}_p^n .

The following Sanov-type LDP is the analogue of Theorem 1.1 and Theorem 1.3 for the non self-adjoint case.

Theorem 1.5. Fix $\beta \in \{1, 2, 4\}$ and $0 < p < +\infty$. For every $n \in \mathbb{N}$, let Z_n be a random matrix chosen uniformly from \mathbb{SB}_p^n or according to the cone probability measure from $\partial \mathbb{SB}_p^n$. Then the sequence of random probability measures

$$\mu_n := \frac{1}{n} \sum_{i=1}^n \delta_{n^{2/p} s_j^2(Z_n)}, \quad n \in \mathbb{N},$$

satisfies an LDP on the space $\mathcal{M}(\mathbb{R}_+)$ of Borel probability measures on \mathbb{R}_+ , endowed with the weak topology, with speed n^2 and good rate function $\mathcal{J} : \mathcal{M}(\mathbb{R}_+) \rightarrow [0, +\infty]$ given by

$$\mathcal{J}(\mu) := \begin{cases} -\frac{\beta}{2} \int_{\mathbb{R}_+} \int_{\mathbb{R}_+} \log |x - y| \mu(dx) \mu(dy) + \frac{\beta}{p} \log \left(\frac{\sqrt{\pi} p \Gamma(\frac{p}{2})}{2^p \sqrt{e} \Gamma(\frac{p+1}{2})} \right), & \int_{\mathbb{R}_+} |x|^{p/2} \mu(dx) \leq 1, \\ +\infty, & \int_{\mathbb{R}_+} |x|^{p/2} \mu(dx) > 1. \end{cases}$$

The result continues to hold in the case $p = +\infty$ if the constant term in the rate function is replaced by its limiting value, as $p \rightarrow \infty$, which is given by $-\frac{\beta}{2} \log 2$.

Again as a corollary, we derive a law of large numbers for the empirical singular-value distribution (and not for the squares of the singular values as in Theorem 1.5). This is the analogue of Corollary 1.4 for the Schatten p -balls.

Corollary 1.6. Fix $\beta \in \{1, 2, 4\}$ and $0 < p < +\infty$. Let $\eta_\infty^{(p)}$ be the probability measure on $[0, b_p]$ with density $x \mapsto 2b_p^{-1} h_p(x/b_p)$ with $h_p(x)$ and b_p given by (2). Further, for each $n \in \mathbb{N}$, let Z_n be uniformly distributed in \mathbb{SB}_p^n or distributed according to the cone probability measure on $\partial \mathbb{SB}_p^n$. Then

$$\mathbb{P} \left[\frac{1}{n} \sum_{j=1}^n \delta_{n^{1/p} s_j(Z_n)} \xrightarrow[n \rightarrow \infty]{w} \eta_\infty^{(p)} \right] = 1.$$

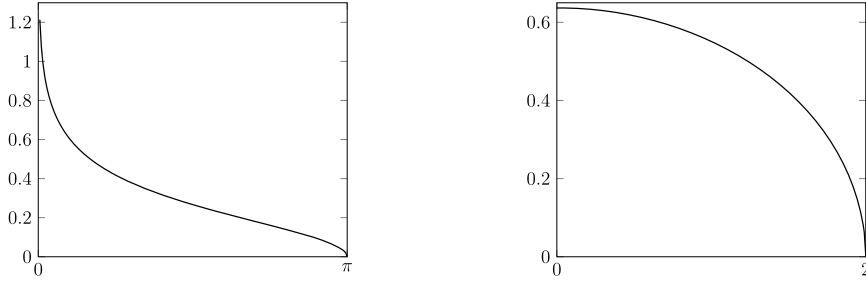


Fig. 2. Plots of the densities $\frac{\eta_\infty^{(1)}(dx)}{dx}$ (left panel) and $\frac{\eta_\infty^{(2)}(dx)}{dx}$ (right panel).

The result also holds in the case $p = +\infty$ with $\eta_\infty^{(\infty)}$ being the absolute arcsine distribution with density $x \mapsto \frac{2}{\pi}(1 - t^2)^{-1/2}$, $t \in (0, 1)$.

For example, if $p = 1$ the limiting distribution has density

$$\frac{d\eta_\infty^{(1)}}{dx}(x) = \frac{2}{\pi^2} \log \frac{\pi + \sqrt{\pi^2 - x^2}}{x} \mathbb{1}_{(0, \pi)}(x),$$

and if $p = 2$ we get the ‘quarter-circle distribution’ with density

$$\frac{d\eta_\infty^{(2)}}{dx}(x) = \frac{1}{\pi} \sqrt{4 - x^2} \mathbb{1}_{(0, 2)}(x),$$

see Figure 2.

1.4. Organization of the paper

The rest of the paper is organized as follows. In Section 2, we introduce the concept of an LDP and some fundamental results used later. In Section 3.1, we present a Schechtman-Zinn type probabilistic representation recently obtained in [21] and which is essential in our argumentation as well. The strategy of the proof is then outlined in Section 3.2. The remaining part of the manuscript is devoted to the proof of Theorem 1.1 with exception of Section 7 in which we prove Corollary 1.4. The proof of Theorem 1.3 is omitted because it follows the same lines as the proof of Theorem 1.1. Since the proofs of Theorem 1.5 and Corollary 1.6 are very similar to the self-adjoint cases, we only sketch the differences in Section 8.

2. Large deviation principles

Keeping a broad readership from both probability theory and geometric analysis in mind, we provide in this section the necessary background material from the theory of large deviations, which may be found in [9, 10, 23], for example. We start directly with the definition of what is understood by a full and a weak large deviations principle. In this paper we fix an underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$, which we implicitly assume to be rich enough to carry all the random objects we consider. Also, for a subset A of a topological (or metric) space we write A° and \bar{A} for the interior and the closure of A , respectively.

Definition 2.1. Let $(\xi_n)_{n \in \mathbb{N}}$ be a sequence of random elements taking values in some metric space M . Further, let $(s_n)_{n \in \mathbb{N}}$ be a positive sequence and $\mathcal{I} : M \rightarrow [0, +\infty]$ be a lower semi-continuous function. We say that $(\xi_n)_{n \in \mathbb{N}}$ satisfies a (full) large deviations principle (LDP) with speed s_n and a rate function \mathcal{I} if

$$-\inf_{x \in A^\circ} \mathcal{I}(x) \leq \liminf_{n \rightarrow \infty} \frac{1}{s_n} \log \mathbb{P}[\xi_n \in A] \leq \limsup_{n \rightarrow \infty} \frac{1}{s_n} \log \mathbb{P}[\xi_n \in A] \leq -\inf_{x \in \bar{A}} \mathcal{I}(x) \quad (3)$$

for all Borel sets $A \subseteq M$. The rate function \mathcal{I} is called good if its level sets $\{x \in M : \mathcal{I}(x) \leq \alpha\}$ are compact for all $\alpha \geq 0$. We say that $(\xi_n)_{n \in \mathbb{N}}$ satisfies a weak LDP with speed s_n and rate function \mathcal{I} if the rightmost upper bound in (3) is valid only for compact sets $A \subseteq M$.

We notice that on the class of all \mathcal{I} -continuity sets, that is, on the class of Borel sets $A \subseteq M$ for which $\mathcal{I}(A^\circ) = \mathcal{I}(\bar{A})$ with $\mathcal{I}(A) := \inf\{\mathcal{I}(x) : x \in A\}$, one has the exact limit relation

$$\lim_{n \rightarrow \infty} \frac{1}{s_n} \log \mathbb{P}[\xi_n \in A] = -\mathcal{I}(A).$$

What separates a weak from a full LDP is the so-called exponential tightness of the sequence of random variables (see, e.g., [9, Lemma 1.2.18] and [23, Lemma 27.9]).

Proposition 2.2. *Let $(\xi_n)_{n \in \mathbb{N}}$ be a sequence of random elements taking values in M . Suppose that it satisfies a weak LDP with speed s_n and rate function \mathcal{I} . Then $(\xi_n)_{n \in \mathbb{N}}$ satisfies a full LDP if and only if the sequence is exponentially tight, that is, if and only if*

$$\inf_K \limsup_{n \rightarrow \infty} \frac{1}{s_n} \log \mathbb{P}[\xi_n \notin K] = -\infty,$$

where the infimum is running over all compact sets $K \subseteq M$. In this case, the rate function \mathcal{I} is good.

The following proposition (see, for instance, [9, Theorem 4.1.11]) shows that it is sufficient to prove a weak LDP for a sequence of random elements for sets in a base of the topology underlying M .

Proposition 2.3. *Let \mathcal{T} be a base of the topology in a metric space M . Let $(\xi_n)_{n \in \mathbb{N}}$ be a sequence of M -valued random elements. If for every $x \in M$,*

$$\mathcal{I}(x) = - \inf_{A \in \mathcal{T}: x \in A} \limsup_{n \rightarrow \infty} \frac{1}{s_n} \log \mathbb{P}[\xi_n \in A] = - \inf_{A \in \mathcal{T}: x \in A} \liminf_{n \rightarrow \infty} \frac{1}{s_n} \log \mathbb{P}[\xi_n \in A],$$

then $(\xi_n)_{n \in \mathbb{N}}$ satisfies a weak LDP with speed s_n and rate function \mathcal{I} .

It will be rather important for us to deduce from an already existing large deviations principle a new one by applying various transformations. We first consider the large deviations behavior under direct products. Let M_1 and M_2 be metric spaces. Assume that $(\xi_n)_{n \in \mathbb{N}}$ is a sequence of M_1 -valued random elements and that $(\eta_n)_{n \in \mathbb{N}}$ is a sequence of M_2 -valued random elements. Assuming that both satisfy LDPs with the same speed, does then also the sequence $((\xi_n, \eta_n))_{n \in \mathbb{N}}$ of $M_1 \times M_2$ -valued random elements satisfy an LDP and, if so, what is its rate function? The following result can be found in [1].

Proposition 2.4. *Assume that $(\xi_n)_{n \in \mathbb{N}}$ satisfies an LDP on a metric space M_1 with speed s_n and good rate function \mathcal{I}_ξ and that $(\eta_n)_{n \in \mathbb{N}}$ satisfies an LDP on a metric space M_2 with the same speed s_n and good rate function \mathcal{I}_η . Then, if ξ_n and η_n are independent for every $n \in \mathbb{N}$, the sequence of pairs $((\xi_n, \eta_n))_{n \in \mathbb{N}}$ satisfies an LDP on $M_1 \times M_2$ with speed s_n and rate function*

$$\mathcal{I}_{(\xi, \eta)}(x) = \mathcal{I}_\xi(x_1) + \mathcal{I}_\eta(x_2), \quad x = (x_1, x_2) \in M_1 \times M_2.$$

Finally, we consider the possibility to ‘transport’ a large deviations principle to another one by means of a continuous function. This device is known as the contraction principle and we refer to [9, Theorem 4.2.1] or [23, Theorem 27.11(i)].

Proposition 2.5. *Let M_1 and M_2 be two metric spaces and let $F : M_1 \rightarrow M_2$ be a continuous function. Further, let $(\xi_n)_{n \in \mathbb{N}}$ be a sequence of M_1 -valued random elements that satisfies an LDP with speed s_n and good rate function \mathcal{I}_ξ . Then the sequence $(F(\xi_n))_{n \in \mathbb{N}}$ of M_2 -valued random elements satisfies an LDP with the same speed and with good rate function $\mathcal{I} = \mathcal{I}_\xi \circ F^{-1}$, i.e.,*

$$\mathcal{I}(y) := \inf\{\mathcal{I}_\xi(x) : x \in M_1, F(x) = y\}, \quad y \in M_2,$$

with the convention that $\mathcal{I}(y) = +\infty$ if $F^{-1}(\{y\}) = \emptyset$.

3. Distributional representation, free energy and strategy of the proof

We present here a Schechtman-Zinn type probabilistic representation for a random matrix chosen uniformly at random from $\mathbb{B}_{p, \beta}^n$ together with some results and concepts from potential theory used later. Also, before we continue with the technical details, we shall briefly explain the strategy of our proof.

3.1. Distributional representation and free energy

Let Z_n be a random matrix uniformly distributed in the ball $\mathbb{B}_{p,\beta}^n$, $\beta \in \{1, 2, 4\}$. The basic fact we rely on is the following distributional representation of the empirical measure of Z_n , see [21, Corollary 4.3]:

$$\mu_n = \frac{1}{n} \sum_{i=1}^n \delta_{n^{1/p} \lambda_i(Z_n)} \stackrel{d}{=} \frac{1}{n} \sum_{i=1}^n \delta_{n^{1/p} U^{1/\ell} \frac{X_{i,n}}{\|X_n\|_p}}, \quad (4)$$

where $\stackrel{d}{=}$ denotes equality in distribution and

(a) $X_n = (X_{1,n}, \dots, X_{n,n})$ is a random vector on \mathbb{R}^n with joint Lebesgue density of the form

$$\frac{1}{C_{n,\beta,p}} e^{-n \sum_{i=1}^n |x_i|^p} \prod_{1 \leq i < j \leq n} |x_i - x_j|^\beta, \quad (x_1, \dots, x_n) \in \mathbb{R}^n \quad (5)$$

with a suitable normalization constant $C_{n,\beta,p} > 0$ depending on n , β and p ,

(b) U is a random variable with uniform distribution on $[0, 1]$ that is independent of X_n ,

(c) $\ell = \ell(n, \beta) = \frac{\beta n(n-1)}{2} + \beta n$ is the (real) dimension of $\mathcal{H}_n(\mathbb{F}_\beta)$.

The above representation can be seen as a non-commutative counterpart to the representation of the uniform distribution on the ℓ_p^n -ball found by Schechtman and Zinn [32]. Similarly, if Z_n is distributed according to the cone measure on $\mathbb{S}_{p,\beta}^{n-1}$, then

$$\mu_n = \frac{1}{n} \sum_{i=1}^n \delta_{n^{1/p} \lambda_i(Z_n)} \stackrel{d}{=} \frac{1}{n} \sum_{i=1}^n \delta_{n^{1/p} \frac{X_{i,n}}{\|X_n\|_p}}. \quad (6)$$

The distribution of X_n has been intensely studied in the literature. Let us recall some results relevant to us. The empirical distribution of X_n is the random probability measure

$$\nu_n := \frac{1}{n} \sum_{i=1}^n \delta_{X_{i,n}}.$$

We consider ν_n as a random element of the space $\mathcal{M}(\mathbb{R})$ of Borel probability measures on \mathbb{R} endowed with the weak topology. For the following result we refer to [18, Theorem 5.4.3].

Theorem 3.1. *The sequence $(\nu_n)_{n \in \mathbb{N}}$ satisfies a large deviations principle on $\mathcal{M}(\mathbb{R})$ with speed n^2 and a good rate function*

$$\mathcal{I}_0(\mu) = \begin{cases} -\frac{\beta}{2} \int_{\mathbb{R}} \int_{\mathbb{R}} \log |x - y| \mu(dx) \mu(dy) + \int_{\mathbb{R}} |x|^p \mu(dx) - B, & \int_{\mathbb{R}} |x|^p \mu(dx) < +\infty, \\ +\infty, & \int_{\mathbb{R}} |x|^p \mu(dx) = +\infty, \end{cases}$$

where $B := \lim_{n \rightarrow \infty} \frac{1}{n^2} \log C_{n,\beta,p}$ with the constants $C_{n,\beta,p}$ given by (5).

The constant B in the theorem is the limit of the so-called free energy $n^{-2} \log C_{n,\beta,p}$ as $n \rightarrow \infty$. Let us determine its precise value. By [18, Theorem 5.4.3], the function $\mathcal{I}_0 : \mathcal{M}(\mathbb{R}) \rightarrow [0, +\infty]$ has a unique minimizer $\mu_*^{(p)}$ satisfying $\mathcal{I}_0(\mu_*^{(p)}) = 0$, which means that

$$B = -\frac{\beta}{2} \int_{\mathbb{R}} \int_{\mathbb{R}} \log |x - y| \mu_*^{(p)}(dx) \mu_*^{(p)}(dy) + \int_{\mathbb{R}} |x|^p \mu_*^{(p)}(dx).$$

By Proposition 5.3.4 from [18, p. 196], the Lebesgue density of the minimizer is $x \mapsto h_p(x/r_p)/r_p$, where h_p is the Ullman density given in (2), and

$$r_p = \left(\frac{\beta}{2p\alpha_p} \right)^{1/p} \quad \text{and} \quad \alpha_p = \frac{\Gamma(\frac{p+1}{2})}{p\sqrt{\pi}\Gamma(\frac{p}{2})} = \int_{-1}^1 |x|^p h_p(x) dx.$$

Using this and the fact that by [18, Proposition 5.3.4]

$$\int_{\mathbb{R}} \int_{\mathbb{R}} \log |x - y| \mu_*^{(p)}(dx) \mu_*^{(p)}(dy) = \log \frac{r_p}{2} - \frac{1}{2p},$$

we obtain the following explicit formula for the limiting free energy:

$$B = \lim_{n \rightarrow \infty} \frac{1}{n^2} \log C_{n,\beta,p} = \frac{\beta}{2p} \log \left(\frac{\beta \sqrt{\pi} \Gamma(\frac{p}{2})}{2\Gamma(\frac{p+1}{2})} \right) - \frac{\beta}{2} \log 2 - \frac{3\beta}{4p}. \quad (7)$$

3.2. Strategy of the proof

We adopt the notation introduced in the previous sections. The outline of our proof of Theorem 1.1 is as follows.

Step 1: We prove an LDP for the sequence of pairs

$$\left(v_n, \int_{\mathbb{R}} |x|^p v_n(dx) \right) = \left(\frac{1}{n} \sum_{i=1}^n \delta_{X_{i,n}}, \frac{1}{n} \sum_{i=1}^n |X_{i,n}|^p \right), \quad n \in \mathbb{N}, \quad (8)$$

of empirical measures of the vector X_n and empirical p th moments of these measures. In a first attempt, it is natural to try to apply the contraction principle to the LDP for v_n (stated above in Theorem 3.1) with the mapping $\mu \mapsto (\mu, \int_{\mathbb{R}} |x|^p \mu(dx))$. However, this mapping is not continuous in the weak topology. At first sight, this may look like a merely technical issue, but it is not. As we shall see, the correct rate function for the pair $(v_n, \int_{\mathbb{R}} |x|^p v_n(dx))$ does not coincide with what one would expect by a naive application of the contraction principle.

Step 2: Using the contraction principle, we derive an LDP for the sequence of random measures

$$\frac{1}{n} \sum_{i=1}^n \delta_{n^{1/p} \frac{X_{i,n}}{\|X\|_p}}, \quad n \in \mathbb{N},$$

thus proving Theorem 1.1 in the case when Z_n is sampled according to the cone measure on $\mathbb{S}_{p,\beta}^{n-1}$.

Step 3: If U is uniformly distributed on $[0, 1]$, then the sequence $(U^{1/\ell})_{n \in \mathbb{N}}$, where we recall that $\ell = \frac{\beta n(n-1)}{2} + \beta n$, satisfies an LDP with rate function $x \mapsto -\frac{\beta}{2} \log x$ if $x \in (0, 1]$ and $+\infty$ otherwise. Applying the contraction principle, we derive an LDP for the sequence of random measures

$$\frac{1}{n} \sum_{i=1}^n \delta_{n^{1/p} U^{1/\ell} \frac{X_{i,n}}{\|X_n\|_p}}, \quad n \in \mathbb{N},$$

thus proving Theorem 1.1 in the case when Z_n is distributed uniformly on the ball $\mathbb{B}_{p,\beta}^n$.

4. Step 1 – A large deviations principle for the empirical pair

The following theorem captures the large deviations behavior for the sequence of pairs of empirical measures v_n of the vector X_n and empirical p th moments of these measures.

Theorem 4.1. *Let $0 < p < +\infty$ and $\beta \in (0, +\infty)$. Assume that for each $n \in \mathbb{N}$, the random vector $X_n = (X_{1,n}, \dots, X_{n,n})$ has joint density given by (5). Then the sequence of random elements*

$$\left(v_n, \int_{\mathbb{R}} |x|^p v_n(dx) \right) = \left(\frac{1}{n} \sum_{i=1}^n \delta_{X_{i,n}}, \frac{1}{n} \sum_{i=1}^n |X_{i,n}|^p \right), \quad n \in \mathbb{N},$$

satisfies an LDP with speed n^2 and good rate function $\mathcal{I}_1 : \mathcal{M}(\mathbb{R}) \times [0, +\infty) \rightarrow [0, +\infty]$ given by

$$\mathcal{I}_1(\mu, m) := \begin{cases} -\frac{\beta}{2} \int_{\mathbb{R}} \int_{\mathbb{R}} \log |x - y| \mu(dx) \mu(dy) + m + B, & m \geq \int_{\mathbb{R}} |x|^p \mu(dx), \\ +\infty, & m < \int_{\mathbb{R}} |x|^p \mu(dx), \end{cases} \quad (9)$$

where B is the same constant as in (7).

The proof is split into several parts which are considered in the subsections below. The main task is to establish a weak large deviations principle by verifying the conditions of Proposition 2.3 on the space $\mathcal{M}(\mathbb{R}) \times [0, +\infty)$ equipped with its product topology arising from the weak topology on $\mathcal{M}(\mathbb{R})$ and the standard topology on $[0, +\infty)$. More precisely, our aim is to show that for every $(\mu, m) \in \mathcal{M}(\mathbb{R}) \times [0, +\infty)$, we have

$$\inf_{G: (\mu, m) \in G} \limsup_{n \rightarrow \infty} \frac{1}{n^2} \log \mathbb{P} \left[\left(v_n, \int_{\mathbb{R}} |x|^p v_n(dx) \right) \in G \right] \leq -\mathcal{I}_1(\mu, m)$$

and

$$\inf_{G: (\mu, m) \in G} \liminf_{n \rightarrow \infty} \frac{1}{n^2} \log \mathbb{P} \left[\left(v_n, \int_{\mathbb{R}} |x|^p v_n(dx) \right) \in G \right] \geq -\mathcal{I}_1(\mu, m),$$

where the infimum is taken over G 's belonging to some neighborhood base of the pair (μ, m) with respect to the product topology just described. This will be done in Lemmas 4.2, 4.3, and 4.7. The exponential tightness, needed to complete the proof of Theorem 4.1, will be verified in Lemma 4.8.

4.1. The case $m < \int_{\mathbb{R}} |x|^p \mu(dx)$

We shall prove the following lemma, which then implies that the rate function $\mathcal{I}_1(\mu, m)$ of the empirical pair is $+\infty$ when m is smaller than the p th moment of the measure μ .

Lemma 4.2. *Let $\mu \in \mathcal{M}(\mathbb{R})$ and $m \in [0, +\infty)$ be such that $m < \int_{\mathbb{R}} |x|^p \mu(dx)$. Then there is a neighborhood G of (μ, m) in the product space $\mathcal{M}(\mathbb{R}) \times [0, +\infty)$ such that, for all $n \in \mathbb{N}$,*

$$\mathbb{P} \left[\left(v_n, \int_{\mathbb{R}} |x|^p v_n(dx) \right) \in G \right] = 0.$$

Proof. Our goal is to construct explicitly a neighborhood $\mathcal{O}(\mu) \subseteq \mathcal{M}(\mathbb{R})$ of μ (in the weak topology) and a neighborhood $\mathcal{O}(m) \subseteq [0, +\infty)$ of m such that

$$\left\{ v \in \mathcal{M}(\mathbb{R}) : \left(v, \int_{\mathbb{R}} |x|^p v(dx) \right) \in \mathcal{O}(\mu) \times \mathcal{O}(m) \right\} = \emptyset.$$

Putting $G := \mathcal{O}(\mu) \times \mathcal{O}(m)$ then completes the proof of the lemma.

Case 1. Assume first that $\int_{\mathbb{R}} |x|^p \mu(dx) < +\infty$. Since $m < \int_{\mathbb{R}} |x|^p \mu(dx)$, there exists some $\alpha > 0$ such that

$$m + \alpha = \int_{\mathbb{R}} |x|^p \mu(dx). \quad (10)$$

By monotone convergence, we can find $A > 0$ (sufficiently large) such that

$$\left| \int_{\mathbb{R}} \min\{A, |x|^p\} \mu(dx) - \int_{\mathbb{R}} |x|^p \mu(dx) \right| < \frac{\alpha}{3}. \quad (11)$$

Now we define the weak neighborhood of μ as

$$\mathcal{O}(\mu) := \left\{ v \in \mathcal{M}(\mathbb{R}) : \left| \int_{\mathbb{R}} \min\{A, |x|^p\} v(dx) - \int_{\mathbb{R}} \min\{A, |x|^p\} \mu(dx) \right| < \frac{\alpha}{3} \right\}.$$

Indeed, this is a neighborhood in the weak topology because the function $x \mapsto \min\{A, |x|^p\}$ is continuous and bounded. Assume now that $v \in \mathcal{O}(\mu)$. Then it follows from (11) and the triangle inequality that

$$\left| \int_{\mathbb{R}} \min\{A, |x|^p\} v(dx) - \int_{\mathbb{R}} |x|^p \mu(dx) \right| \leq \frac{2\alpha}{3}.$$

From this and (10) we infer

$$\int_{\mathbb{R}} |x|^p v(dx) \geq \int_{\mathbb{R}} \min\{A, |x|^p\} v(dx) \geq \int_{\mathbb{R}} |x|^p \mu(dx) - \frac{2\alpha}{3} = m + \frac{\alpha}{3},$$

which implies that $\int_{\mathbb{R}} |x|^p v(dx) \notin \mathcal{O}(m)$ if we define $\mathcal{O}(m) := (m - \frac{\alpha}{6}, m + \frac{\alpha}{6}) \cap [0, +\infty)$.

Case 2. The case $\int_{\mathbb{R}} |x|^p \mu(dx) = +\infty$ is similar. By monotone convergence, we can find a sufficiently large $A > 0$ such that

$$\int_{\mathbb{R}} \min\{A, |x|^p\} \mu(dx) > 3m. \quad (12)$$

Consider the following weak neighborhood of μ defined as

$$\mathcal{O}(\mu) := \left\{ \nu \in \mathcal{M}(\mathbb{R}) : \left| \int_{\mathbb{R}} \min\{A, |x|^p\} \nu(dx) - \int_{\mathbb{R}} \min\{A, |x|^p\} \mu(dx) \right| < m \right\}.$$

It follows directly from (12) that every $\nu \in \mathcal{O}(\mu)$ satisfies

$$\int_{\mathbb{R}} \min\{A, |x|^p\} \nu(dx) > 2m.$$

In particular, we have $\int_{\mathbb{R}} |x|^p \nu(dx) > 2m$. Defining the neighborhood $\mathcal{O}(m) := (\frac{1}{2}m, \frac{3}{2}m)$, it follows that $\int_{\mathbb{R}} |x|^p \nu(dx) \notin \mathcal{O}(m)$. \square

4.2. The case $m \geq \int_{\mathbb{R}} |x|^p \mu(dx)$

We consider here the case where $m \geq \int_{\mathbb{R}} |x|^p \mu(dx)$. We prove lower and upper bounds in the weak LDP separately. Let us start with the lower bound.

4.2.1. The lower bound

First we note that since we are in the case $m \geq \int_{\mathbb{R}} |x|^p \mu(dx)$, if we assume $m = 0$, then necessarily $\mu = \delta_0$ (the Dirac measure at 0) and so $\mathcal{I}_1(\delta_0, 0) = +\infty$. In particular, for any neighborhood G of $(\delta_0, 0)$,

$$\liminf_{n \rightarrow \infty} \frac{1}{n^2} \log \mathbb{P} \left[\left(\nu_n, \int_{\mathbb{R}} |x|^p \nu_n(dx) \right) \in G \right] \geq -\infty = -\mathcal{I}_1(\delta_0, 0),$$

holds trivially. We therefore assume from now on that $m > 0$. Our aim is to prove the following result.

Lemma 4.3. *Let $\mu \in \mathcal{M}(\mathbb{R})$ and $m \in (0, +\infty)$ be such that $m > \int_{\mathbb{R}} |x|^p \mu(dx)$. Then, for every neighborhood G of (μ, m) in the space $\mathcal{M}(\mathbb{R}) \times [0, +\infty)$, we have*

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \frac{1}{n^2} \log \mathbb{P} \left[\left(\nu_n, \int_{\mathbb{R}} |x|^p \nu_n(dx) \right) \in G \right] \\ & \geq \frac{\beta}{2} \int_{\mathbb{R}} \int_{\mathbb{R}} \log |x - y| \mu(dx) \mu(dy) - B - m. \end{aligned}$$

We may assume that the neighborhood G is of the form

$$G = \mathcal{O}_{\varepsilon, d}(\mu) \times (m - \delta, m + \delta)$$

with $\varepsilon > 0$, $\delta > 0$, $d \in \mathbb{N}$ and

$$\begin{aligned} \mathcal{O}_{\varepsilon, d}(\mu) &= \mathcal{O}_{\varepsilon, f_1, \dots, f_d}(\mu) \\ &= \left\{ \nu \in \mathcal{M}(\mathbb{R}) : \left| \int_{\mathbb{R}} f_i(y) \mu(dy) - \int_{\mathbb{R}} f_i(y) \nu(dy) \right| < \varepsilon \forall i = 1, \dots, d \right\} \end{aligned}$$

for some functions $f_1, \dots, f_d \in \mathcal{C}_b(\mathbb{R})$. Here and in what follows, $\mathcal{C}_b(\mathbb{R})$ denotes the space of bounded continuous functions on \mathbb{R} . We may further assume that $\|f_i\|_{\infty} \leq 1$ for all $i \in \{1, \dots, d\}$.

Before proceeding to precise statements, let us give a heuristic explanation of what follows. We need a lower bound on the probability of the event

$$\left\{ \nu_n \in \mathcal{O}_{\varepsilon, d}(\mu), \int_{\mathbb{R}} |x|^p \nu_n(dx) \in (m - \delta, m + \delta) \right\}.$$

At first sight, the conditions $v_n \in \mathcal{O}_{\varepsilon,d}(\mu)$ and $\int_{\mathbb{R}} |x|^p v_n(dx) \in (m - \delta, m + \delta)$ seem to contradict each other for sufficiently small $\varepsilon > 0$ and $\delta > 0$. Indeed, the first condition states that $v_n \approx \mu$, which seems to imply that $\int_{\mathbb{R}} |x|^p v_n(dx) \approx \int_{\mathbb{R}} |x|^p \mu(dx)$, which in turn contradicts the condition $\int_{\mathbb{R}} |x|^p v_n(dx) \approx m$. However, as we already mentioned above, the map $v \mapsto \int_{\mathbb{R}} |x|^p v(dx)$ is not weakly continuous, so that this argument is incorrect.

As we are aiming for a lower bound, we shall provide an explicit description of how the event $\{v_n \in \mathcal{O}_{\varepsilon,d}(\mu), \int_{\mathbb{R}} |x|^p v_n(dx) \in (m - \delta, m + \delta)\}$ can be realized. The next lemma states that this event occurs if the following three events A_n , B_n and C_n occur simultaneously:

A_n : The one-leave-out empirical measure

$$v'_{n-1} := \frac{1}{n-1} \sum_{i=1}^{n-1} \delta_{X_{i,n}}$$

is “close” to μ in the weak topology,

B_n : the p th moment of v'_{n-1} is “close” to the p th moment of μ ,

C_n : the last element $X_{n,n}$ is an outlier, which is “close” to

$$n^{1/p} \left(m - \int_{\mathbb{R}} |x|^p \mu(dx) \right).$$

The event C_n is crucial because it ensures that $\int_{\mathbb{R}} |x|^p v_n(dx)$ is “close” to m rather than to the strictly smaller number $\int_{\mathbb{R}} |x|^p \mu(dx)$. The precise statement is as follows.

Lemma 4.4. *Let $\varepsilon, \delta > 0$, let $\mu \in \mathcal{M}(\mathbb{R})$ and assume that $m \in (0, +\infty)$ is such that $m > \int_{\mathbb{R}} |x|^p \mu(dx)$. Let us define the interval*

$$D_n := \left(n^{1/p} \left(m - \int_{\mathbb{R}} |y|^p \mu(dy) \right)^{1/p} - n^{1/p-2}, n^{1/p} \left(m - \int_{\mathbb{R}} |y|^p \mu(dy) \right)^{1/p} \right). \quad (13)$$

Then, for all sufficiently large $n \in \mathbb{N}$,

$$\begin{aligned} & \mathbb{P} \left[v_n \in \mathcal{O}_{\varepsilon,d}(\mu), \int_{\mathbb{R}} |x|^p v_n(dx) \in (m - \delta, m + \delta) \right] \\ & \geq \mathbb{P} \left[v'_{n-1} \in \mathcal{O}_{\varepsilon/3,d}(\mu), \left| \int_{\mathbb{R}} |y|^p v'_{n-1}(dy) - \int_{\mathbb{R}} |y|^p \mu(dy) \right| < \frac{\delta}{3}, X_{n,n} \in D_n \right]. \end{aligned}$$

Proof. Recall the definition of $\mathcal{O}_{\varepsilon,d}(\mu)$ (in particular the functions f_1, \dots, f_d involved there) and consider the events

$$A_n := \{v'_{n-1} \in \mathcal{O}_{\varepsilon/3,d}(\mu)\}, \quad (14)$$

$$B_n := \left\{ \left| \int_{\mathbb{R}} |y|^p v'_{n-1}(dy) - \int_{\mathbb{R}} |y|^p \mu(dy) \right| < \frac{\delta}{3} \right\}, \quad (15)$$

$$C_n := \{X_{n,n} \in D_n\}. \quad (16)$$

Step 1. We prove that on the event $A_n \cap B_n \cap C_n$ it holds that $v_n \in \mathcal{O}_{\varepsilon,d}(\mu)$. Indeed, for all $i \in \{1, \dots, d\}$, we observe that

$$\begin{aligned} & \left| \int_{\mathbb{R}} f_i(y) v_n(dy) - \int_{\mathbb{R}} f_i(y) \mu(dy) \right| = \left| \frac{1}{n} \sum_{j=1}^n f_i(X_{j,n}) - \int_{\mathbb{R}} f_i(y) \mu(dy) \right| \\ & = \left| \frac{1}{n} f_i(X_{n,n}) + \frac{1}{n} \sum_{j=1}^{n-1} f_i(X_{j,n}) - \int_{\mathbb{R}} f_i(y) \mu(dy) \right| \\ & = \left| \frac{1}{n} f_i(X_{n,n}) + \frac{n-1}{n} \int_{\mathbb{R}} f_i(y) v'_{n-1}(dy) - \int_{\mathbb{R}} f_i(y) \mu(dy) \right| \\ & \leq \frac{1}{n} + \left| \frac{n-1}{n} \int_{\mathbb{R}} f_i(y) v'_{n-1}(dy) - \int_{\mathbb{R}} f_i(y) \mu(dy) \right|, \end{aligned}$$

where in the last estimate we used the triangle inequality and that $\|f_i\|_\infty \leq 1$ for all $i \in \{1, \dots, d\}$. Using again the triangle inequality (now in the second step), the assumption $v'_{n-1} \in \mathcal{O}_{\varepsilon/3,d}(\mu)$ and the fact that $\|f_i\|_\infty \leq 1$ for all $i \in \{1, \dots, d\}$, we arrive at

$$\begin{aligned} & \left| \frac{n-1}{n} \int_{\mathbb{R}} f_i(y) v'_{n-1}(dy) - \int_{\mathbb{R}} f_i(y) \mu(dy) \right| \\ & \leq \left| \frac{n-1}{n} \int_{\mathbb{R}} f_i(y) v'_{n-1}(dy) - \int_{\mathbb{R}} f_i(y) v'_{n-1}(dy) \right| + \left| \int_{\mathbb{R}} f_i(y) v'_{n-1}(dy) - \int_{\mathbb{R}} f_i(y) \mu(dy) \right| \\ & \leq \frac{1}{n} \left| \int_{\mathbb{R}} f_i(y) v'_{n-1}(dy) \right| + \frac{\varepsilon}{3} \leq \frac{1}{n} + \frac{\varepsilon}{3}. \end{aligned}$$

Obviously, we may choose $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ we have $\frac{1}{n} \leq \frac{\varepsilon}{3}$. Putting everything together, we obtain for all $n \geq n_0$ and $i \in \{1, \dots, d\}$ that

$$\left| \int_{\mathbb{R}} f_i(y) v_n(dy) - \int_{\mathbb{R}} f_i(y) \mu(dy) \right| < \frac{1}{n} + \frac{1}{n} + \frac{\varepsilon}{3} = \varepsilon.$$

In other words, this means that $v_n \in \mathcal{O}_{\varepsilon,d}(\mu)$.

Step 2. Next we prove that on the event $A_n \cap B_n \cap C_n$, we have $\int_{\mathbb{R}} |x|^p v_n(dx) \in (m - \delta, m + \delta)$. Note that we may write

$$\int_{\mathbb{R}} |x|^p v_n(dx) = \frac{n-1}{n} \int_{\mathbb{R}} |y|^p v'_{n-1}(dy) + \frac{1}{n} |X_{n,n}|^p.$$

In order to assure that $\int_{\mathbb{R}} |x|^p v_n(dx) \in (m - \delta, m + \delta)$ it is therefore enough to have

$$\left| \int_{\mathbb{R}} |y|^p \mu(dy) + \frac{1}{n} |X_{n,n}|^p - m \right| \leq \frac{\delta}{3}. \quad (17)$$

Indeed, using the triangle inequality in the second, the definition (15) of B_n and (17) in the third and once again the definition (15) of B_n in the fourth step, would imply that

$$\begin{aligned} & \left| \int_{\mathbb{R}} |x|^p v_n(dx) - m \right| \\ & = \left| \frac{n-1}{n} \int_{\mathbb{R}} |y|^p v'_{n-1}(dy) + \frac{1}{n} |X_{n,n}|^p - m \right| \\ & \leq \left| \frac{n-1}{n} \int_{\mathbb{R}} |y|^p v'_{n-1}(dy) - \int_{\mathbb{R}} |y|^p v'_{n-1}(dy) \right| + \left| \int_{\mathbb{R}} |y|^p v'_{n-1}(dy) - \int_{\mathbb{R}} |y|^p \mu(dy) \right| \\ & \quad + \left| \int_{\mathbb{R}} |y|^p \mu(dy) + \frac{1}{n} |X_{n,n}|^p - m \right| \\ & \leq \frac{1}{n} \int_{\mathbb{R}} |y|^p v'_{n-1}(dy) + \frac{\delta}{3} + \frac{\delta}{3} \\ & \leq \frac{1}{n} \left(\int_{\mathbb{R}} |y|^p \mu(dy) + \frac{\delta}{3} \right) + \frac{\delta}{3} + \frac{\delta}{3} \\ & \leq \delta, \end{aligned}$$

as we may choose $n_1 \in \mathbb{N}$ such that for all $n \geq n_1$,

$$\frac{1}{n} \left(\int_{\mathbb{R}} |y|^p \mu(dy) + \frac{\delta}{3} \right) \leq \frac{\delta}{3}.$$

It remains to prove that (17) holds on the event $A_n \cap B_n \cap C_n$. By definition (16) of the event C_n and the definition (13) of the interval D_n , we have

$$\frac{1}{n} |X_{n,n}|^p - \left(m - \int_{\mathbb{R}} |y|^p \mu(dy) \right) \leq 0.$$

On the other hand, the same definitions of the event C_n and the interval D_n imply that

$$\begin{aligned} & \frac{1}{n} |X_{n,n}|^p - \left(m - \int_{\mathbb{R}} |y|^p \mu(dy) \right) \\ & \geq \left(m - \int_{\mathbb{R}} |y|^p \mu(dy) \right) \cdot \left[\left(1 - \frac{1}{n^2} \left(m - \int_{\mathbb{R}} |y|^p \mu(dy) \right)^{-1/p} \right)^p - 1 \right] \\ & \geq -\frac{\delta}{3} \end{aligned}$$

for all sufficiently large $n \in \mathbb{N}$. This concludes the proof of (17). \square

We shall now prove the lower bound for the probability of the previously discussed event $A_n \cap B_n \cap C_n$, which is based on a decomposition technique. For now, we restrict ourselves to probability measures which are supported on a compact interval and have continuous density on this interval. As we shall see later, it is actually enough to consider this case.

Lemma 4.5. *Let $\mu \in \mathcal{M}(\mathbb{R})$ be supported on an interval $[a, b]$ with $-\infty < a < b < +\infty$ and let $m \in (0, +\infty)$ be such that $m > \int_{\mathbb{R}} |x|^p \mu(dx)$. Assume that μ has Lebesgue density h that is continuous on $[a, b]$ and satisfies $\inf_{x \in [a, b]} h(x) > 0$. Let $d \in \mathbb{N}$, $f_1, \dots, f_d \in C_b(\mathbb{R})$ and $\varepsilon, \delta > 0$. Then*

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \frac{1}{n^2} \log \mathbb{P} \left[v'_{n-1} \in \mathcal{O}_{\varepsilon, d}(\mu), \left| \int_{\mathbb{R}} |y|^p v'_{n-1}(dy) - \int_{\mathbb{R}} |y|^p \mu(dy) \right| < \delta, X_{n,n} \in D_n \right] \\ & \geq \frac{\beta}{2} \int_{\mathbb{R}} \int_{\mathbb{R}} \log |x - y| \mu(dx) \mu(dy) - B - m. \end{aligned}$$

Proof. We use a classical approach similar to the proof of Lemma 5.4.6 in [18, pp. 216–217], but the “outlier” $X_{n,n}$ needs to be treated differently. Denote by $g : [0, 1] \rightarrow [a, b]$ the inverse function of the continuous, strictly monotone function

$$t \mapsto \int_a^t h(x) dx, \quad t \in [a, b].$$

Note that $g(0) = a$ and $g(1) = b$. For each $k \in \{1, \dots, n-1\}$, we define

$$a_k^{(n-1)} := g\left(\frac{2k-1}{2(n-1)}\right) \quad \text{and} \quad b_k^{(n-1)} := g\left(\frac{2k}{2(n-1)}\right).$$

Note that

$$a = b_0^{(n-1)} < a_1^{(n-1)} < b_1^{(n-1)} < a_2^{(n-1)} < b_2^{(n-1)} < \dots < a_{n-1}^{(n-1)} < b_{n-1}^{(n-1)} = b.$$

This way we obtain a decomposition of the interval $[a, b]$ into $2(n-1)$ intervals. Let

$$\Delta_{n-1} := \{(t_1, \dots, t_{n-1}) \in \mathbb{R}^{n-1} : a_k^{(n-1)} \leq t_k \leq b_k^{(n-1)} \text{ for all } k = 1, \dots, n-1\}.$$

For sufficiently large $n \in \mathbb{N}$, we have that, for all $t = (t_1, \dots, t_n) \in \Delta_{n-1}$,

$$\frac{1}{n-1} \sum_{k=1}^{n-1} \delta_{t_k} \in \mathcal{O}_{\varepsilon, d}(\mu) \quad \text{and} \quad \left| \frac{1}{n-1} \sum_{k=1}^{n-1} |t_k|^p - \int_{\mathbb{R}} |x|^p \mu(dx) \right| < \delta.$$

Therefore, recalling from (5) the density of the random vector X_n , we can write

$$\begin{aligned} & \mathbb{P} \left[v'_{n-1} \in \mathcal{O}_{\varepsilon, d}(\mu), \left| \int_{\mathbb{R}} |y|^p v'_{n-1}(dy) - \int_{\mathbb{R}} |y|^p \mu(dy) \right| < \delta, X_{n,n} \in D_n \right] \\ & \geq \frac{1}{C_{n, \beta, p}} \int_{\Delta_{n-1} \times D_n} e^{-n \sum_{k=1}^n |t_k|^p} \prod_{1 \leq i < j \leq n} |t_j - t_i|^\beta d(t_1, \dots, t_n) \\ & = \frac{1}{C_{n, \beta, p}} \int_{\Delta_{n-1}} \int_{D_n} e^{-n |t_n|^p} e^{-n \sum_{k=1}^{n-1} |t_k|^p} \prod_{1 \leq i < j < n} |t_j - t_i|^\beta \prod_{1 \leq i < n} |t_n - t_i|^\beta dt_n d(t_1, \dots, t_{n-1}). \end{aligned}$$

Now we define, for each $k \in \{1, \dots, n-1\}$,

$$\xi_k^{(n-1)} := \max\{|t|^p : t \in [a_k^{(n-1)}, b_k^{(n-1)}]\}.$$

Then

$$\begin{aligned} & \mathbb{P}\left[v'_{n-1} \in \mathcal{O}_{\varepsilon, d}(\mu), \left|\int_{\mathbb{R}} |y|^p v'_{n-1}(\mathrm{d}y) - \int_{\mathbb{R}} |y|^p \mu(\mathrm{d}y)\right| < \delta, X_{n,n} \in D_n\right] \\ & \geq \frac{1}{C_{n,\beta,p}} \int_{\Delta_{n-1}} \int_{D_n} e^{-n|t_n|^p} e^{-n \sum_{k=1}^{n-1} |\xi_k^{(n-1)}|^p} \prod_{1 \leq i < j < n} |a_j^{(n-1)} - b_i^{(n-1)}|^\beta \\ & \quad \times \prod_{1 \leq i < n} |t_n - t_i|^\beta \mathrm{d}t_n \mathrm{d}(t_1, \dots, t_{n-1}) \\ & \geq \frac{1}{C_{n,\beta,p}} e^{-n \sum_{k=1}^{n-1} |\xi_k^{(n-1)}|^p} \prod_{1 \leq i < j < n} |a_j^{(n-1)} - b_i^{(n-1)}|^\beta e^{-n^2(m - \int_{\mathbb{R}} |x|^p \mu(\mathrm{d}x))} \mathrm{vol}_n(\Delta_{n-1} \times D_n), \end{aligned}$$

where we used that on $\Delta_{n-1} \times D_n$ we have $|t_n|^p < n(m - \int_{\mathbb{R}} |x|^p \mu(\mathrm{d}x))$ by (13), and

$$\prod_{1 \leq i < n} |t_n - t_i|^\beta \geq 1 \tag{18}$$

for sufficiently large $n \in \mathbb{N}$. Indeed, on $\Delta_{n-1} \times D_n$ we have $t_1, \dots, t_{n-1} \in [a, b]$, whereas $t_n > n^{1/p}(m - \int_{\mathbb{R}} |y|^p \mu(\mathrm{d}y))^{1/p} - n^{1/p-2}$ by (13), which exceeds $b + 1$ for sufficiently large n . This proves (18). Observe also that

$$\mathrm{vol}_n(\Delta_{n-1} \times D_n) \geq n^{1/p-2} \prod_{k=1}^{n-1} (b_k^{(n-1)} - a_k^{(n-1)}) \geq \left(\frac{1}{(2n-2) \sup_{x \in [a,b]} h(x)} \right)^{n-1} n^{1/p-2},$$

which implies

$$\liminf_{n \rightarrow \infty} \frac{1}{n^2} \log \mathrm{vol}_n(\Delta_{n-1} \times D_n) = 0.$$

Recall also that $\lim_{n \rightarrow \infty} \frac{1}{n^2} \log C_{n,\beta,p} = B$. Using that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^{n-1} |\xi_k^{(n-1)}|^p = \int_{\mathbb{R}} |x|^p \mu(\mathrm{d}x)$$

and

$$\liminf_{n \rightarrow \infty} \frac{1}{n^2} \sum_{1 \leq i < j < n} \log |a_j^{(n-1)} - b_i^{(n-1)}| \geq \frac{1}{2} \int_{\mathbb{R}} \int_{\mathbb{R}} \log |x - y| \mu(\mathrm{d}x) \mu(\mathrm{d}y),$$

we obtain

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \frac{1}{n^2} \log \mathbb{P}\left[v'_{n-1} \in \mathcal{O}_{\varepsilon, d}(\mu), \left|\int_{\mathbb{R}} |y|^p v'_{n-1}(\mathrm{d}y) - \int_{\mathbb{R}} |y|^p \mu(\mathrm{d}y)\right| < \delta, X_{n,n} \in D_n\right] \\ & \geq -B - m + \frac{\beta}{2} \int_{\mathbb{R}} \int_{\mathbb{R}} \log |x - y| \mu(\mathrm{d}x) \mu(\mathrm{d}y), \end{aligned}$$

which is the required estimate. □

Next we shall lift the previous lemma to arbitrary measures $\mu \in \mathcal{M}(\mathbb{R})$ and prove Lemma 4.3 stating the lower bound in the weak LDP for the pair $(v_n, \int_{\mathbb{R}} |x|^p v_n(\mathrm{d}x))$.

Lemma 4.6. Let $\mu \in \mathcal{M}(\mathbb{R})$ and $m \in (0, +\infty)$ be such that $m \geq \int_{\mathbb{R}} |x|^p \mu(dx)$. Let $\varepsilon, \delta > 0$ and $d \in \mathbb{N}$, $f_1, \dots, f_d \in \mathcal{C}_b(\mathbb{R})$. Then

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \frac{1}{n^2} \log \mathbb{P} \left[v_n \in \mathcal{O}_{\varepsilon, d}(\mu), \int_{\mathbb{R}} |y|^p v_n(dy) \in (m - \delta, m + \delta) \right] \\ & \geq \frac{\beta}{2} \int_{\mathbb{R}} \int_{\mathbb{R}} \log |x - y| \mu(dx) \mu(dy) - B - m. \end{aligned}$$

Proof. Following the argumentation of [18, Lemma 5.4.6] there is a sequence $(\mu_k)_{k \in \mathbb{N}}$ of probability measures on \mathbb{R} such that, for all $k \in \mathbb{N}$,

- (i) μ_k is supported on some interval $[a_k, b_k]$ and has continuous Lebesgue density h_k on this interval such that $\inf_{x \in [a_k, b_k]} h_k(x) > 0$,
- (ii) μ_k converges to μ weakly, as $k \rightarrow \infty$,
- (iii) for $k \rightarrow \infty$,

$$\int_{\mathbb{R}} |x|^p \mu_k(dx) \rightarrow \int_{\mathbb{R}} |x|^p \mu(dx).$$

As a consequence of (ii) and (iii), for all sufficiently large $k \in \mathbb{N}$ (in particular depending on the functions f_1, \dots, f_d defining the weak neighborhood),

$$\mathcal{O}_{\varepsilon/3, d}(\mu_k) \subseteq \mathcal{O}_{\varepsilon, d}(\mu).$$

By (iii), we can construct a sequence $(m_k)_{k \in \mathbb{N}}$ such that $\lim_{k \rightarrow \infty} m_k = m$ and $m_k > \int_{\mathbb{R}} |x|^p \mu_k(dx)$ for all $k \in \mathbb{N}$. Indeed, we may simply take

$$m_k := m - \left(\int_{\mathbb{R}} |x|^p \mu(dx) - \int_{\mathbb{R}} |x|^p \mu_k(dx) \right) + \frac{1}{k}, \quad k \in \mathbb{N}.$$

Therefore, for large enough $k \in \mathbb{N}$,

$$\begin{aligned} & \mathbb{P} \left[v_n \in \mathcal{O}_{\varepsilon, d}(\mu), \int_{\mathbb{R}} |y|^p v_n(dy) \in (m - \delta, m + \delta) \right] \\ & \geq \mathbb{P} \left[v_n \in \mathcal{O}_{\varepsilon/3, d}(\mu_k), \int_{\mathbb{R}} |y|^p v_n(dy) \in \left(m_k - \frac{\delta}{2}, m_k + \frac{\delta}{2} \right) \right] \\ & \geq \mathbb{P} \left[v'_{n-1} \in \mathcal{O}_{\varepsilon/9, d}(\mu_k), \left| \int_{\mathbb{R}} |y|^p v'_{n-1}(dy) - \int_{\mathbb{R}} |y|^p \mu_k(dy) \right| < \frac{\delta}{6}, X_{n,n} \in D_{n,k} \right], \end{aligned}$$

where the last inequality follows from Lemma 4.4 and $D_{n,k}$ is defined in the same way as D_n , see (13), but with m replaced by m_k . We can now apply Lemma 4.5 with μ replaced by μ_k there and obtain, for every sufficiently large fixed $k \in \mathbb{N}$,

$$\liminf_{n \rightarrow \infty} \frac{1}{n^2} \log \mathbb{P} \left[v_n \in \mathcal{O}_{\varepsilon, d}(\mu), \int_{\mathbb{R}} |y|^p v_n(dy) \in (m - \delta, m + \delta) \right] \geq -\mathcal{J}_1(\mu_k, m_k).$$

It is now left to show that

$$\limsup_{k \rightarrow \infty} \mathcal{J}_1(\mu_k, m_k) \leq \mathcal{J}_1(\mu, m).$$

But this follows from the upper semi-continuity of the free entropy in the following form: under conditions (ii) and (iii) we have

$$\limsup_{k \rightarrow \infty} \int_{\mathbb{R}} \int_{\mathbb{R}} \log |x - y| \mu_k(dx) \mu_k(dy) \leq \int_{\mathbb{R}} \int_{\mathbb{R}} \log |x - y| \mu(dx) \mu(dy).$$

This standard fact can be verified by the same argument as on p. 214 of [18]. □

4.2.2. The upper bound

Again we recall that $m \geq \int_{\mathbb{R}} |x|^p \mu(dx)$. To obtain the upper bound, we follow a classical idea and consider an appropriate kernel function together with its truncated version (see, e.g., [16–18]).

Lemma 4.7. *Let $\mu \in \mathcal{M}(\mathbb{R})$ and $m \in (0, +\infty)$ be such that $m \geq \int_{\mathbb{R}} |x|^p \mu(dx)$. Let $\varepsilon, \delta > 0$ and $d \in \mathbb{N}$, $f_1, \dots, f_d \in \mathcal{C}_b(\mathbb{R})$. Then*

$$\begin{aligned} & \lim_{\varepsilon \downarrow 0, \delta \downarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{n^2} \log \mathbb{P} \left[v_n \in \mathcal{O}_{\varepsilon, d}(\mu), \int_{\mathbb{R}} |y|^p v_n(dy) \in (m - \delta, m + \delta) \right] \\ & \leq \frac{\beta}{2} \int_{\mathbb{R}} \int_{\mathbb{R}} \log |x - y| \mu(dx) \mu(dy) - B - m. \end{aligned}$$

Proof. For some $\gamma > 0$ define the following weighted logarithmic kernel

$$F(x, y; \gamma) := \begin{cases} -\frac{\beta}{2} \log |x - y| + \gamma \frac{|x|^p + |y|^p}{2}, & x \neq y, \\ +\infty, & x = y, \end{cases}$$

and for $\alpha > 0$ its α -truncated version by

$$F_{\alpha}(x, y; \gamma) := \min\{F(x, y; \gamma), \alpha\}.$$

We observe that $F_{\alpha}(x, x; \gamma) = \alpha$. For $t = (t_1, \dots, t_n) \in \mathbb{R}^n$, let us write $\mu_t = \frac{1}{n} \sum_{i=1}^n \delta_{t_i}$ for the empirical measure of t and denote by H the set

$$H := \left\{ t \in \mathbb{R}^n : \mu_t \in \mathcal{O}_{\varepsilon, d}(\mu), \frac{1}{n} \sum_{i=1}^n |t_i|^p \in (m - \delta, m + \delta) \right\},$$

where we recall that the definition of $\mathcal{O}_{\varepsilon, d}(\mu)$ depends on the functions f_1, \dots, f_d . We have

$$\begin{aligned} & \mathbb{P} \left[v_n \in \mathcal{O}_{\varepsilon, d}(\mu), \int_{\mathbb{R}} |y|^p v_n(dy) \in (m - \delta, m + \delta) \right] \\ & = \frac{1}{C_{n, \beta, p}} \int_H \exp \left\{ -n \sum_{i=1}^n |t_i|^p \right\} \exp \left\{ \beta \sum_{1 \leq i < j \leq n} \log |t_j - t_i| \right\} d(t_1, \dots, t_n) \\ & = \frac{1}{C_{n, \beta, p}} \int_H \exp \left\{ -\sum_{i=1}^n |t_i|^p (n + \gamma - \gamma n) \right\} \exp \left\{ -2 \sum_{1 \leq i < j \leq n} F(t_i, t_j; \gamma) \right\} d(t_1, \dots, t_n) \\ & \leq \frac{1}{C_{n, \beta, p}} \int_H e^{(m - \delta)n(\gamma n - \gamma - n)} \exp \left\{ -n^2 \int_{\mathbb{R}} \int_{\mathbb{R}} F_{\alpha}(x, y; \gamma) \mu_t(dx) \mu_t(dy) + n\alpha \right\} d(t_1, \dots, t_n) \\ & \leq \frac{1}{C_{n, \beta, p}} e^{(m - \delta)n^2(\gamma - 1) - (m - \delta)n\gamma} \exp \left\{ -n^2 \inf_{\mu' \in \mathcal{O}_{\varepsilon, d}(\mu)} \int_{\mathbb{R}} \int_{\mathbb{R}} F_{\alpha}(x, y; \gamma) \mu'(dx) \mu'(dy) + n\alpha \right\}, \end{aligned}$$

where in the second step we used that

$$\begin{aligned} & -n \sum_{i=1}^n |t_i|^p + \beta \sum_{1 \leq i < j \leq n} \log |t_i - t_j| \\ & = -n \sum_{i=1}^n |t_i|^p - 2 \sum_{1 \leq i < j \leq n} F(t_i, t_j; \gamma) + \gamma \sum_{1 \leq i < j \leq n} (|t_i|^p + |t_j|^p) \\ & = -n \sum_{i=1}^n |t_i|^p - 2 \sum_{1 \leq i < j \leq n} F(t_i, t_j; \gamma) + \gamma \sum_{i=1}^{n-1} \left((n - i) |t_i|^p + \sum_{j=i+1}^n |t_j|^p \right) \end{aligned}$$

$$\begin{aligned}
&= -n \sum_{i=1}^n |t_i|^p - 2 \sum_{1 \leq i < j \leq n} F(t_i, t_j; \gamma) + \gamma(n-1) \sum_{i=1}^n |t_i|^p \\
&= - \sum_{i=1}^n |t_i|^p (n + \gamma - \gamma n) - 2 \sum_{1 \leq i < j \leq n} F(t_i, t_j; \gamma).
\end{aligned}$$

Taking the logarithm and dividing by n^2 yields

$$\begin{aligned}
&\frac{1}{n^2} \log \mathbb{P} \left[v_n \in \mathcal{O}_{\varepsilon, d}(\mu), \int_{\mathbb{R}} |y|^p v_n(dy) \in (m - \delta, m + \delta) \right] \\
&\leq -\frac{1}{n^2} \log C_{n, \beta, p} + (m - \delta)(\gamma - 1) - \frac{\gamma(m - \delta)}{n} - \inf_{\mu' \in \mathcal{O}_{\varepsilon, d}(\mu)} \int_{\mathbb{R}} \int_{\mathbb{R}} F_{\alpha}(x, y; \gamma) \mu'(dx) \mu'(dy) + \frac{\alpha}{n}.
\end{aligned}$$

Letting $n \rightarrow \infty$, we arrive at

$$\begin{aligned}
&\limsup_{n \rightarrow \infty} \frac{1}{n^2} \log \mathbb{P} \left[v_n \in \mathcal{O}_{\varepsilon, d}(\mu), \int_{\mathbb{R}} |y|^p v_n(dy) \in (m - \delta, m + \delta) \right] \\
&\leq -B + (m - \delta)(\gamma - 1) - \inf_{\mu' \in \mathcal{O}_{\varepsilon, d}(\mu)} \int_{\mathbb{R}} \int_{\mathbb{R}} F_{\alpha}(x, y; \gamma) \mu'(dx) \mu'(dy).
\end{aligned}$$

The function $(x, y) \mapsto F_{\alpha}(x, y; \gamma)$ is continuous and bounded on \mathbb{R}^2 . It follows that the functional

$$\mathcal{M}(\mathbb{R}) \rightarrow \mathbb{R}, \quad \mu' \mapsto \int_{\mathbb{R}} \int_{\mathbb{R}} F_{\alpha}(x, y; \gamma) \mu'(dx) \mu'(dy)$$

is weakly continuous. Now, if we take $\varepsilon \downarrow 0$ as well as $\delta \downarrow 0$, then

$$\begin{aligned}
&\lim_{\varepsilon \downarrow 0, \delta \downarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{n^2} \log \mathbb{P} \left[v_n \in \mathcal{O}_{\varepsilon, d}(\mu), \int_{\mathbb{R}} |y|^p v_n(dy) \in (m - \delta, m + \delta) \right] \\
&\leq -B + m(\gamma - 1) - \int_{\mathbb{R}} \int_{\mathbb{R}} F_{\alpha}(x, y; \gamma) \mu(dx) \mu(dy).
\end{aligned}$$

This holds for all $\alpha > 0$ and $\gamma > 0$. Fix any $\alpha > 0$. As $\gamma \downarrow 0$, we have

$$F_{\alpha}(x, y; \gamma) \downarrow \min \left\{ -\frac{\beta}{2} \log |x - y|, \alpha \right\}$$

for every $(x, y) \in \mathbb{R}^2$. Moreover, we have the bound $F_{\alpha}(x, y; \gamma) \leq \alpha$. The monotone convergence theorem, applied to the functions $(-F_{\alpha}(x, y; \gamma))_{0 < \gamma < 1}$, yields

$$\begin{aligned}
&\lim_{\varepsilon \downarrow 0, \delta \downarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{n^2} \log \mathbb{P} \left[v_n \in \mathcal{O}_{\varepsilon, d}(\mu), \int_{\mathbb{R}} |y|^p v_n(dy) \in (m - \delta, m + \delta) \right] \\
&\leq -B - m - \int_{\mathbb{R}} \int_{\mathbb{R}} \min \left\{ -\frac{\beta}{2} \log |x - y|, \alpha \right\} \mu(dx) \mu(dy).
\end{aligned}$$

Next we let $\alpha \rightarrow +\infty$, use the monotone convergence theorem this time for the double integral over the set $\{(x, y) \in \mathbb{R}^2 : |x - y| \leq 1\}$, and observe that the integral over $\{(x, y) \in \mathbb{R}^2 : |x - y| \geq 1\}$ does not depend on α , to get

$$\begin{aligned}
&\lim_{\varepsilon \downarrow 0, \delta \downarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{n^2} \log \mathbb{P} \left[v_n \in \mathcal{O}_{\varepsilon, d}(\mu), \int_{\mathbb{R}} |y|^p v_n(dy) \in (m - \delta, m + \delta) \right] \\
&\leq -B - m + \frac{\beta}{2} \int_{\mathbb{R}} \int_{\mathbb{R}} \log |x - y| \mu(dx) \mu(dy).
\end{aligned}$$

This completes the proof of the lemma. \square

4.3. Exponential tightness

Having established the weak LDP for the sequence of pairs $(\nu_n, \int_{\mathbb{R}} |x|^p \nu_n(dx))$, we proceed to the exponential tightness. The next lemma is the last missing part of the proof of Theorem 4.1.

Lemma 4.8. *The sequence of random elements $(\nu_n, \int_{\mathbb{R}} |x|^p \nu_n(dx))$, $n \in \mathbb{N}$, is exponentially tight on $\mathcal{M}(\mathbb{R}) \times [0, +\infty)$.*

Proof. We start by showing that for each $A > 0$ the set $K_A := \{\nu \in \mathcal{M}(\mathbb{R}) : \int_{\mathbb{R}} |x|^p \nu(dx) \leq A\}$ is compact in $\mathcal{M}(\mathbb{R})$. The closedness of K_A follows by considering the cut-off $\min\{|x|^p, n\}$ and using the definition of the weak topology together with the monotone convergence theorem. It therefore remains to show that K_A is relatively compact in $\mathcal{M}(\mathbb{R})$. To do this, we prove that K_A is a tight family of measures and then use Prohorov's theorem (see [28], [23, Theorem 16.3] or [26, Theorem 13.29]), which states that a tight family of probability measures is relatively compact in the weak topology. Given some $\varepsilon > 0$, putting $B := (A/\varepsilon)^{1/p}$ and using Markov's inequality, we find that

$$\nu(\{x \in \mathbb{R} : |x| > B\}) \leq \frac{\int_{\mathbb{R}} |x|^p \nu(dx)}{B^p} \leq \frac{A}{B^p} = \varepsilon.$$

Thus,

$$\sup_{\nu \in K_A} \nu([-B, B]^c) \leq \varepsilon,$$

which shows that the family K_A of probability measures is tight. Hence, the set K_A is compact in $\mathcal{M}(\mathbb{R})$. In particular, this means that the set $K_A^* := K_A \times [0, A]$ is compact in $\mathcal{M}(\mathbb{R}) \times [0, +\infty)$.

To complete the proof of the lemma it suffices to show that

$$\lim_{A \rightarrow +\infty} \limsup_{n \rightarrow \infty} \frac{1}{n^2} \log \mathbb{P} \left[\left(\nu_n, \int_{\mathbb{R}} |x|^p \nu_n(dx) \right) \notin K_A^* \right] = -\infty,$$

which is equivalent to

$$\lim_{A \rightarrow +\infty} \limsup_{n \rightarrow \infty} \frac{1}{n^2} \log \mathbb{P} \left[\int_{\mathbb{R}} |x|^p \nu_n(dx) > A \right] = -\infty. \quad (19)$$

However, the latter property has been established in the proof of [18, Lemma 5.4.8]. \square

Remark 4.9. Theorem 4.1 continues to hold if the space $\mathcal{M}(\mathbb{R}) \times [0, +\infty)$ is replaced by $\mathcal{M}(\mathbb{R}) \times (0, +\infty)$. Clearly, we can consider $(\nu_n, \int_{\mathbb{R}} |x|^p \nu_n(dx))$ as an element of the latter space since the probability of the event $\{\int_{\mathbb{R}} |x|^p \nu_n(dx) = 0\}$ is zero. The proof of the weak LDP does not change. However, in the proof of exponential tightness the set $K_A \times (0, A]$ is not compact and has to be replaced by $K_A \times [A^{-1}, A]$. It remains to check that

$$\lim_{A \rightarrow +\infty} \limsup_{n \rightarrow \infty} \frac{1}{n^2} \log \mathbb{P} \left[\int_{\mathbb{R}} |x|^p \nu_n(dx) < \frac{1}{A} \right] = -\infty. \quad (20)$$

Using the formula for the joint density of $(X_{1,n}, \dots, X_{n,n})$ given in (5), we can write

$$\begin{aligned} \mathbb{P} \left[\int_{\mathbb{R}} |x|^p \nu_n(dx) < \frac{1}{A} \right] &= \mathbb{P} \left[\sum_{i=1}^n |X_{i,n}|^p < \frac{n}{A} \right] \\ &= \frac{1}{C_{n,\beta,p}} \int_{\mathbb{R}^n} \left(e^{-n \sum_{i=1}^n |x_i|^p} \prod_{1 \leq i < j \leq n} |x_j - x_i|^\beta \right) \mathbb{1}_{\{\sum_{i=1}^n |x_i|^p < \frac{n}{A}\}} d(x_1, \dots, x_n) \\ &\leq \frac{1}{C_{n,\beta,p}} \int_{\mathbb{R}^n} \left(\prod_{1 \leq i < j \leq n} |x_j - x_i|^\beta \right) \mathbb{1}_{\{\sum_{i=1}^n |x_i|^p < \frac{n}{A}\}} d(x_1, \dots, x_n) \\ &\leq \frac{1}{C_{n,\beta,p}} \left(\frac{n}{A} \right)^{\frac{\beta n(n-1)}{2p} + \frac{n}{p}} \int_{\mathbb{R}^n} \left(\prod_{1 \leq i < j \leq n} |\lambda_j - \lambda_i|^\beta \right) \mathbb{1}_{\{\sum_{i=1}^n |\lambda_i|^p < 1\}} d(\lambda_1, \dots, \lambda_n), \end{aligned}$$

where we used the change of variables $x_i = (n/A)^{1/p} \lambda_i$, $1 \leq i \leq n$. The asymptotic behavior of the integral on the right-hand side has been determined in [21, Lemma 3.9], where this integral was denoted by $I_{n,\beta,p}$. For some constant $\Delta(p) > 0$ (whose explicit value can be found in [21, Theorem 3.1]), we have

$$\begin{aligned} \mathbb{P} \left[\int_{\mathbb{R}} |x|^p \nu_n(dx) < \frac{1}{A} \right] &\leq \frac{1}{C_{n,\beta,p}} \left(\frac{n}{A} \right)^{\frac{\beta n(n-1)}{2p} + \frac{n}{p}} n^{-\frac{\beta n^2}{2p}} (\Delta^\beta(p) (1 + o(1)))^{\frac{n^2}{2}} \\ &\leq e^{O(1)n^2} \left(\frac{1}{A} \right)^{\frac{\beta n(n-1)}{2p} + \frac{n}{p}}, \end{aligned}$$

where the $O(1)$ -constant does not depend on A . It follows that

$$\limsup_{n \rightarrow \infty} \frac{1}{n^2} \log \mathbb{P} \left[\int_{\mathbb{R}} |x|^p \nu_n(dx) < \frac{1}{A} \right] \leq O(1) - \frac{\beta \log A}{2p},$$

which implies (20).

5. Step 2 – LDP for the cone measure

We shall now use the contraction principle (see Proposition 2.5) to prove the part of Theorem 1.1 relating to the cone measure.

Proposition 5.1. *For every $n \in \mathbb{N}$, let Z_n be a random matrix sampled according to the cone measure on $\mathbb{S}_{p,\beta}^{n-1}$. Then the random probability measure*

$$\mu_n = \frac{1}{n} \sum_{i=1}^n \delta_{n^{1/p} \lambda_i(Z_n)}$$

satisfies an LDP on $\mathcal{M}(\mathbb{R})$ with speed n^2 and good rate function $\mathcal{J} : \mathcal{M}(\mathbb{R}) \rightarrow [0, +\infty]$ given by (1).

Proof. Recall from (6) the distributional equality

$$\mu_n = \frac{1}{n} \sum_{i=1}^n \delta_{n^{1/p} \lambda_i(Z_n)} \stackrel{d}{=} \frac{1}{n} \sum_{i=1}^n \delta_{\frac{n^{1/p} X_{i,n}}{\|X_n\|_p}}.$$

We shall prove an LDP for the right-hand side. To this end, we shall apply the contraction principle to the LDP obtained in Theorem 4.1. Let us define the function

$$F_p : \mathcal{M}(\mathbb{R}) \times (0, +\infty) \rightarrow \mathcal{M}(\mathbb{R}), \quad F_p(\mu, c)(A) := \mu(c^{1/p} A), \quad (21)$$

where A is any Borel subset of \mathbb{R} . Note that the value $c = 0$ is excluded. It follows from [25, Lemma 3.1] that this function is continuous. In the following, it will be convenient to use the notation

$$m_p(\nu) := \int_{\mathbb{R}} |x|^p \nu(dx), \quad \nu \in \mathcal{M}(\mathbb{R}).$$

Recall that $\nu_n := \frac{1}{n} \sum_{i=1}^n \delta_{X_{i,n}}$ and observe that, for every $n \in \mathbb{N}$ and all Borel sets $A \subseteq \mathbb{R}$,

$$F_p(\nu_n, m_p(\nu_n))(A) = \nu_n(m_p^{1/p}(\nu_n) A) = \frac{1}{n} \sum_{i=1}^n \delta_{X_{i,n}}(n^{-1/p} \|X_n\|_p A) = \frac{1}{n} \sum_{i=1}^n \delta_{\frac{n^{1/p} X_{i,n}}{\|X_n\|_p}}(A).$$

Since the mapping F_p defined in (21) is continuous, it follows from the contraction principle (see Proposition 2.5) in combination with Theorem 4.1 above (see also Remark 4.9) that the sequence of random elements

$$F_p(\nu_n, m_p(\nu_n)) = \frac{1}{n} \sum_{i=1}^n \delta_{\frac{n^{1/p} X_{i,n}}{\|X_n\|_p}}, \quad n \in \mathbb{N},$$

satisfies an LDP with speed n^2 and good rate function $\mathcal{J}_2 : \mathcal{M}(\mathbb{R}) \rightarrow [0, +\infty]$ given by

$$\mathcal{J}_2(\mu) = \inf_{(v,m): F_p(v,m)=\mu} \mathcal{J}_1(v,m) = \inf_{(v,m): v(m^{1/p}\cdot)=\mu(\cdot)} \mathcal{J}_1(v,m),$$

where $\mathcal{J}_1 : \mathcal{M}(\mathbb{R}) \times (0, +\infty) \rightarrow [0, +\infty]$ is the rate function from Theorem 4.1. It remains to check that $\mathcal{J}_2(\mu) = \mathcal{J}(\mu)$ for all $\mu \in \mathcal{M}(\mathbb{R})$.

Case 1. Let $\mu \in \mathcal{M}(\mathbb{R})$ be such that $m_p(\mu) > 1$. If (v, m) is such that $\mu(\cdot) = v(m^{1/p}\cdot)$, then

$$1 < m_p(\mu) = \int_{\mathbb{R}} |x|^p \mu(dx) = \int_{\mathbb{R}} |m^{-1/p} y|^p v(dy),$$

which implies $m < \int_{\mathbb{R}} |y|^p v(dy)$. Consequently, $\mathcal{J}_1(v, m) = +\infty$, which shows that in this case $\mathcal{J}_2(\mu) = +\infty = \mathcal{J}(\mu)$.

Case 2. Now assume that $\mu \in \mathcal{M}(\mathbb{R})$ satisfies $m_p(\mu) \leq 1$. If (v, m) is such that $\mu(\cdot) = v(m^{1/p}\cdot)$, then $m \geq \int_{\mathbb{R}} |y|^p v(dy)$. We obtain

$$\begin{aligned} \mathcal{J}(\mu) &= \inf_{(v,m): v(m^{1/p}\cdot)=\mu(\cdot)} \mathcal{J}_1(v,m) \\ &= \inf_{(v,m): v(m^{1/p}\cdot)=\mu(\cdot)} \left(-\frac{\beta}{2} \int_{\mathbb{R}} \int_{\mathbb{R}} \log |x-y| v(dx) v(dy) + m + B \right) \\ &= \inf_{(v,m): v(m^{1/p}\cdot)=\mu(\cdot)} \left(-\frac{\beta}{2} \int_{\mathbb{R}} \int_{\mathbb{R}} \log \frac{|x-y|}{m^{1/p}} v(dx) v(dy) - \frac{\beta}{2p} \log m + m + B \right) \\ &= \inf_{(v,m): v(m^{1/p}\cdot)=\mu(\cdot)} \left(-\frac{\beta}{2} \int_{\mathbb{R}} \int_{\mathbb{R}} \log |x-y| \mu(dx) \mu(dy) - \frac{\beta}{2p} \log m + m + B \right) \\ &= -\frac{\beta}{2} \int_{\mathbb{R}} \int_{\mathbb{R}} \log |x-y| \mu(dx) \mu(dy) - \frac{\beta}{2p} \log \frac{\beta}{2p} + \frac{\beta}{2p} + B. \end{aligned}$$

Recalling the formula for B given in (7), we can compute the constant term as follows:

$$\begin{aligned} B - \frac{\beta}{2p} \log \frac{\beta}{2p} + \frac{\beta}{2p} &= \left(\frac{\beta}{2p} \log \left(\frac{\beta \sqrt{\pi} \Gamma(\frac{p}{2})}{2 \Gamma(\frac{p+1}{2})} \right) - \frac{\beta}{2} \log 2 - \frac{3\beta}{4p} \right) - \frac{\beta}{2p} \log \frac{\beta}{2p} + \frac{\beta}{2p} \\ &= \frac{\beta}{2p} \log \left(\frac{\sqrt{\pi} p \Gamma(\frac{p}{2})}{2^p \sqrt{e} \Gamma(\frac{p+1}{2})} \right), \end{aligned}$$

which completes the proof that $\mathcal{J}_2(\mu) = \mathcal{J}(\mu)$. □

6. Step 3 – LDP for the uniform distribution on the ball

We shall now apply once more the contraction principle to prove the remaining part of Theorem 1.1. We prepare the proof with the following auxiliary result on large deviations for sequences of beta random variables. We fix $a, b > 0$ and recall that a random variable ξ is beta distributed with parameters a and b , we write $\xi \sim \beta_{a,b}$, provided that its Lebesgue density is given by

$$x \mapsto \frac{1}{B(a,b)} x^{a-1} (1-x)^{b-1}, \quad x \in [0, 1],$$

where $B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$ is Euler's beta function. The next result is a slightly more general version of [1, Lemma 4.1]. The proof is almost literally the same and for this reason omitted.

Lemma 6.1. *Fix some $s > 0$ and let $(a_n)_{n \in \mathbb{N}}$ and $(b_n)_{n \in \mathbb{N}}$ be positive sequences such that the limits*

$$\lim_{n \rightarrow \infty} \frac{a_n}{n^s} = a \in [0, \infty) \quad \text{and} \quad \lim_{n \rightarrow \infty} \frac{b_n}{n^s} = b \in [0, \infty)$$

exist and are not both zero. For each $n \in \mathbb{N}$ assume that ξ_n is beta distributed with parameters a_n and b_n . Then the sequence $(\xi_n)_{n \in \mathbb{N}}$ satisfies a LDP with speed n^s and rate function $\mathcal{I}_{a,b}$ given by

$$\mathcal{I}_{a,b}(y) = \begin{cases} -a \log \frac{y}{a} - b \log \frac{1-y}{b} - (a+b) \log(a+b), & y \in (0, 1) \text{ and } a > 0, b > 0, \\ -a \log y, & y \in (0, 1] \text{ and } a > 0, b = 0, \\ -b \log(1-y), & y \in [0, 1) \text{ and } a = 0, b > 0. \end{cases}$$

We are now prepared to present the last step in the proof of Theorem 1.1.

Proposition 6.2. *For every $n \in \mathbb{N}$, let Z_n be a random matrix sampled uniformly from $\mathbb{B}_{p,\beta}^n$. Then the random probability measure*

$$\mu_n = \frac{1}{n} \sum_{i=1}^n \delta_{n^{1/p} \lambda_i(Z_n)}$$

satisfies an LDP on $\mathcal{M}(\mathbb{R})$ with speed n^2 and good rate function $\mathcal{I} : \mathcal{M}(\mathbb{R}) \rightarrow [0, +\infty]$ given by (1).

Proof. Recall from (4) the representation

$$\mu_n = \frac{1}{n} \sum_{i=1}^n \delta_{n^{1/p} \lambda_i(Z_n)} \stackrel{d}{=} \frac{1}{n} \sum_{i=1}^n \delta_{U^{1/\ell} \frac{X_{i,n}}{\|X_n\|_p}},$$

where $\ell = \ell(n, \beta) = \beta \frac{n-1}{2} + \beta n$, $X_n = (X_{1,n}, \dots, X_{n,n})$ has joint Lebesgue density given in (5) and U is independent from X_n and uniformly distributed on $[0, 1]$. The ultimate goal is to prove an LDP for the sequence

$$\frac{1}{n} \sum_{i=1}^n \delta_{U^{1/\ell} n^{1/p} \frac{X_{i,n}}{\|X_n\|_p}}.$$

Note that for each $n \in \mathbb{N}$ we have that $U^{1/\ell} \sim \beta_{\ell,1}$. Thus, Lemma 6.1 with $a_n = \ell$, $b_n = 0$, $a = \beta/2$, $b = 0$ and $s = 2$ implies that the sequence $(U^{1/\ell})_{n \in \mathbb{N}}$ satisfies an LDP with speed n^2 and rate function

$$\mathcal{I}_{\beta/2,0}(x) = \begin{cases} -\frac{\beta}{2} \log x, & x \in (0, 1], \\ +\infty, & \text{otherwise.} \end{cases}$$

In the next step towards the final LDP, we observe that as a direct consequence of Proposition 2.4, Proposition 5.1 and the previous LDP for $(U^{1/\ell})_{n \in \mathbb{N}}$, the sequence of pairs

$$\left(\left(\frac{1}{n} \sum_{i=1}^n \delta_{n^{1/p} \frac{X_{i,n}}{\|X_n\|_p}}, U^{1/\ell} \right) \right)_{n \in \mathbb{N}}$$

satisfies an LDP with speed n^2 and rate function $\mathcal{I}_4 : \mathcal{M}(\mathbb{R}) \times (0, +\infty) \rightarrow [0, +\infty]$ given by

$$\begin{aligned} \mathcal{I}_4(v, z) &= \mathcal{I}(v) + \mathcal{I}_{\beta/2,0}(z) \\ &= -\frac{\beta}{2} \int_{\mathbb{R}} \int_{\mathbb{R}} \log |x - y| v(dx) v(dy) + \frac{\beta}{2p} \log \left(\frac{\sqrt{\pi} p \Gamma(\frac{p}{2})}{2^p \sqrt{e} \Gamma(\frac{p+1}{2})} \right) - \frac{\beta}{2} \log z. \end{aligned}$$

This time we shall apply the contraction principle with a continuous function similar to the one defined in (21) but for $p = -1$, i.e., with

$$F : \mathcal{M}(\mathbb{R}) \times (0, +\infty) \rightarrow \mathcal{M}(\mathbb{R}), \quad F(\mu, c)(A) = \mu \left(\frac{1}{c} A \right),$$

where A is any Borel subset of \mathbb{R} . This map is continuous. Note that, for any Borel set $A \subseteq \mathbb{R}$,

$$F \left(\left(\frac{1}{n} \sum_{i=1}^n \delta_{n^{1/p} \frac{X_{i,n}}{\|X_n\|_p}}, U^{1/\ell} \right) \right) (A) = \frac{1}{n} \sum_{i=1}^n \delta_{n^{1/p} \frac{X_{i,n}}{\|X_n\|_p}} (U^{-1/\ell} A) = \frac{1}{n} \sum_{i=1}^n \delta_{U^{1/\ell} n^{1/p} \frac{X_{i,n}}{\|X_n\|_p}} (A). \quad (22)$$

The contraction principle (see Proposition 2.5) shows that the empirical measures in (22) satisfy an LDP on $\mathcal{M}(\mathbb{R})$ with speed n^2 and rate function

$$\mathcal{I}_5(\mu) = \inf_{(v,z): F(v,z)=\mu} \mathcal{I}_4(v,z) = \inf_{(v,z): F(v,z)=\mu} \left(\mathcal{I}(v) - \frac{\beta}{2} \log z \right).$$

Here and in the following, z is restricted to the interval $(0, 1]$. It remains to check that $\mathcal{I}_5(\mu) = \mathcal{I}(\mu)$ for all $\mu \in \mathcal{M}(\mathbb{R})$.

Case 1. Let $\mu \in \mathcal{M}(\mathbb{R})$ be such that $m_p(\mu) > 1$. If (v, z) is such that $\mu(\cdot) = v(z^{-1}\cdot)$, then $m_p(v) = z^{-p}m_p(\mu) > 1$. Consequently, $\mathcal{I}(v) = +\infty$ and it follows that in this case $\mathcal{I}_5(\mu) = +\infty = \mathcal{I}(\mu)$.

Case 2. Let $\mu \in \mathcal{M}(\mathbb{R})$ be such that $m_p(\mu) \leq 1$. If (v, z) is such that $\mu(\cdot) = v(z^{-1}\cdot)$, then $m_p(v) = z^{-p}m_p(\mu)$ and we have two possibilities. If $m_p(v) > 1$, then $\mathcal{I}(v) - \frac{\beta}{2} \log z = +\infty$. So we let in the following $m_p(v) \leq 1$, which means that z is restricted to the non-empty interval $[m_p^{1/p}(\mu), 1]$. Then,

$$\begin{aligned} \mathcal{I}(\mu) &= \inf_{(v,z): F(v,z)=\mu} \mathcal{I}_4(v,z) \\ &= \inf_{(v,z): v(z^{-1}\cdot)=\mu(\cdot)} \left(-\frac{\beta}{2} \int_{\mathbb{R}} \int_{\mathbb{R}} \log |x-y| v(dx) v(dy) \right. \\ &\quad \left. + \frac{\beta}{2p} \log \left(\frac{\sqrt{\pi} p \Gamma(\frac{p}{2})}{2^p \sqrt{e} \Gamma(\frac{p+1}{2})} \right) - \frac{\beta}{2} \log z \right) \\ &= \inf_{(v,z): v(z^{-1}\cdot)=\mu(\cdot)} \left(\frac{\beta}{2} \log z - \frac{\beta}{2} \int_{\mathbb{R}} \int_{\mathbb{R}} \log |x-y| \mu(dx) \mu(dy) \right. \\ &\quad \left. + \frac{\beta}{2p} \log \left(\frac{\sqrt{\pi} p \Gamma(\frac{p}{2})}{2^p \sqrt{e} \Gamma(\frac{p+1}{2})} \right) - \frac{\beta}{2} \log z \right) \\ &= -\frac{\beta}{2} \int_{\mathbb{R}} \int_{\mathbb{R}} \log |x-y| \mu(dx) \mu(dy) + \frac{\beta}{2p} \log \left(\frac{\sqrt{\pi} p \Gamma(\frac{p}{2})}{2^p \sqrt{e} \Gamma(\frac{p+1}{2})} \right), \end{aligned}$$

which shows that again $\mathcal{I}_5(\mu) = \mathcal{I}(\mu)$. □

7. Proof of Corollary 1.4

By [18, Proposition 5.3.4], the probability measure $\mu_\infty^{(p)}$ is the only minimizer of \mathcal{I} and it holds that $\mathcal{I}(\mu_\infty^{(p)}) = 0$.

Let $d(\cdot, \cdot)$ be any metric on $\mathcal{M}(\mathbb{R})$ metrizing the topology of weak convergence. Then, for every $\varepsilon > 0$,

$$\limsup_{n \rightarrow \infty} \frac{1}{n^2} \log \mathbb{P}[d(\mu_n, \mu_\infty^{(p)}) \geq \varepsilon] \leq - \inf_{\tau \in \mathcal{M}(\mathbb{R}): d(\tau, \mu_\infty^{(p)}) \geq \varepsilon} \mathcal{I}(\tau).$$

We claim that the infimum on the right-hand side is strictly negative. Indeed, if it were 0, we could find a sequence $(\tau_k)_{k \in \mathbb{N}}$ such that $d(\tau_k, \mu_\infty^{(p)}) \geq \varepsilon$ for all $k \in \mathbb{N}$ and $\lim_{k \rightarrow \infty} \mathcal{I}(\tau_k) = 0$. Since the rate function \mathcal{I} is good, the set $\{\tau_k : k \in \mathbb{N}\}$ is weakly compact, and by passing to a subsequence we may assume that $\tau_k \rightarrow \tau$ weakly, as $k \rightarrow \infty$. By the lower semicontinuity and non-negativity of \mathcal{I} , this implies that $\mathcal{I}(\tau) = 0$. This means that $\tau = \mu_\infty^{(p)}$, which contradicts the assumption that $d(\tau_k, \mu_\infty^{(p)}) \geq \varepsilon$ for all $k \in \mathbb{N}$.

Hence, the infimum is negative and we have, for all sufficiently large $n > n_0(\varepsilon) \in \mathbb{N}$ and some $c = c(\varepsilon) > 0$,

$$\mathbb{P}[d(\mu_n, \mu_\infty^{(p)}) \geq \varepsilon] \leq e^{-cn^2}.$$

From the Borel–Cantelli Lemma it follows that $d(\mu_n, \mu_\infty^{(p)}) \rightarrow 0$ with probability 1.

8. Sketch of the proofs in the non self-adjoint case

The proof of Theorem 1.5 is essentially the same as the one of Theorem 1.1 (if $0 < p < +\infty$) and Theorem 1.3 if $p = +\infty$. For this reason, we only sketch the argument and highlight the points, where differences to the self-adjoint case occur. As above, we restrict our attention to the case $p < +\infty$.

Key ingredients in the self-adjoint case were the distributional representations (4) (for the uniform distribution) and (6) (for the cone measure) taken from [21]. The proof of these representations in [21] was based on a change-of-variables formula (Proposition 4.1.1 in [3]). In the non self-adjoint case, a reformulation of this formula reads as follows (see [3, Proposition 4.1.3]):

Lemma 8.1. *Let $0 < p < +\infty$ and $\beta \in \{1, 2, 4\}$. For each $n \in \mathbb{N}$ let $Z_n \in \mathbb{SB}_p^n$ be uniformly distributed and, independently, let π be a uniform random permutation on $\{1, \dots, n\}$. Then the random vector $(s_{\pi(1)}^2(Z_n), \dots, s_{\pi(n)}^2(Z_n))$ has density on \mathbb{R}_+^n proportional to*

$$(x_1, \dots, x_n) \mapsto \prod_{1 \leq i < j \leq n} |x_i - x_j|^\beta \prod_{i=1}^n x_i^{\frac{\beta}{2}-1} \mathbb{1}_{\{\sum_{i=1}^n x_i^{p/2} \leq 1\}}.$$

Noting that the function

$$\mathbb{R}_+^n \rightarrow \mathbb{R}, (x_1, \dots, x_n) \mapsto h_\beta(x_1, \dots, x_n) := \prod_{1 \leq i < j \leq n} |x_i - x_j|^\beta \prod_{i=1}^n x_i^{\frac{\beta}{2}-1}$$

is homogeneous of degree

$$m = m(n) := \frac{\beta n(n-1)}{2} + n \left(\frac{\beta}{2} - 1 \right)$$

we can now argue as in the proof of Lemma 4.2 and Corollary 4.3 in [21] to conclude the following Schechtman-Zinn-type distributional representations for Schatten p -classes (note that instead of p the value $p/2$ occurs at many places since we consider the squares of the singular values and not the singular values themselves).

Proposition 8.2. *Let $0 < p < +\infty$ and $\beta \in \{1, 2, 4\}$. For each $n \in \mathbb{N}$ let $Z_n \in \mathbb{SB}_p^n$ be uniformly distributed and, independently, let π be a uniform random permutation of $\{1, \dots, n\}$. Then*

$$(s_{\pi(1)}^2(Z_n), \dots, s_{\pi(n)}^2(Z_n)) \stackrel{d}{=} U^{\frac{1}{n+m}} \frac{X_n}{\|X_n\|_{p/2}},$$

where the random vector $X_n \in \mathbb{R}_+^n$ has joint density proportional to

$$(x_1, \dots, x_n) \mapsto e^{-n \sum_{i=1}^n |x_i|^{p/2}} h_\beta(x_1, \dots, x_n) \quad (23)$$

and, independently of X_n , U is uniformly distributed on $(0, 1)$. Also, if $Z_n \in \partial \mathbb{SB}_p^n$ is distributed according to the cone probability measure, one has that

$$(s_{\pi(1)}^2(Z_n), \dots, s_{\pi(n)}^2(Z_n)) \stackrel{d}{=} \frac{X_n}{\|X_n\|_{p/2}}.$$

In a next step, one proves an LDP for the sequence of empirical measures

$$\frac{1}{n} \sum_{i=1}^n \delta_{X_n^{(i)}}, \quad n \in \mathbb{N}$$

determined by the coordinates $X_n^{(i)}$ of the random vector $X_n = (X_n^{(1)}, \dots, X_n^{(n)})$ with joint density (23). In fact, from [18, Theorem 5.5.1] (with the choice $m(n) = n$, $\gamma = \beta/2$, $\alpha = 1$ and $\gamma = 0$ there) we conclude that this sequence satisfies an LDP on the space $\mathcal{M}(\mathbb{R}_+)$ of Borel probability measures on \mathbb{R}_+ endowed with the weak topology with speed n^2 and good rate function

$$\mu \mapsto -\frac{\beta}{2} \int_{\mathbb{R}_+} \int_{\mathbb{R}_+} \log |x - y| \mu(dx) \mu(dy) + \int_{\mathbb{R}_+} |x|^{p/2} \mu(dx) + B.$$

As in the proof for the self-adjoint case it is then possible to prove that the sequence of pairs

$$\left(\sum_{i=1}^n \delta_{X_n^{(i)}}, \sum_{i=1}^n (X_n^{(i)})^{p/2} \right), \quad n \in \mathbb{N},$$

satisfies an LDP on the product space $\mathcal{M}(\mathbb{R}_+) \times \mathbb{R}_+$ supplied with the product of the weak topology on $\mathcal{M}(\mathbb{R}_+)$ and the standard topology on \mathbb{R}_+ with good rate function

$$(\mu, m) \mapsto \begin{cases} -\frac{\beta}{2} \int_{\mathbb{R}_+} \int_{\mathbb{R}_+} \log |x - y| \mu(\mathrm{d}x) \mu(\mathrm{d}y) + \int_{\mathbb{R}_+} |x|^{p/2} \mu(\mathrm{d}x) + B, & m \geq \int_{\mathbb{R}_+} |x|^{p/2} \mu(\mathrm{d}x), \\ +\infty, & m < \int_{\mathbb{R}_+} |x|^{p/2} \mu(\mathrm{d}x). \end{cases}$$

Continuing the argument along the lines of the self-adjoint case one concludes that if Z_n is uniformly distributed in \mathbb{SB}_p^n or distributed according to the cone measure on $\partial \mathbb{SB}_p^n$ then the sequence of random probability measures

$$\frac{1}{n} \sum_{i=1}^n \delta_{n^{2/p} s_i^2(Z_n)}, \quad n \in \mathbb{N},$$

satisfies an LDP on $\mathcal{M}(\mathbb{R}_+)$ with speed n^2 and good rate function

$$\mu \mapsto \begin{cases} -\frac{\beta}{2} \int_{\mathbb{R}_+} \int_{\mathbb{R}_+} \log |x - y| \mu(\mathrm{d}x) \mu(\mathrm{d}y) + C_{p,\beta}, & \int_{\mathbb{R}_+} |x|^{p/2} \mu(\mathrm{d}x) \leq 1, \\ +\infty, & \int_{\mathbb{R}_+} |x|^{p/2} \mu(\mathrm{d}x) > 1, \end{cases} \quad (24)$$

where $C_{p,\beta}$ is a constant depending on p and on β .

To determine the value of $C_{p,\beta}$ we use the same symmetrization trick as in the proof of [20, Proposition 6]. Namely, in order to find a Borel probability measure $\mu \in \mathcal{M}(\mathbb{R}_+)$ which minimizes the functional $\mu \mapsto -\frac{\beta}{2} \int_{\mathbb{R}_+} \int_{\mathbb{R}_+} \log |x - y| \mu(\mathrm{d}x) \mu(\mathrm{d}y)$ under the condition that $\int_{\mathbb{R}_+} |x|^{p/2} \mu(\mathrm{d}x) \leq 1$, we maximize the expression $\frac{\beta}{2} \mathbb{E} \log |V - \tilde{V}|$ over all random variables $V \geq 0$ under the condition that $\mathbb{E} V^{p/2} \leq 1$, where \tilde{V} is an independent copy of V . Let ε be a Rademacher random variable taking the values $+1$ and -1 with probability $1/2$ and assume that ε and V are independent. Putting $U := \varepsilon \sqrt{V}$ and denoting by \tilde{U} an independent copy of U we have that

$$\frac{\beta}{2} \mathbb{E} \log |V - \tilde{V}| = \frac{\beta}{2} \mathbb{E} \log |U^2 - \tilde{U}^2| = \frac{\beta}{2} \mathbb{E} \log |U - \tilde{U}| + \frac{\beta}{2} \mathbb{E} \log |U + \tilde{U}| = \beta \mathbb{E} \log |U - \tilde{U}|.$$

This means that we need to maximize the functional $\beta \mathbb{E} \log |U - \tilde{U}|$ over all symmetric random variables U on \mathbb{R} under the condition that $\mathbb{E} |U|^p \leq 1$. It is known that the maximizer of this problem is given by a random variable with Lebesgue density $b_p^{-1} h_p(x/b_p)$, where $h_p(x)$ and b_p are given by (2), see [30]. Using now the computations carried out in [21, Section 2.5] we conclude that the precise value of the constant $C_{p,\beta}$ in the above rate function (24) equals

$$C_{p,\beta} = \frac{\beta}{p} \log \left(\frac{\sqrt{\pi} p \Gamma(\frac{p}{2})}{2^p \sqrt{e} \Gamma(\frac{p+1}{2})} \right).$$

This completes the argument leading to 1.5. Finally, the proof of the law of large numbers in Corollary 1.6 is literally the same as that of Corollary 1.4.

Acknowledgements

JP has been supported by a *Visiting International Professor Fellowship* from the Ruhr University Bochum and its Research School PLUS. ZK and CT were supported by the DFG Scientific Network *Cumulants, Concentration and Superconcentration*. We also thank an anonymous referee for helpful comments and suggestions.

References

- [1] D. Alonso-Gutiérrez, J. Prochno and C. Thäle. Large deviations for high-dimensional random projections of ℓ_p^n -balls. *Adv. in Appl. Math.* **99** (2018) 1–35. [MR3806754 https://doi.org/10.1016/j.aam.2018.04.003](https://doi.org/10.1016/j.aam.2018.04.003)

- [2] D. Alonso-Gutierrez, J. Prochno and C. Thäle. Gaussian fluctuations for high-dimensional random projections of ℓ_p^n -balls. *Bernoulli*. To appear, 2019.
- [3] G. W. Anderson, A. Guionnet and O. Zeitouni. *An Introduction to Random Matrices*. Cambridge Studies in Advanced Mathematics **118**. Cambridge University Press, Cambridge, 2010. [MR2760897](#)
- [4] M. Anttila, K. Ball and I. Perissinaki. The central limit problem for convex bodies. *Trans. Amer. Math. Soc.* **355** (12) (2003) 4723–4735. [MR1997580](#) <https://doi.org/10.1090/S0002-9947-03-03085-X>
- [5] S. Artstein-Avidan, A. Giannopoulos and V. D. Milman. *Asymptotic Geometric Analysis. Part I. Mathematical Surveys and Monographs* **202**. American Mathematical Society, Providence, RI, 2015. [MR3331351](#) <https://doi.org/10.1090/surv/202>
- [6] F. Barthe and D. Cordero-Erausquin. Invariances in variance estimates. *Proc. Lond. Math. Soc.* (3) **106** (1) (2013) 33–64. [MR3020738](#) <https://doi.org/10.1112/plms/pds011>
- [7] G. Ben Arous and A. Guionnet. Large deviations for Wigner’s law and Voiculescu’s non-commutative entropy. *Probab. Theory Related Fields* **108** (4) (1997) 517–542. [MR1465640](#) <https://doi.org/10.1007/s004400050119>
- [8] S. Brazitikos, A. Giannopoulos, P. Valettas and B.-H. Vritsiou. *Geometry of Isotropic Convex Bodies*. Mathematical Surveys and Monographs **196**. American Mathematical Society, Providence, RI, 2014. [MR3185453](#) <https://doi.org/10.1090/surv/196>
- [9] A. Dembo and O. Zeitouni. *Large Deviations. Techniques and Applications. Stochastic Modelling and Applied Probability* **38**. Springer-Verlag, Berlin, 2010. Corrected reprint of the second (1998) edition. [MR2571413](#) <https://doi.org/10.1007/978-3-642-03311-7>
- [10] F. den Hollander. *Large Deviations. Fields Institute Monographs* **14**. American Mathematical Society, Providence, RI, 2000. [MR1739680](#)
- [11] N. Gantert, S. S. Kim and K. Ramanan. Large deviations for random projections of ℓ^p balls. *Ann. Probab.* **45** (6B) (2017) 4419–4476. [MR3737915](#) <https://doi.org/10.1214/16-AOP1169>
- [12] Y. Gordon and D. R. Lewis. Absolutely summing operators and local unconditional structures. *Acta Math.* **133** (1974) 27–48. [MR0410341](#) <https://doi.org/10.1007/BF02392140>
- [13] O. Guédon. Concentration phenomena in high dimensional geometry. In *Journées MAS 2012* 47–60. *ESAIM Proc.* **44**. EDP Sci., Les Ulis, 2014. [MR3178607](#) <https://doi.org/10.1051/proc/201444002>
- [14] O. Guédon, P. Nayar and T. Tkocz. Concentration inequalities and geometry of convex bodies. In *Analytical and Probabilistic Methods in the Geometry of Convex Bodies* 9–86. *IMPAN Lect. Notes* **2**. Polish Acad. Sci. Inst. Math., Warsaw, 2014. [MR3329056](#)
- [15] O. Guédon and G. Paouris. Concentration of mass on the Schatten classes. *Ann. Inst. Henri Poincaré Probab. Stat.* **43** (1) (2007) 87–99. [MR2288271](#) <https://doi.org/10.1016/j.anihpb.2006.01.002>
- [16] F. Hiai and D. Petz. Eigenvalue density of the Wishart matrix and large deviations. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **1** (4) (1998) 633–646. [MR1665279](#) <https://doi.org/10.1142/S021902579800034X>
- [17] F. Hiai and D. Petz. A large deviation theorem for the empirical eigenvalue distribution of random unitary matrices. *Ann. Inst. Henri Poincaré Probab. Stat.* **36** (1) (2000) 71–85. [MR1743092](#) [https://doi.org/10.1016/S0246-0203\(00\)00116-3](https://doi.org/10.1016/S0246-0203(00)00116-3)
- [18] F. Hiai and D. Petz. *The Semicircle Law, Free Random Variables and Entropy*. Mathematical Surveys and Monographs **77**. American Mathematical Society, Providence, RI, 2000. [MR1746976](#)
- [19] A. Hinrichs, J. Prochno and J. Vybíral. Entropy numbers of embeddings of Schatten classes. *J. Funct. Anal.* **273** (10) (2017) 3241–3261. [MR3695893](#) <https://doi.org/10.1016/j.jfa.2017.08.008>
- [20] Z. Kabluchko, J. Prochno and C. Thäle. Exact asymptotic volume and volume ratio of Schatten unit balls. Preprint, 2018. Available at [arXiv:1804.03467](#).
- [21] Z. Kabluchko, J. Prochno and C. Thäle. Intersection of unit balls in classical matrix ensembles. Preprint, 2018. Available at [arXiv:1804.03466](#).
- [22] Z. Kabluchko, J. Prochno and C. Thäle. High-dimensional limit theorems for random vectors in ℓ_p^n -balls. *Commun. Contemp. Math.* **21** (1) (2019) 1750092. [MR3904638](#) <https://doi.org/10.1142/S0219199717500924>
- [23] O. Kallenberg. *Foundations of Modern Probability*, 2nd edition. Probability and Its Applications. Springer-Verlag, New York, 2002. [MR1876169](#) <https://doi.org/10.1007/978-1-4757-4015-8>
- [24] S. S. Kim. Problems at the interface of probability and convex geometry: Random projections and constrained processes. Ph.D. thesis, Brown University, 2017.
- [25] S. S. Kim and K. Ramanan. A conditional limit theorem for high-dimensional ℓ^p -spheres. *J. Appl. Probab.* **55** (4) (2018) 1060–1077. [MR3899928](#) <https://doi.org/10.1017/jpr.2018.71>
- [26] A. Klenke. *Probability Theory: A Comprehensive Course*, 2nd edition. Universitext. Springer, London, 2014. [MR3112259](#) <https://doi.org/10.1007/978-1-4471-5361-0>
- [27] H. König, M. Meyer and A. Pajor. The isotropy constants of the Schatten classes are bounded. *Math. Ann.* **312** (4) (1998) 773–783. [MR1660231](#) <https://doi.org/10.1007/s002080050245>
- [28] Y. V. Prohorov. Convergence of random processes and limit theorems in probability theory. *Teor. Veroyatn. Primen.* **1** (1956) 177–238. [MR0084896](#)
- [29] J. Radke and B.-H. Vritsiou. On the thin-shell conjecture for the Schatten classes. Preprint, 2016. Available at [arXiv:1602.06934](#).
- [30] E. B. Saff and V. Totik. *Logarithmic Potentials with External Fields*. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences] **316**. Springer-Verlag, Berlin, 1997. Appendix B by Thomas Bloom. [MR1485778](#) <https://doi.org/10.1007/978-3-662-03329-6>
- [31] G. Schechtman and M. Schmuckenschläger. Another remark on the volume of the intersection of two L_p^n balls. In *Geometric Aspects of Functional Analysis (1989–90)* 174–178. *Lecture Notes in Math.* **1469**. Springer, Berlin, 1991. [MR1122622](#) <https://doi.org/10.1007/BFb0089224>
- [32] G. Schechtman and J. Zinn. On the volume of the intersection of two L_p^n balls. *Proc. Amer. Math. Soc.* **110** (1) (1990) 217–224. [MR1015684](#) <https://doi.org/10.2307/2048262>
- [33] S. Szarek and N. Tomczak-Jaegermann. On nearly Euclidean decomposition for some classes of Banach spaces. *Compos. Math.* **40** (3) (1980) 367–385. [MR0571056](#)
- [34] N. Tomczak-Jaegermann. The moduli of smoothness and convexity and the Rademacher averages of trace classes S_p ($1 \leq p < \infty$). *Studia Math.* **50** (1974) 163–182. [MR0355667](#)
- [35] D. Voiculescu. The analogues of entropy and of Fisher’s information measure in free probability theory. I. *Comm. Math. Phys.* **155** (1) (1993) 71–92. [MR1228526](#)
- [36] B.-H. Vritsiou. Selberg-type integrals and the variance conjecture for the operator norm. Preprint, 2018. Available at [arXiv:1805.03129](#).