

# Sharp asymptotics for $q$ -norms of random vectors in high-dimensional $\ell_p^n$ -balls

Tom Kaufmann

*Ruhr University Bochum, Universitätsstraße 150, 44801 Bochum, Germany*

[tom.kaufmann@rub.de](mailto:tom.kaufmann@rub.de) (T. Kaufmann)

Received: 29 January 2021, Revised: 5 May 2021, Accepted: 22 May 2021,  
Published online: 22 June 2021

**Abstract** Sharp large deviation results of Bahadur–Ranga Rao type are provided for the  $q$ -norm of random vectors distributed on the  $\ell_p^n$ -ball  $\mathbb{B}_p^n$  according to the cone probability measure or the uniform distribution for  $1 \leq q < p < \infty$ , thereby furthering previous large deviation results by Kabluchko, Prochno and Thäle in the same setting. These results are then applied to deduce sharp asymptotics for intersection volumes of different  $\ell_p^n$ -balls in the spirit of Schechtman and Schmuckenschläger, and for the length of the projection of an  $\ell_p^n$ -ball onto a line with uniform random direction. The sharp large deviation results are proven by providing convenient probabilistic representations of the  $q$ -norms, employing local limit theorems to approximate their densities, and then using geometric results for asymptotic expansions of Laplace integrals to integrate these densities and derive concrete probability estimates.

**Keywords** Asymptotic geometric analysis, Bahadur Ranga Rao, high-dimensional convexity, intersection volume,  $\ell_p^n$ -balls,  $\ell_p^n$ -spheres, large deviation principles, precise large deviations, sharp asymptotics, sharp large deviations, strong large deviations, volume of convex bodies

**2010 MSC** 52A23, 60F10, 46B09, 60D05

## 1 Introduction

The study of convex bodies in high dimensions, known today as asymptotic geometric analysis, has arisen from the local theory of Banach spaces, which aimed at analyzing infinite-dimensional normed spaces via their finite-dimensional substructures, such as their unit balls. Despite having its origin in the realm of functional analysis, the field

© 2021 The Author(s). Published by VTeX. Open access article under the [CC BY](https://creativecommons.org/licenses/by/4.0/) license.

has since established itself in its own right, considering problems also beyond the study of centrally symmetric convex bodies that occur naturally as the unit balls of Banach spaces. In high dimensions convex bodies exhibit certain regularities, such as concentration of measure phenomena (see, e.g., [23]), which make it highly useful to approach them from a probabilistic perspective. As pointed out in [6], it might seem counter-intuitive to analyze something exhibiting regularities from a probabilistic perspective, as probability concerns itself with studying the nature of irregularity, i.e. randomness, of given quantities. But as with well-known limit theorems from probability such as the law of large numbers and the central limit theorem, with large sample sizes (and analogously – with high dimensionality) random objects exhibit interesting patterns well characterized in the language of probability and vice versa. Many results analogous to those from classic probability have been found for high-dimensional convex sets, such as the central limit theorem (see, e.g., Anttila, Ball and Perissinaki [5], Klartag [34, 35]). For further background on high-dimensional convexity, see [6, 11, 23, 24].

The  $\ell_p^n$ -ball  $\mathbb{B}_p^n$ ,  $n \in \mathbb{N}$ , has been a prominent object of study, as it is the unit ball of the (finite-dimensional) sequence space  $\ell_p^n$ , and has been the subject of a multitude of results. We will name only a select few and refer to the survey by Prochno, Thäle and Turchi [40] for a comprehensive summary of classic and contemporary results. Let us denote by  $\mathbf{U}_{n,p}$  the uniform distribution on the Euclidean  $\ell_p^n$ -ball  $\mathbb{B}_p^n$  and by  $\mathbf{C}_{n,p}$  the cone probability measure on the  $\ell_p^n$ -sphere  $\mathbb{S}_p^{n-1}$ . Schechtman and Zinn [46] and Rachev and Rüschendorf [41] showed a generalization of the Poincaré–Maxwell–Borel lemma, proving that, for  $k \in \mathbb{N}$  with  $k < n$ , the  $k$ -dimensional marginal distribution of a random vector distributed according to  $\mathbf{C}_{n,p}$  converges to a  $k$ -dimensional generalized Gaussian distribution as  $n$  increases. They also provided a probabilistic representation for such random vectors in terms of these generalized Gaussian distributions, which will be a key building block in our main results. The primary quantity of interest of this paper however is the behaviour of the  $q$ -norm  $\|Z\|_q$  of a random vector  $Z$  in  $\mathbb{S}_p^{n-1}$  and  $\mathbb{B}_p^n$ . This was first studied by Schechtman and Zinn [46], who derived concentration inequalities for  $\|Z\|_q$  with  $Z \sim \mathbf{C}_{n,p}$  and  $Z \sim \mathbf{U}_{n,p}$  for  $q > p$ . This is closely related to the intersection volume of  $t$ -multiples of volume-normalized  $\ell_p^n$ -balls  $\mathbb{D}_p^n := \text{vol}_n(\mathbb{B}_p^n)^{-1/n} \mathbb{B}_p^n$ , i.e.  $\text{vol}_n(\mathbb{D}_p^n \cap t\mathbb{D}_p^n)$  with  $t \in [0, \infty)$ , for which Schechtman and Schmuckenschläger [45] gave the asymptotics for  $t \neq 1$ . Schechtman and Zinn [47] expanded their previous results in [46], by not only considering the  $q$ -norm, but also images of random vectors under Lipschitz functions in general. Thus, they gave concentration inequalities for  $f(Z)$ , with  $Z \sim \mathbf{C}_{n,p}$  and  $Z \sim \mathbf{U}_{n,p}$ ,  $p \in [1, 2)$ , and  $f$  a Lipschitz function with respect to the Euclidean norm. Schmuckenschläger [48] provided a central limit theorem (CLT) for  $\|Z\|_q$  with  $Z \sim \mathbf{C}_{n,p}$  and  $Z \sim \mathbf{U}_{n,p}$  and used it to refine the previous intersection results in [45] for all  $t \in (0, \infty)$ . Naor [38] gave concentration inequalities for  $\|Z\|_q^q$  with  $Z \sim \mathbf{C}_{n,p}$ , showed that the total variation distance between  $\mathbf{C}_{n,p}$  and the normalized surface measure  $\sigma_{n,p}$  on  $\mathbb{S}_p^{n-1}$  tends to zero proportional to  $n^{-1/2}$ , and used the previously mentioned results to show a concentration inequality for  $\|Z\|_q^q$  with  $Z \sim \sigma_{n,p}$ . He also discussed how concentration results similar to Schechtman and Zinn [47] for  $\|Z\|_q$  could already be derived from previous results of Gromov and Milman [22] for the concentration of Lipschitz functions on convex bodies. Kabluchko, Prochno

and Thäle [28] gave a multivariate CLT for  $(\|Z\|_{q_1}, \dots, \|Z\|_{q_d})$  with  $Z \sim \mathbf{U}_{n,p}$  in the spirit of [48] and also considered the asymptotics for the intersection volume of multiple  $\ell_p^n$ -balls, i.e.  $\text{vol}_n(\mathbb{D}_p^n \cap t_1 \mathbb{D}_{q_1}^n \cap \dots \cap t_d \mathbb{D}_{q_d}^n)$  with  $t_i \in [0, \infty)$ . This CLT was furthermore applied by the same authors to infer a central limit theorem for the length of  $\mathbb{B}_p^n$  projected onto a line with uniform random direction. Moreover, they provided a large deviation principle (LDP) for  $\|Z\|_q$  with  $Z \sim \mathbf{C}_{n,p}$  and  $Z \sim \mathbf{U}_{n,p}$ . In a follow-up paper [30], the same authors showed a CLT for  $\|Z\|_q$ , where the distribution of  $Z$  is taken from a wider class of  $p$ -radial distributions  $\mathbf{P}_{n,p,\mathbf{W}}$ , introduced by Barthe, Guédon, Mendelson and Naor [8], consisting of mixtures of  $\mathbf{U}_{n,p}$  and  $\mathbf{C}_{n,p}$ , combined via a measure  $\mathbf{W}$  on  $[0, \infty)$ . This class contains both  $\mathbf{U}_{n,p}$  and  $\mathbf{C}_{n,p}$ , but also distributions corresponding with geometrically interesting projections (see, e.g., [30, Introduction, (iii)]). Finally, they gave a moderate and a large deviation principle for  $\|Z\|_q$  with  $Z \sim \mathbf{P}_{n,p,\mathbf{W}}$ .

Generally, studying large deviations within asymptotic geometric analysis has started fairly recently with Gantert, Kim and Ramanan [20], who gave an LDP for projections of random points in  $\ell_p^n$ -balls with distributions  $\mathbf{C}_{n,p}$  and  $\mathbf{U}_{n,p}$  onto both random and fixed one-dimensional subspaces. Today, large deviations theory has become a well-established toolbox in high-dimensional convex geometry, giving rise to a plethora of results (see, e.g., [3, 4, 28–30, 32, 33]). Recently, a new tool from large deviations theory was introduced to asymptotic geometric analysis by Liao and Ramanan [37]. They gave sharp large deviation (SLD) results in the spirit of Bahadur and Ranga Rao [7] and Petrov [39] for the projections of random points in  $\ell_p^n$ -balls with distributions  $\mathbf{C}_{n,p}$  and  $\mathbf{U}_{n,p}$  onto a fixed one-dimensional subspace. Other works in asymptotic geometric analysis have also employed methods from sharp large deviations theory as well, such as Kabluchko and Prochno [27], who derived asymptotic volumes for generalizations of  $\ell_p^n$ -balls, known as Orlicz balls, and showed a Schechtman and Schmuckenschläger-type result by considering intersection volumes of Orlicz balls. Their results on Orlicz balls were then expanded upon by Alonso-Gutiérrez and Prochno in [2], who gave the exact asymptotic volume of Orlicz balls and provided thin-shell concentrations for them, augmenting their results into sharp asymptotics under certain conditions. While LDPs only give tail asymptotics on a logarithmic scale, the sharp asymptotics provided by sharp large deviations theory can give tail estimates for concrete values of  $n \in \mathbb{N}$ , which makes them significantly more useful for practical applications. Moreover, a lot of idiosyncrasies of the underlying distributions, that are drowned out on the LDP scale, are still visible on the SLD scale, thus giving a deeper understanding of the geometric interpretation of the quantities involved. This paper will follow closely in the footsteps of Liao and Ramanan [37] and establish SLD results for the  $q$ -norms of random vectors with distribution  $\mathbf{C}_{n,p}$  and  $\mathbf{U}_{n,p}$ . Furthermore, we will use these results to expand on works of Schechtman and Schmuckenschläger [45], Schmuckenschläger [48], and Kabluchko, Prochno and Thäle [28] for intersection volumes of  $\ell_p^n$ -balls by giving sharp asymptotics for  $\text{vol}_n(\mathbb{D}_p^n \cap t \mathbb{D}_q^n)$  at a considerably improved rate for  $1 \leq q < p < \infty$  and  $t > C(p, q)$  bigger than some constant dependent on  $p$  and  $q$  only. Additionally, we will also apply our results for  $\ell_p^n$ -spheres to retain sharp asymptotics for the length of the projection of an  $\ell_p^n$ -ball onto the line spanned by a uniform random direction.

The paper will proceed as follows: in Section 2 some basic notation and definitions will be provided while also giving some appropriate background on the involved large deviations theory. Furthermore, we will recapitulate some existing results that are relevant to this paper. In Section 3 we will present our main results regarding the  $q$ -norms of random vectors on  $\ell_p^n$ -spheres and  $\ell_p^n$ -balls. Also, we will present and prove their application to intersections and one-dimensional projections of  $\ell_p^n$ -balls, and outline the idea of the two central proofs. In Section 4 we will reformulate the target probabilities from the main results in terms of useful probabilistic representations, using well-established representations of random vectors in  $\ell_p^n$ -balls of Schechtman and Zinn [46] and Rachev and Rüschendorf [41]. In Section 5 local density approximations of these probabilistic representations will be provided. In Sections 6 and 7 we will prove the SLD results for  $\ell_p^n$ -spheres and  $\ell_p^n$ -balls, respectively, by integrating over the density estimates. For that, we will utilize some geometric results for asymptotic expansions of Laplace integrals from Adriani and Baldi [1] and Breitung and Hohenbichler [12].

## 2 Preliminaries

### 2.1 Notation and important distributions

We denote by  $\text{vol}_d$  the  $d$ -dimensional Lebesgue measure on  $\mathbb{R}^d$  and write  $\mathcal{B}(\mathbb{R}^d)$  for the  $\sigma$ -field of Borel sets in  $\mathbb{R}^d$ . For a set  $A \in \mathcal{B}(\mathbb{R}^d)$  we write  $A^\circ$ ,  $\bar{A}$ ,  $\partial A$ , and  $A^c$  for the interior, closure, boundary and complement of  $A$ , respectively. Furthermore, we write  $\langle \cdot, \cdot \rangle$  for the standard scalar product in  $\mathbb{R}^d$ . For  $g : \mathbb{R}^d \rightarrow \mathbb{R}^d$ , we denote by  $J_x g(x^*)$  the Jacobian of  $g$  with respect to the vector  $x$  evaluated at  $x^* \in \mathbb{R}^d$ , and for  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  by  $\nabla_x f(x^*)$  and  $\mathcal{H}_x f(x^*)$  the gradient and the Hessian of  $f$  with respect to the vector  $x$  evaluated at  $x^* \in \mathbb{R}^d$ , respectively, and use the shorthand notation

$$f_{[i_1, \dots, i_d]}(x^*) = \frac{\partial^{i_1}}{\partial x_1^{i_1}} \dots \frac{\partial^{i_d}}{\partial x_1^{i_d}} f(x) \Big|_{x=x^*}. \quad (1)$$

We write  $(x_1, \dots, x_d) \in \mathbb{R}^d$  for a standard column vector and for  $x, y \in \mathbb{R}^d$ , we write their product  $x^T y$  as  $xy$ , skipping the explicit transpose notation for brevity. Given a random variable  $X$  with distribution  $\mathbf{P}$ , we write  $X \sim \mathbf{P}$  and denote by  $\mathbb{E}X$  its expectation. For two random variables  $X, Y$  with the same distribution we write  $X \stackrel{d}{=} Y$ . For a random vector  $X$  in  $\mathbb{R}^d$  and  $s \in \mathbb{R}^d$ , denote by  $\varphi_X(s) := \mathbb{E}[e^{(s, X)}]$  and  $\Lambda_X(s) := \log \varphi_X(s)$  the moment generating function and cumulant generating function (m.g.f. and c.g.f.), respectively. We call the set of  $s \in \mathbb{R}^d$  for which  $\Lambda_X(s) < \infty$  the effective domain  $\mathcal{D}_X$  of  $\Lambda_X$ . Moreover, for  $x \in \mathbb{R}^d$  we denote by  $\Lambda_X^*(x) := \sup_{s \in \mathbb{R}^d} \langle x, s \rangle - \Lambda_X(s)$  the Legendre–Fenchel transform of the c.g.f.  $\Lambda_X$ . When considering sequences in  $n \in \mathbb{N}$ , we denote by  $o(1)$  a sequence that tends to zero as  $n \rightarrow \infty$ .

Let us consider the class of distributions at the core of the probabilistic constructions throughout this paper. We say a real-valued random variable  $X$  has a generalized Gaussian distribution if its distribution has the Lebesgue density

$$f_{\text{gen}}(x) := \frac{b}{2a \Gamma(\frac{1}{b})} e^{-(|x-m|/a)^b}, \quad x \in \mathbb{R},$$

where  $m \in \mathbb{R}$  and  $a, b > 0$ , and denote this by  $X \sim \mathbf{N}_{\text{gen}}(m, a, b)$ . As mentioned in the introduction, the generalized Gaussian distributions are essential for constructing probabilistically equivalent representations of the quantities of interest, based on results of Schechtman and Zinn [46] and Rachev and Rüschendorf [41]. For these constructions we will be using the specific generalized Gaussian distribution  $\mathbf{N}_p := \mathbf{N}_{\text{gen}}(0, p^{1/p}, p)$ ,  $p \in [1, \infty)$ , with the density

$$f_p(x) := \frac{1}{2 p^{1/p} \Gamma(1 + \frac{1}{p})} e^{-|x|^p/p}, \quad x \in \mathbb{R}.$$

For  $X \sim \mathbf{N}_p$  and  $r > 0$ , we write  $M_p(r) := \mathbb{E}|X|^r$  for the  $r$ -th absolute moment of  $X$ , for which

$$M_p(r) := \mathbb{E}|X|^r = \left( \frac{p^{r/p}}{r+1} \frac{\Gamma(1 + \frac{r+1}{p})}{\Gamma(1 + \frac{1}{p})} \right). \tag{2}$$

### 2.2 Background material from (sharp) large deviations theory

We will give some basic notions and definitions from large deviations theory. To keep this paper self-contained, we will present them here, while referring the reader to [17, 18, 31] for additional background material on large deviations. Furthermore, we want to give some insight into the methods of the lesser known theory of sharp large deviations.

**Definition 1.** Let  $(\mathbf{P}_n)_{n \in \mathbb{N}}$  be a sequence of probability measures on  $\mathbb{R}^d$ . We say that  $(\mathbf{P}_n)_{n \in \mathbb{N}}$  satisfies a large deviation principle (LDP) if there are two functions  $s : \mathbb{N} \rightarrow \mathbb{R}$  and  $\mathcal{I} : \mathbb{R}^d \rightarrow [0, \infty]$ , such that  $\mathcal{I}$  is lower semi-continuous and

- a)  $\limsup_{n \rightarrow \infty} \frac{1}{s(n)} \log \mathbf{P}_n(C) \leq -\mathcal{I}(C) \quad \text{for all } C \subset \mathbb{R}^n \text{ closed,}$
- b)  $\liminf_{n \rightarrow \infty} \frac{1}{s(n)} \log \mathbf{P}_n(O) \geq -\mathcal{I}(O) \quad \text{for all } O \subset \mathbb{R}^n \text{ open,}$

where for  $B \subset \mathbb{R}^d$  we define  $\mathcal{I}(B) := \inf_{x \in B} \mathcal{I}(x)$ . We call  $s$  the speed and  $\mathcal{I}$  the rate function. We say that  $\mathcal{I}$  is a good rate function, if it has compact sub-level sets.

We apply the definition of LDPs to sequences of random variables as well by applying the above definition to the sequence of their distributions. In our setting the sequence parameter  $n \in \mathbb{N}$  will furthermore coincide with the space dimension  $d \in \mathbb{N}$ , as we are considering the effects of increasing dimensionality. Given a sequence  $(X^{(n)})_{n \in \mathbb{N}}$  of i.i.d. random vectors in  $\mathbb{R}^d$ , one is frequently interested in the behaviour of the sequence  $(S^{(n)})_{n \in \mathbb{N}}$  of the empirical averages  $S^{(n)} := \frac{1}{n} \sum_{i=1}^n X^{(i)} \in \mathbb{R}^d$ . One of the most well known and most frequently used results in the theory of large deviations is the theorem of Cramér, which states that if the c.g.f.  $\Lambda_X$  is finite in an open neighbourhood of the origin, then  $(S^{(n)})_{n \in \mathbb{N}}$  satisfies an LDP in  $\mathbb{R}^d$  with speed  $n$  and rate function  $\Lambda_X^*$  (see, e.g., [17, Theorem 2.2.30, Theorem 6.1.3, Corollary 6.1.6]). Hence, under suitable exponential moment assumptions for the  $X^{(n)}$ , we can already infer the large deviation behaviour of  $(S^{(n)})_{n \in \mathbb{N}}$ .

The classic LDP gives us an idea of the asymptotic deviation behaviour of a sequence of distributions on a logarithmic scale. By doing this however, a lot of subtleties of the underlying distributions can be drowned out. Many small and medium scale properties of a given sequence of distributions are often missed in the asymptotic analysis of LDPs, since they either disappear for very large  $n \in \mathbb{N}$  or are drowned out by other, more significant phenomena of the distribution. Thus, one is also interested in considering large deviations on a nonlogarithmic scale, which we refer to as “sharp” large deviations (also called “precise” or “strong” large deviations in the literature). One of the first and most prominent results in this regard was shown by Bahadur and Ranga Rao [7]. They showed that for a sequence  $(X^{(n)})_{n \in \mathbb{N}}$  of i.i.d. random variables and any  $z > \mathbb{E}[X^{(n)}]$  with  $\Lambda_X^*(z) < \infty$ , it holds that

$$\mathbb{P}\left(S^{(n)} > z\right) = \frac{1}{\sqrt{2\pi n} \kappa(z) \xi(z)} e^{-n\Lambda_X^*(z)}(1 + o(1)),$$

where  $\kappa(z)$  and  $\xi(z)$  are only dependent on the distribution of the  $X^{(n)}$  and the deviation size  $z$ . This is proven via a (somewhat implicit) application of the so-called saddle point method (or method of steepest descents), which was established by Debye [16], and brought to the realm of probability by Esscher [19] and Daniels [14]. The saddle point method generalizes Laplace’s method for integral approximation to the complex plane, and is therefore highly useful when dealing with integrals of characteristic functions. In general, for appropriate functions  $f, g$  and  $n \in \mathbb{N}$  large, the saddle point method gives a way to approximate Laplace-type integrals  $\int_P g(z)e^{-nf(z)}dz$  along complex paths  $P$ , by deforming the integration path using Cauchy’s theorem, into some  $\tilde{P}$  that passes through a saddle point of  $f$ . The mass of the reformulated integral is then heavily concentrated around the saddle point and standard integral expansion methods, such as Edgeworth expansion, can be used to great effect. In the realm of probability, this has been used for both tail probabilities (e.g., Esscher [19], Cramér [13]) and densities of random variables (e.g., Daniels [14], Richter [42, 43]), by writing them as an integral of their characteristic functions, using the Fourier inversion formula, and then approximating those integrals via the use of a complex saddle point. We say that this was used “somewhat implicitly” in certain results, such as those of Esscher [19], Cramér [13] and Bahadur and Ranga Rao [7], since the technique used therein, which is a certain change of measure, often called exponential tilting or Esscher/Cramér transform, under the surface employs saddle points as well. For further background on this method, we refer to the book of Jensen [26].

As mentioned in the introduction, Section 5 will provide density estimates for our probabilistic representations from Section 4, which are derived using the saddle point method. However, since our probabilistic representations are given as sums of i.i.d. random vectors, we will refer to previous results where this was done explicitly, while making sure that the conditions for their application are still met in our setting. Generally, the core idea of the saddle point method, which is reformulating an integral so that all of its mass heavily concentrates around a critical point, around which we can then employ approximation methods, is used in the overall proof of our main results in a broader sense as well. We reformulate our target probabilities via some convenient representations, whose densities we also provide, such that the remaining

integrals then heavily concentrate their mass around a given critical point, such that approximations at that point yield accurate results, as we will see in Sections 6 and 7.

### 2.3 Distributions on $\ell_p^n$ -balls

For  $p \in [1, \infty]$ ,  $n \in \mathbb{N}$ , and  $x = (x_1, \dots, x_n) \in \mathbb{R}^n$  we denote by

$$\|x\|_p := \begin{cases} \left(\sum_{i=1}^n |x_i|^p\right)^{1/p} & : p < \infty \\ \max\{|x_1|, \dots, |x_n|\} & : p = \infty \end{cases} \quad (3)$$

the  $\ell_p^n$ -norm of  $x$ . Let  $\mathbb{B}_p^n := \{x \in \mathbb{R}^n : \|x\|_p \leq 1\}$  be the unit  $\ell_p^n$ -ball and  $\mathbb{S}_p^{n-1} := \{x \in \mathbb{R}^n : \|x\|_p = 1\}$  be the unit  $\ell_p^n$ -sphere. We define the uniform distribution on  $\mathbb{B}_p^n$  and cone probability measure on  $\mathbb{S}_p^{n-1}$  as

$$\mathbf{U}_{n,p}(\cdot) := \frac{\text{vol}_n(\cdot)}{\text{vol}_n(\mathbb{B}_p^n)} \quad \text{and} \quad \mathbf{C}_{n,p}(\cdot) := \frac{\text{vol}_n(\{rx : r \in [0, 1], x \in \cdot\})}{\text{vol}_n(\mathbb{B}_p^n)}.$$

The following result is the basis of our probabilistic representations for random vectors with distributions  $\mathbf{C}_{n,p}$  and  $\mathbf{U}_{n,p}$  and is due to [41] and [46].

**Lemma 2.** *Let  $p \in [1, \infty)$ ,  $Y = (Y_1, \dots, Y_n)$  be a random vector in  $\mathbb{R}^n$  with  $Y_i \sim \mathbf{N}_p$  i.i.d., and  $U$  be an independent random variable uniformly distributed on  $[0, 1]$ . Then,*

- i) *the random vector  $Y/\|Y\|_p$  has distribution  $\mathbf{C}_{n,p}$  and is independent of  $\|Y\|_p$ ,*
- ii) *the random vector  $U^{1/n} Y/\|Y\|_p$  has distribution  $\mathbf{U}_{n,p}$ .*

### 2.4 LDPs for $q$ -norms in $\ell_p^n$ -balls

Throughout this paper we assume  $1 \leq q < p < \infty$ . The main variables of interest will be the  $q$ -norms of the random vectors  $Z^{(n)}, \mathcal{Z}^{(n)} \in \mathbb{B}_p^n$  with  $Z^{(n)} \sim \mathbf{C}_{n,p}$  and  $\mathcal{Z}^{(n)} \sim \mathbf{U}_{n,p}$ . Note, that we will always denote quantities related to  $\mathcal{Z}^{(n)} \sim \mathbf{U}_{n,p}$  cursorily. To get nontrivial results, our target variables also need to be appropriately rescaled. Thus, for random vectors  $Z^{(n)}, \mathcal{Z}^{(n)} \in \mathbb{B}_p^n$  with  $Z^{(n)} \sim \mathbf{C}_{n,p}$  and  $\mathcal{Z}^{(n)} \sim \mathbf{U}_{n,p}$ , our target variables will be  $n^{1/p-1/q} \|Z^{(n)}\|_q$  and  $n^{1/p-1/q} \|\mathcal{Z}^{(n)}\|_q$ , respectively. We set

$$\|Z\| := \left(n^{1/p-1/q} \|Z^{(n)}\|_q\right)_{n \in \mathbb{N}} \quad \text{and} \quad \|\mathcal{Z}\| := \left(n^{1/p-1/q} \|\mathcal{Z}^{(n)}\|_q\right)_{n \in \mathbb{N}}.$$

It follows via the strong law of large numbers and the continuous mapping theorem applied to the probabilistic representations in (4) and (5) that the expectations of  $\|Z\|$  and  $\|\mathcal{Z}\|$  converge to  $m_{p,q} := M_p(q)^{1/q}$  as  $n \in \mathbb{N}$  increases. For fixed  $n \in \mathbb{N}$  we will denote

$$\mathbb{E} \left[ n^{1/p-1/q} \|Z^{(n)}\|_q \right] := m_{n,p,q} \quad \text{and} \quad \mathbb{E} \left[ n^{1/p-1/q} \|\mathcal{Z}^{(n)}\|_q \right] := m_{n,p,q}.$$

Furthermore, LDPs for  $\|Z\|$  and  $\|\mathcal{Z}\|$  have been given in previous works, which we want to include here explicitly. But first, let us look at the following probabilistic

representations of  $\|Z\|$  and  $\|\mathcal{Z}\|$ , since the LDPs are given with respect to the c.g.f. of these representations: Let  $(Y^{(n)})_{n \in \mathbb{N}}$  be a sequence of i.i.d. random vectors  $Y^{(n)} := (Y_1^{(n)}, \dots, Y_n^{(n)})$  with  $Y_i^{(n)} \sim \mathbf{N}_p$ , and  $U$  be a random variable, independent of the  $Y_i^{(n)}$ , and uniformly distributed on  $[0, 1]$ . Then, we can see via Lemma 2 that

$$n^{1/p-1/q} \|Z^{(n)}\|_q \stackrel{d}{=} n^{1/p-1/q} \frac{\|Y^{(n)}\|_q}{\|Y^{(n)}\|_p} = \frac{\left(\frac{1}{n} \sum_{i=1}^n |Y_i^{(n)}|^q\right)^{1/q}}{\left(\frac{1}{n} \sum_{i=1}^n |Y_i^{(n)}|^p\right)^{1/p}}, \tag{4}$$

and

$$n^{1/p-1/q} \|\mathcal{Z}^{(n)}\|_q \stackrel{d}{=} n^{1/p-1/q} U^{1/n} \frac{\|Y^{(n)}\|_q}{\|Y^{(n)}\|_p} = U^{1/n} \frac{\left(\frac{1}{n} \sum_{i=1}^n |Y_i^{(n)}|^q\right)^{1/q}}{\left(\frac{1}{n} \sum_{i=1}^n |Y_i^{(n)}|^p\right)^{1/p}}. \tag{5}$$

Define

$$V^{(n)} := \left(V_1^{(n)}, \dots, V_n^{(n)}\right) \in \mathbb{R}^{2n} \tag{6}$$

with

$$V_i^{(n)} := \left(|Y_i^{(n)}|^q, |Y_i^{(n)}|^p\right),$$

and

$$\mathcal{V}^{(n)} := \left(\mathcal{V}_1^{(n)}, \dots, \mathcal{V}_n^{(n)}\right) \in \mathbb{R}^{3n} \tag{7}$$

with

$$\mathcal{V}_i^{(n)} := \left(|Y_i^{(n)}|^q, |Y_i^{(n)}|^p, U^{1/n}\right).$$

We denote the m.g.f. and c.g.f. of the  $V_i^{(n)}$  as

$$\varphi_p(\tau) := \int_{\mathbb{R}} e^{\tau_1|y|^q + \tau_2|y|^p} f_p(y) dy,$$

and

$$\Lambda_p(\tau) := \log \int_{\mathbb{R}} e^{\tau_1|y|^q + \tau_2|y|^p} f_p(y) dy, \tag{8}$$

for  $\tau = (\tau_1, \tau_2) \in \mathbb{R}^2$ , and the Legendre–Fenchel transform of  $\Lambda_p$  as

$$\Lambda_p^*(x) := \sup_{\tau \in \mathbb{R}^2} [\langle x, \tau \rangle - \Lambda_p(\tau)], \quad x \in \mathbb{R}^2.$$

Let  $\mathcal{D}_p$  be the effective domain of  $\Lambda_p$ . Since  $q < p$ , for the integral in both  $\varphi_p$  and  $\Lambda_p$  to be finite, the sign of the dominant term in the exponent must be negative. Remembering the definition of  $f_p$ , one can see that this is valid for  $\tau_2 < 1/p$ , thus  $\mathcal{D}_p = \mathbb{R} \times (-\infty, 1/p)$ . Now, we want to characterize the points  $x \in \mathbb{R}^2$  for which there exists a  $\tau(x) \in \mathcal{D}_p$ , such that

$$\Lambda_p^*(x) = \langle x, \tau(x) \rangle - \Lambda_p(\tau(x)), \tag{9}$$



i.e. for which the function  $g_x(\tau) := \langle x, \tau \rangle - \Lambda_p(\tau)$  does not attain its supremum at infinity. We will do this along the lines of [1, Section 2]. It holds that  $\Lambda_p$  is convex in  $\tau$  (see standard properties of the c.g.f. in, e.g., [17, Lemma 2.2.31]), hence  $g_x(\tau) := \langle x, \tau \rangle - \Lambda_p(\tau)$  is concave as a sum of concave functions. Then, for a given  $x \in \mathbb{R}^2$ , there either exists a  $\tau(x) \in \mathbb{R}^2$  that satisfies  $x = \nabla_\tau \Lambda_p(\tau)$ , i.e. that is a root of  $\nabla_\tau g_x(\tau)$ , or the supremum of  $g_x$  will be attained at infinity. If such a  $\tau(x)$  exists and lies in  $\mathcal{D}_p$ , then  $\Lambda_p^*(x) = \langle x, \tau(x) \rangle - \Lambda_p(\tau(x)) < \infty$  (see [17, Lemma 2.2.31]). Since the  $V_i^{(n)}$  are not concentrated on a hyperplane (as  $1 \leq q < p < \infty$ ), the covariance matrix of their distribution given by  $\mathcal{H}_\tau \Lambda_p(0, 0)$  is positive definite and thereby invertible.  $\mathcal{H}_\tau \Lambda_p(\tau)$  for  $\tau \in \mathcal{D}_p$  can also be interpreted as the covariance matrix of an exponentially shifted distribution of  $V_i^{(n)}$  (see [1, p. 374]), which, by the same argument, is also not concentrated on any hyperplane, hence the  $\mathcal{H}_\tau \Lambda_p(\tau)$  are positive definite as well. This implies that  $\Lambda_p(\tau)$  is strictly convex on  $\mathcal{D}_p$ , thereby also making  $g_x$  strictly concave on  $\mathcal{D}_p$ . Hence, the strict concavity of  $g_x$  then ensures that  $\tau(x)$  is unique in  $\mathcal{D}_p$  in the above described property. We denote the set of  $x \in \mathbb{R}^2$  for which such a  $\tau(x) \in \mathcal{D}_p$  exists as  $\mathcal{J}_p$  and call it the admissible domain of  $\Lambda_p^*$ .

**Remark 3.** Note that the admissible domain  $\mathcal{J}_p$  is the image of  $\mathcal{D}_p$  under the derivative of the c.g.f.  $\Lambda_p$ . It actually holds that  $\nabla_\tau \Lambda_p(\tau)$  is a bijection from the interior of the effective domain of  $\Lambda_p$  into the interior of the effective domain of  $\Lambda_p^*$ , by the properties of the Legendre transform (see [44, Theorem 26.5]). Since  $\mathcal{D}_p$  is open and  $\nabla_\tau \Lambda_p$  continuous on  $\mathcal{D}_p$ , we thereby get that the effective domain of  $\Lambda_p^*$  is also open and thus,  $\mathcal{J}_p$  is simply the effective domain of  $\Lambda_p^*$ .

For the sequence  $\|Z\|$  the following LDP has already been shown by Kabluchko, Prochno and Thäle [28, Section 5.1]:

**Proposition 4.** *Let  $1 \leq q < p < \infty$  and  $Z^{(n)} \sim \mathbf{C}_{n,p}$  be a random vector in  $\mathbb{S}_p^{n-1}$ . Then the sequence  $(n^{1/p-1/q} \|Z^{(n)}\|_q)_{n \in \mathbb{N}}$  satisfies an LDP with speed  $n$  and good rate function*

$$\mathcal{I}_{\|Z\|}(z) := \begin{cases} \inf_{\substack{t_1, t_2 > 0 \\ t_1^{1/q} t_2^{-1/p} = z}} \Lambda_p^*(t_1, t_2) & : z > 0 \\ +\infty & : z \leq 0. \end{cases}$$

In [37, Lemma 2.1, Appendix A] Liao and Ramanan established a simplification of a similar rate function in a different setting. Their arguments can be analogously applied in our setting to derive the following result.

**Lemma 5.** *Let  $z > m_{p,q}$  be such that  $z^* := (z^q, 1) \in \mathcal{J}_p$ . Then*

$$\mathcal{I}_{\|Z\|}(z) = \inf_{\substack{t_1, t_2 > 0 \\ t_1^{1/q} t_2^{-1/p} = z}} \Lambda_p^*(t_1, t_2) = \Lambda_p^*(z^*),$$

with  $z^*$  being the unique point at which  $\Lambda_p^*$  attains its infimum under the above conditions.

To keep this paper self-contained, we will present the analogous proof of this in the Appendix. For the sequence  $\|\mathcal{Z}\|$ , the following LDP was also provided by Kabluchko, Prochno and Thäle in [28, Theorem 1.2].

**Proposition 6.** *Let  $1 \leq q < p < \infty$  and  $\mathcal{X}^{(n)} \sim \mathbf{U}_{n,p}$  be a random vector in  $\mathbb{B}_p^n$ . Then the sequence  $(n^{1/p-1/q} \|\mathcal{X}^{(n)}\|_q)_{n \in \mathbb{N}}$  satisfies an LDP with speed  $n$  and good rate function*

$$\mathcal{I}_{\|\mathcal{X}\|}(z) := \begin{cases} \inf_{\substack{z = z_1 z_2 \\ z_1, z_2 > 0}} [\mathcal{I}_{\|Z\|}(z_1) + \mathcal{I}_U(z_2)] & : z > 0 \\ +\infty & : z \leq 0, \end{cases}$$

with  $\mathcal{I}_{\|Z\|}$  as in Proposition 4 and

$$\mathcal{I}_U(z_2) := \begin{cases} -\log(z_2) & : z_2 \in (0, 1] \\ +\infty & : \text{otherwise.} \end{cases}$$

We again show that the above infimum is attained at a unique point satisfying the infimum condition.

**Lemma 7.** *Assume the same setting as in Proposition 6. For  $z > m_{p,q}$ , we can simplify the rate function by combining the two infimum operations to get*

$$\mathcal{I}_{\|\mathcal{X}\|}(z) = \inf_{\substack{z = t_1^{1/q} t_2^{-1/p} t_3 \\ t_1, t_2 > 0, t_3 \in (0, 1)}} [\Lambda_p^*(t_1, t_2) - \log(t_3)].$$

We define

$$\mathcal{I}_{\mathcal{S}}(t) := [\Lambda_p^*(t_1, t_2) - \log(t_3)], \quad t_1, t_2 \in \mathbb{R}, \quad t_3 \in (0, 1),$$

and set  $z^* := (z^q, 1) \in \mathbb{R}^2$ ,  $z^{**} := (z^q, 1, 1) \in \mathbb{R}^3$ . It then holds for  $z > m_{p,q}$  with  $z^* \in \mathcal{J}_p$  that

$$\mathcal{I}_{\|\mathcal{X}\|}(z) = \mathcal{I}_{\mathcal{S}}(z^{**}) = \Lambda_p^*(z^*),$$

with  $z^{**}$  being the unique point at which  $\mathcal{I}_{\mathcal{S}}$  attains its infimum under the above conditions.

Thus, for  $z > m_{p,q}$  with  $z^* \in \mathcal{J}_p$  both  $\|Z\|$  and  $\|\mathcal{X}\|$  satisfy LDPs with the same speed and rate function. Again, the proof of this is relegated to the Appendix.

**Remark 8.** Note that in the results within this paper, deviations from the ‘‘limit expectation’’  $m_{p,q}$  are considered, even though the sequences  $\|Z\|$  and  $\|\mathcal{X}\|$  have respective expectations  $m_{n,p,q}$  and  $\mathcal{m}_{n,p,q}$ , that only converge to  $m_{p,q}$  as  $n \in \mathbb{N}$  increases. This, however, is not an issue for our results. As seen in (4) and (5), the sequences are represented via the empirical averages of probabilistic representations seen in (6) and (7). The expectations of these representations only ever play a role in our proofs regarding the behaviour of the corresponding c.g.f.s, specifically only in the case of  $\|Z\|$  (e.g., in the proofs of Lemma 7 and Lemma 17, or implicitly in the proof of the density approximations in Section 5). As the  $V_i^{(n)}$  in (6) are i.i.d., they all share the same c.g.f. as given in (8) and the same expectation  $\mathbb{E}[V_i^{(n)}] = (M_p(q), M_p(p)) = (m_{p,q}^q, 1)$ . Hence, the fact that the expectation  $m_{n,p,q}$  only converges to  $m_{p,q}$  does not affect our proofs. This is in keeping with classical results from large deviations theory like the

Gärtner–Ellis theorem (see [18, Theorem V.6]), where an arbitrary (i.e. not necessarily i.i.d.) sequence of random variables is not required to have a shared expectation, but rather that the sequence of the (appropriately rescaled) c.g.f.s of the individual random variables in the sequence converge to a fixed function with the origin in the interior of its effective domain. The resulting LDP then considers deviation probabilities from the limiting expectation as well. In the case of  $\|\mathcal{X}\|$  the c.g.f.s of the  $\mathcal{Z}_i^{(n)}$  are not employed at all (neither themselves nor their limit in  $n$ ). Instead, we simply use the density approximation in Proposition 18 for the sum of the  $V_i^{(n)}$  and make use of the independence of  $U^{1/n}$  from the coordinates of  $V_i^{(n)}$ . Since our main results assume  $n \in \mathbb{N}$  to be sufficiently large (that is, large enough for the local density approximations in Section 5 to hold), this effectively means that for  $n \in \mathbb{N}$  sufficiently large, the difference of  $m_{p,q}$  and  $m_{n,p,q}, \mathfrak{m}_{n,p,q}$  is of order at most  $o(1)$  and therefore does not affect our SLD estimates.

### 2.5 A few remarks on Weingarten maps and curvature

As outlined in the introduction, we will finish the proof of our first main result in Theorem 11 by integrating over a previously established density estimate via a result of Adriani and Baldi [1] for Laplacian integral expansions. This result has a heavily geometric flavour and relies on the Weingarten maps of certain hypersurfaces, which in our case are simply curves in  $\mathbb{R}^2$ . We will therefore just give a brief reminder of the Weingarten map in this setting, recall some of its properties, and refer to the relevant literature (e.g., [25, 36]) or Adriani and Baldi [1] for a more in-depth discussion of the topic.

In general, the Weingarten map of a smooth hypersurface  $M \subset \mathbb{R}^d$  at a point  $p \in M$  is an endomorphism of the tangent space  $T_p M$  at  $p$ , mapping any  $y \in T_p M$  to the directional derivative of a normal field of  $M$  in  $p$  in the direction of  $y$ . However, as remarked in [1, Example 4.3], for  $d = 2$ , hypersurfaces simplify to planar curves and the Weingarten map at a point  $p$  simplifies to the absolute value of the curvature  $K(p)$  of the curve at  $p$ . For implicit curves, i.e. curves given as the zero set of a function, we have the following formula for its curvature from [21, Proposition 3.1].

**Lemma 9.** *Let  $F : \mathbb{R}^2 \rightarrow \mathbb{R}$  be a smooth function. For a curve  $\mathcal{C} := \{x \in \mathbb{R}^2 : F(x) = 0\}$  given as the zero set of  $F$ , and a point  $p \in \mathcal{C}$ , where  $\nabla_x F(p) \neq 0$ , it holds that*

$$K(p) = \frac{(-F_{[0,1]}, F_{[1,0]}) \begin{pmatrix} F_{[2,0]} & F_{[1,1]} \\ F_{[1,1]} & F_{[0,2]} \end{pmatrix} (-F_{[0,1]}, F_{[1,0]})}{(F_{[1,0]}^2 + F_{[0,1]}^2)^{3/2}},$$

with derivatives  $F_{[i,j]} = F_{[i,j]}(p)$  as in (1).

**Remark 10.**

- i) Given the set-up of the previous lemma, straightforward calculation of the above fraction gives that

$$K(p) = \frac{F_{[0,1]}^2 F_{[2,0]} - 2F_{[0,1]} F_{[1,0]} F_{[1,1]} + F_{[1,0]}^2 F_{[0,2]}}{(F_{[1,0]}^2 + F_{[0,1]}^2)^{3/2}}.$$

ii) In case that  $\mathcal{C}$  is the graph of a smooth function  $f : \mathbb{R} \rightarrow \mathbb{R}$ , i.e.  $\mathcal{C} = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 = f(x_1)\}$ , and  $p = (x, f(x))$ , the above reduces to

$$K(p) = \frac{|f''(x)|}{(1 + f'(x)^2)^{3/2}}.$$

### 3 Main results

Using the concepts and notation established in the previous section, we now proceed to present our main results and their applications.

#### 3.1 Sharp asymptotics for $q$ -norms of random vectors in $\mathbb{S}_p^{n-1}$ and $\mathbb{B}_p^n$

For  $Z^{(n)} \sim \mathbf{C}_{n,p}$ , we want to give sharp asymptotics for the probability  $\mathbb{P}(n^{1/p-1/q} \|Z^{(n)}\|_q > z)$  for  $z > m_{p,q}$  such that  $z^* \in \mathcal{J}_p$ , with  $z^*$  as defined in Lemma 5. Before presenting our results, let us define the deviation-dependent functions  $\xi(z)$  and  $\kappa(z)$ , as mentioned also in the sharp large deviation results of Bahadur and Ranga Rao [7]. For  $x \in \mathbb{R}^2$ , we set

$$\mathfrak{H}_x := \mathcal{H}_\tau \Lambda_p(\tau(x)) \tag{10}$$

to be the Hessian of the c.g.f.  $\Lambda_p(\tau)$  in  $\tau \in \mathbb{R}^2$ , evaluated at  $\tau(x)$ . For  $z > m_{p,q}$  such that  $z^* \in \mathcal{J}_p$ , we then define the deviation-dependent functions as

$$\xi(z)^2 := \langle \mathfrak{H}_{z^*} \tau(z^*), \tau(z^*) \rangle \det \mathfrak{H}_{z^*}, \tag{11}$$

and

$$\kappa(z)^2 := 1 - c_\kappa(z), \tag{12}$$

where  $c_\kappa(z)$  is given by

$$\frac{(\tau(z^*)_1^2 + \tau(z^*)_2^2)^{3/2} |pq(p-q)z^q|}{|\tau(z^*)_2^2 \left(\mathfrak{H}_{z^*}^{-1}\right)_{11} - 2\tau(z^*)_1 \tau(z^*)_2 \left(\mathfrak{H}_{z^*}^{-1}\right)_{12} + \tau(z^*)_1^2 \left(\mathfrak{H}_{z^*}^{-1}\right)_{22}| (z^{2q} + p^2 q^{-2})^{3/2}}.$$

**Theorem 11.** *Let  $1 \leq q < p < \infty$ ,  $n \in \mathbb{N}$ , and  $Z^{(n)}$  be a random vector in  $\mathbb{B}_p^n$  with  $Z^{(n)} \sim \mathbf{C}_{n,p}$ . Then, for  $n$  sufficiently large and any  $z > m_{p,q}$  such that  $z^* \in \mathcal{J}_p$ , it holds that*

$$\mathbb{P}\left(n^{1/p-1/q} \|Z^{(n)}\|_q > z\right) = \frac{1}{\sqrt{2\pi n} \kappa(z) \xi(z)} e^{-n \Lambda_p^*(z^*)} (1 + o(1)).$$

We want to do the same for  $\mathbb{P}(n^{1/p-1/q} \|\mathcal{Z}^{(n)}\|_q > z)$  with  $\mathcal{Z}^{(n)} \sim \mathbf{U}_{n,p}$  and  $z > m_{p,q}$ . Again, we start by defining our deviation-dependent function for  $z > m_{p,q}$  as

$$\begin{aligned} \gamma(z)^2 &:= \det \mathfrak{H}_{z^*} \tau(z^*)_1^2 (qz^q \tau(z^*)_1 + 1)^2 \\ &\times \left[ \frac{z^{2q} q^2}{p^2} \left(\mathfrak{H}_{z^*}^{-1}\right)_{11} + \frac{2z^q q}{p} \left(\mathfrak{H}_{z^*}^{-1}\right)_{12} + \left(\mathfrak{H}_{z^*}^{-1}\right)_{22} + \tau(z^*)_1 \frac{z^q q (q-p)}{p^2} \right]. \end{aligned} \tag{13}$$

**Theorem 12.** Let  $1 \leq q < p < \infty$ ,  $n \in \mathbb{N}$ , and  $\mathcal{X}^{(n)}$  be a random vector in  $\mathbb{B}_p^n$  with  $\mathcal{X}^{(n)} \sim \mathbf{U}_{n,p}$ . Then, for  $n$  sufficiently large and any  $z > m_{p,q}$  such that  $z^* \in \mathcal{J}_p$ , it holds that

$$\mathbb{P}\left(n^{1/p-1/q} \|\mathcal{X}^{(n)}\|_q > z\right) = \frac{1}{\sqrt{2\pi n} \gamma(z)} e^{-n \Lambda_p^*(z^*)} (1 + o(1)).$$

We have seen in Section 2.4 that  $\|Z\|$  and  $\|\mathcal{X}\|$  both satisfy LDPs with the same speed and rate function for  $z > m_{p,q}$  such that  $z^* \in \mathcal{J}_p$ , despite the underlying distributions being different. Comparing Theorem 11 and Theorem 12 now paints a different picture, with the sharp asymptotics for  $\|Z\|$  and  $\|\mathcal{X}\|$  being noticeably different. As mentioned in our introductory statements, idiosyncratic phenomena of underlying distributions, which can be drowned out on the LDP scale, are often still visible on the scale of sharp large deviations. This is in keeping with what was shown in [37, Theorem 2.4, Theorem 2.6] for one-dimensional projections of  $\ell_p^n$ -spheres and  $\ell_p^n$ -balls.

**Remark 13.** Let us draw a brief comparison between our results and the concentration inequality that follows by the Gromov–Milman theorem as discussed in [38, Remark, p. 1062]. Therein, it is shown that the Gromov–Milman theorem from [22] implies that for  $1 < q \leq p < \infty$  and a random vector  $Z^{(n)} \sim \mathbf{C}_{n,p}$ , it holds that

$$\mathbb{P}\left(\left|n^{1/p-1/q} \|Z^{(n)}\|_q - m_{n,p,q}\right| \geq z\right) \leq C \exp\left(-cn z^{\max\{2,p\}}\right),$$

where  $C > 0$  and  $c > 0$  are constants. If we consider the set-up of Theorem 11, i.e.  $1 \leq q < p < \infty$  and  $z > m_{n,p,q}$ , and only consider deviations without the absolute value, we can derive from the above that

$$\mathbb{P}\left(n^{1/p-1/q} \|Z^{(n)}\|_q > z\right) \leq C \exp\left(-cn z^{\max\{2,p\}}\right).$$

Comparing this with our sharp large deviation results from Theorem 11 for  $z > m_{p,q}$  such that  $z^* \in \mathcal{J}_p$ ,

$$\mathbb{P}\left(n^{1/p-1/q} \|Z^{(n)}\|_q > z\right) = \frac{1}{\sqrt{2\pi n} \kappa(z)\xi(z)} e^{-n \Lambda_p^*(z^*)} (1 + o(1)),$$

we can see that our results improve the estimate in terms of  $n \in \mathbb{N}$  by a factor of  $n^{-1/2}$  and give explicit and deviation-dependent terms  $\kappa(z)$  and  $\xi(z)$  instead of fixed constants for all deviations  $z$ .

**Remark 14.** When comparing the SLD results in Theorem 11 and Theorem 12 to those of Liao and Ramanan [37, Theorem 2.4, Theorem 2.6], one directly notices the core difference in the settings. Liao and Ramanan examine projections of random vectors on  $\mathbb{S}_p^{n-1}$  and  $\mathbb{B}_p^n$  with respective distributions  $\mathbf{C}_{n,p}$  and  $\mathbf{U}_{n,p}$  onto fixed one-dimensional subspaces, and therefore have to consider weighted sums of *dependent* random vectors as probabilistic representations. Thus, all their results have to be conditioned on the projection space and include additional terms accounting for the specifics of the subspace. In our case however, the probabilistic representations

are given as sums of *i.i.d.* random variables (see Section 4), which does not necessitate these additional factors. Therefore, when using results from Liao and Ramanan [37], we adapt their usage accordingly to the given probabilistic representations in our setting. Beyond that however, the SLD results share several similarities, especially when comparing the deviation-dependent terms  $\kappa$ ,  $\xi$  and  $\gamma$ , which for  $q = 1$  are almost equal.

Both proofs of Theorem 11 and Theorem 12 contain three essential steps, as already briefly mentioned in the introduction. The first will be rewriting the probabilities in both theorems with respect to convenient probabilistic representations, specifically  $S^{(n)}$  and  $\mathcal{S}^{(n)}$  given in (16) of Section 4 as the respective empirical averages of the  $V_i^{(n)}$  and  $\mathcal{V}_i^{(n)}$  in (6) and (7). The idea is to write the deviation probabilities as an integral of their distribution over a given “deviation area”. The second step is giving local density approximations for these representations. Since the entries of both the  $V^{(n)}$  and the  $\mathcal{V}^{(n)}$  are highly dependent, no canonical joint densities are available to us to easily do so. However, their Fourier transforms can be given explicitly, thus, for  $n \in \mathbb{N}$  large enough one can use the Fourier inversion theorem to write the densities of  $S^{(n)}$  and  $\mathcal{S}^{(n)}$  as integrals of their Fourier transforms. Heuristically speaking, this means that while the individual  $V_i^{(n)}$  and  $\mathcal{V}_i^{(n)}$  do not possess densities in  $\mathbb{R}^2$ , but for  $n \in \mathbb{N}$  sufficiently large their empirical averages  $S^{(n)}$  and  $\mathcal{S}^{(n)}$  asymptotically do. The resulting integrals can then be approximated using the saddle point method. Since our representations are given as sums of *i.i.d.* random vectors, for whom this has been done in previous results (see, e.g., [10, 14, 42, 43]), we will not prove the density approximations here explicitly. The third and final step then is to calculate the integrals of these densities over their respective deviation area. For  $\|Z\|$ , this is done by a result of Adriani and Baldi [1], which construes the boundary of the deviation area and the level sets of the rate function in the corresponding LDP as hypersurfaces, which are just planar curves in our setting, and uses their Weingarten maps to approximate the integral. For  $\|\mathcal{Z}\|$ , this is not applicable, as certain differentiability conditions are no longer met. Thus, a result by Breitung and Hohenbichler [12] is used, which allows for multi-dimensional Laplace integral approximations under less restrictive differentiability conditions.

### 3.2 Intersection volumes of $\ell_p^n$ -balls

We want to use our sharp large deviation results to further the findings of Schechtman and Schmuckenschläger [45] and Schmuckenschläger [48] for intersection volumes of  $t$ -multiples of different  $\ell_p^n$ -balls. We will first give a brief overview of the original results. For  $p \in [1, \infty)$ , we define  $\mathbb{D}_p^n := \text{vol}_n(\mathbb{B}_p^n)^{-1/n} \mathbb{B}_p^n$  to be the volume normalized  $\ell_p^n$ -ball and recall that

$$\text{vol}_n(\mathbb{B}_p^n) = \frac{\left(2\Gamma\left(1 + \frac{1}{p}\right)\right)^n}{\Gamma\left(1 + \frac{n}{p}\right)}.$$

We furthermore set

$$c_{n,p} := n^{1/p} \text{vol}_n\left(\mathbb{B}_p^n\right)^{1/n} \quad \text{and} \quad c_p := 2 e^{1/p} p^{1/p} \Gamma\left(1 + \frac{1}{p}\right),$$

and recall that it was shown in [45] that  $\lim_{n \rightarrow \infty} c_{n,p} = c_p$ . Moreover, for  $p, q \in [1, \infty)$ ,  $p \neq q$ , we set

$$c_{n,p,q} := \frac{c_{n,p}}{c_{n,q}}, \quad A_{n,p,q} := \frac{c_{n,p}}{m_{p,q} c_{n,q}}, \quad \text{and} \quad A_{p,q} := \lim_{n \rightarrow \infty} A_{n,p,q}.$$

Hence, it follows that

$$A_{p,q} = \frac{c_p}{m_{p,q} c_q} = \frac{\Gamma\left(1 + \frac{1}{p}\right)^{1+(1/q)}}{\Gamma\left(1 + \frac{1}{q}\right) \Gamma\left(\frac{q+1}{p}\right)^{1/q}} e^{1/p-1/q}.$$

Lastly, for  $t \geq 0$  and  $n \in \mathbb{N}$ , we define  $t_n \geq 0$  such that

$$t_n \frac{A_{p,q}}{A_{n,p,q}} = t.$$

Having established the necessary notation, we shall now recall the result of Schmuckenschläger [48, Theorem 3.3]. Therein, it was shown that for  $p, q \in [1, \infty)$ ,  $p \neq q$ , and  $t \geq 0$  it holds that

$$\text{vol}_n \left( \mathbb{D}_p^n \cap t \mathbb{D}_q^n \right) \xrightarrow{n \rightarrow \infty} \begin{cases} 1 & : A_{p,q} t > 1 \\ \frac{1}{2} & : A_{p,q} t = 1 \\ 0 & : A_{p,q} t < 1. \end{cases} \tag{14}$$

To prove this, a central limit theorem for  $n^{1/p-1/q} \|\mathcal{Z}^{(n)}\|_q$  with  $\mathcal{Z}^{(n)} \sim \mathbf{U}_{n,p}$  and  $p, q \in [1, \infty)$ ,  $p \neq q$ , is shown in [48, Proposition 2.4, Proof of Theorem 3.2], since  $\text{vol}_n(\mathbb{D}_p^n \cap t \mathbb{D}_q^n)$  can be written as

$$\begin{aligned} \text{vol}_n(\mathbb{D}_p^n \cap t \mathbb{D}_q^n) &= \text{vol}_n \left( \left\{ z \in \mathbb{D}_p^n : z \in t_n \frac{A_{p,q}}{A_{n,p,q}} \mathbb{D}_q^n \right\} \right) \\ &= \text{vol}_n \left( \left\{ z \in \mathbb{D}_p^n : z \in t_n A_{p,q} m_{p,q} \frac{c_{n,q}}{c_{n,p}} \mathbb{D}_q^n \right\} \right) \\ &= \text{vol}_n \left( \left\{ z \in \text{vol}_n(\mathbb{B}_p^n)^{-1/n} \mathbb{B}_p^n : \right. \right. \\ &\quad \left. \left. z \in t_n A_{p,q} m_{p,q} n^{1/q-1/p} \text{vol}_n(\mathbb{B}_p^n)^{-1/n} \mathbb{B}_q^n \right\} \right) \\ &= \text{vol}_n(\mathbb{B}_p^n)^{-1} \\ &\quad \times \text{vol}_n \left( \left\{ z \in \mathbb{B}_p^n : z \in t_n A_{p,q} m_{p,q} n^{1/q-1/p} \mathbb{B}_q^n \right\} \right) \\ &= \mathbb{P} \left( n^{1/p-1/q} \|\mathcal{Z}^{(n)}\|_q \leq t_n A_{p,q} m_{p,q} \right). \end{aligned} \tag{15}$$

However, we know from the Berry–Esseen theorem (see [49, Theorem 2.1.3]) that the error of the Gaussian approximation given by a central limit theorem decreases with rate  $n^{-1/2}$ . Thus, using (15) and the central limit theorem from [48], we can only infer a rate of convergence of  $n^{-1/2}$  in (14). Using Theorem 12, we can considerably refine that rate of convergence in the first of the three cases in (14) from a sublinear rate to an exponential rate for  $1 \leq q < p < \infty$ .

**Proposition 15.** *Let  $1 \leq q < p < \infty$  and  $n \in \mathbb{N}$ . Using the notation established above, it holds for  $t > m_{p,q} c_{n,p,q}^{-1}$  such that  $(t c_{n,p,q})^* \in \mathcal{J}_p$ , and sufficiently large  $n \in \mathbb{N}$  that*

$$\text{vol}_n \left( \mathbb{D}_p^n \cap t \mathbb{D}_q^n \right) = 1 - \frac{1}{\sqrt{2\pi n} \gamma(t c_{n,p,q})} e^{-n \Lambda_p^*((t c_{n,p,q})^*)} (1 + o(1)).$$

**Proof.** Let  $1 \leq q < p < \infty$ ,  $t > m_{p,q} c_{n,p,q}^{-1}$  such that  $(t c_{n,p,q})^* \in \mathcal{J}_p$ , and assume  $\mathcal{Z}^{(n)}$  is a random vector in  $\mathbb{B}_p^n$  with  $\mathcal{Z}^{(n)} \sim \mathbf{U}_{n,p}$ . Using (15), we get that

$$\begin{aligned} \text{vol}_n \left( \mathbb{D}_p^n \cap t \mathbb{D}_q^n \right) &= \mathbb{P} \left( n^{1/p-1/q} \|\mathcal{Z}^{(n)}\|_q \leq t_n A_{p,q} m_{p,q} \right) \\ &= 1 - \mathbb{P} \left( n^{1/p-1/q} \|\mathcal{Z}^{(n)}\|_q > t_n A_{p,q} m_{p,q} \right). \end{aligned}$$

Now, by  $t > m_{p,q} c_{n,p,q}^{-1}$ , we have that  $t m_{p,q}^{-1} c_{n,p,q} = t A_{n,p,q} = t_n A_{p,q} > 1$ , and hence  $t c_{n,p,q} = t_n A_{p,q} m_{p,q} > m_{p,q}$  with  $(t c_{n,p,q})^* \in \mathcal{J}_p$ . Thus, by Theorem 12, it follows that

$$\text{vol}_n \left( \mathbb{D}_p^n \cap t \mathbb{D}_q^n \right) = 1 - \frac{1}{\sqrt{2\pi n} \gamma(t c_{n,p,q})} e^{-n \Lambda_p^*((t c_{n,p,q})^*)} (1 + o(1)),$$

which finishes our proof. □

### 3.3 One-dimensional projections of $\ell_q^n$ -balls

In Remark 14 we have already discussed the differences between the setting of the results of Liao and Ramanan [37] and the setting of this paper. However, a geometrically similar result to those in [37] follows from Theorem 11. In [28, Section 2.4] Kabluchko, Prochno and Thäle derived a central limit theorem for the length of the projection of an  $\ell_p^n$ -ball onto the line spanned by a random vector  $\theta^{(n)} \in \mathbb{S}^{n-1}$  with  $\theta^{(n)} \sim \mathbf{C}_{n,2}$  as a corollary of their main results. We will proceed similarly and derive sharp large deviation results in the same setting. To be specific, in [37] sharp asymptotics were provided for the *scalar product* of a random vector  $\mathbf{Z}^{(n)} \sim \mathbf{C}_{n,p}$  on  $\mathbb{S}_p^{n-1}$  with a random vector  $\theta^{(n)} \sim \mathbf{C}_{n,2}$  on  $\mathbb{S}^{n-1}$ , which can be negative. We, on the other hand, consider the *absolute value* of the scalar product of such random vectors, thereby only considering nonnegative values.

In the following, for  $q \in [1, \infty]$ , define its conjugate  $q^*$  via  $1/q + 1/q^* = 1$ , setting  $1/\infty = 0$  by convention. Furthermore, for a vector  $\theta^{(n)} \in \mathbb{S}^{n-1}$ , we write  $P_{\theta^{(n)}} \mathbb{B}_q^n$  for the projection of  $\mathbb{B}_q^n$  onto the line spanned by  $\theta^{(n)}$ . Then, our quantity of interest is the projection length  $\text{vol}_1 \left( P_{\theta^{(n)}} \mathbb{B}_q^n \right)$ .

**Corollary 16.** *Let  $2 < q \leq \infty$  and  $\theta^{(n)} \in \mathbb{S}^{n-1}$  be a random vector with  $\theta^{(n)} \sim \mathbf{C}_{n,2}$ . Then, for any  $z > 2 m_{2,q^*}$  such that  $(\frac{z}{2})^* \in \mathcal{J}_p$ , and sufficiently large  $n \in \mathbb{N}$ , it holds that*

$$\mathbb{P} \left( n^{1/2-1/q} \text{vol}_1 \left( P_{\theta^{(n)}} \mathbb{B}_q^n \right) > z \right) = \frac{1}{\sqrt{2\pi n} \kappa \left( \frac{z}{2} \right) \xi \left( \frac{z}{2} \right)} e^{-n \Lambda_2^*((\frac{z}{2})^*)} (1 + o(1)),$$

with  $\Lambda_2$  as in (8) and  $\xi, \kappa$  as in (11), (12), respectively, for  $q^*$  and  $p = 2$ .



**Proof.** It holds that

$$\begin{aligned} \mathbb{P}\left(n^{1/2-1/q} \text{vol}_1\left(P_{\theta^{(n)}}\mathbb{B}_q^n\right) > z\right) &= \mathbb{P}\left(n^{1/2-1/q} 2 \sup_{x \in \mathbb{B}_q^n} |\langle x, \theta^{(n)} \rangle| > z\right) \\ &= \mathbb{P}\left(n^{1/2-1/q} \|\theta^{(n)}\|_{q^*} > \frac{z}{2}\right). \end{aligned}$$

Since  $2 < q \leq \infty$ , we have  $1 \leq q^* < 2 = p$ , whereby we can apply Theorem 11 to the above to get that

$$\mathbb{P}\left(n^{1/2-1/q} \text{vol}_1\left(P_{\theta^{(n)}}\mathbb{B}_q^n\right) > z\right) = \frac{1}{\sqrt{2\pi n} \kappa\left(\frac{z}{2}\right) \xi\left(\frac{z}{2}\right)} e^{-n \Lambda_2^*\left(\left(\frac{z}{2}\right)^*\right)} (1 + o(1)),$$

with  $\Lambda_2, \xi, \kappa$  as described above, which concludes our proof. □

### 4 Probabilistic representation

Recalling the definitions of the random vectors  $V^{(n)}$  and  $\mathcal{V}^{(n)}$  from (6) and (7), we define

$$S^{(n)} := \frac{1}{n} \sum_{i=1}^n V_i^{(n)} \quad \text{and} \quad \mathcal{S}^{(n)} := \frac{1}{n} \sum_{i=1}^n \mathcal{V}_i^{(n)} \tag{16}$$

as the empirical averages of their respective coordinates. Furthermore, we define the sets

$$D_z := \{(t_1, t_2) \in \mathbb{R}^2 : t_1, t_2 > 0, t_1^{1/q} t_2^{-1/p} > z\}$$

and

$$\mathcal{D}_z := \{(t_1, t_2, t_3) \in \mathbb{R}^3 : t_1, t_2 > 0, t_3 \in (0, 1], t_3 t_1^{1/q} t_2^{-1/p} > z\}.$$

It then follows from the reformulations of  $\|Z^{(n)}\|_q$  and  $\|\mathcal{Z}^{(n)}\|_q$  in (4) and (5) that we can write the probabilities within Theorem 11 and Theorem 12 with respect to  $S^{(n)}$  and  $\mathcal{S}^{(n)}$ , respectively, as

$$\begin{aligned} \mathbb{P}\left(n^{1/p-1/q} \|Z^{(n)}\|_q > z\right) &= \mathbb{P}\left(\frac{1}{n} \sum_{i=1}^n |Y_i^{(n)}|^q > z^q \left(\frac{1}{n} \sum_{i=1}^n |Y_i^{(n)}|^p\right)^{\frac{q}{p}}\right) \\ &= \mathbb{P}\left(S^{(n)} \in D_z\right) \end{aligned} \tag{17}$$

and

$$\begin{aligned} \mathbb{P}\left(n^{1/p-1/q} \|\mathcal{Z}^{(n)}\|_q > z\right) &= \mathbb{P}\left(U \frac{1}{n} \sum_{i=1}^n |Y_i^{(n)}|^q > z^q \left(\frac{1}{n} \sum_{i=1}^n |Y_i^{(n)}|^p\right)^{\frac{q}{p}}\right) \\ &= \mathbb{P}\left(\mathcal{S}^{(n)} \in \mathcal{D}_z\right). \end{aligned} \tag{18}$$

We refer to these sets as “deviation areas”, as  $S^{(n)}$  or  $\mathcal{S}^{(n)}$  lying in  $D_z$  or  $\mathcal{D}_z$  represents a deviation of  $\|Z^{(n)}\|_q$  and  $\|\mathcal{Z}^{(n)}\|_q$ . Note that the boundaries of the deviation areas

$$\partial D_z = \{(t_1, t_2) \in \mathbb{R}^2 : t_1, t_2 > 0, t_1^{1/q} t_2^{-1/p} = z\}$$

and

$$\partial \mathcal{D}_z = \{(t_1, t_2, t_3) \in \mathbb{R}^3 : t_1, t_2 > 0, t_3 \in (0, 1], t_3 t_1^{1/q} t_2^{-1/p} = z\}$$

are the same sets given by the infimum conditions in the respective LDPs for  $\|Z\|$  and  $\|\mathcal{Z}\|$  in Proposition 4 and Proposition 6. The fact that for  $z > m_{p,q}$  the rate functions of these LDPs both assume a unique minimum on  $\partial D_z$  and  $\partial \mathcal{D}_z$ , respectively, as was shown in Lemma 5 and Lemma 7, will be essential to the proof of our main results in Sections 6 and 7. We can expand this unique infimum property onto the entirety of  $\overline{D_z}$  and  $\overline{\mathcal{D}_z}$ , as the following lemma will show.

**Lemma 17.** *Assume the same set-up as in Lemma 5 and Lemma 7. Let  $z > m_{p,q}$  such that  $z^* \in \mathcal{J}_p$ . Then*

- i)  $z^* = (z^q, 1)$  is the unique point at which  $\Lambda_p^*$  attains its infimum on  $\overline{D_z}$ ,
- ii)  $z^{**} = (z^q, 1, 1)$  is the unique point at which  $\mathcal{I}_{\mathcal{S}}$  attains its infimum on  $\overline{\mathcal{D}_z}$ .

**Proof.** We start off by showing i). Let  $t \in \mathbb{R}^2$  such that  $t \in D_z^{\circ}$ , meaning  $t_1^{1/q} t_2^{-1/p} > z$ . Then, for  $\tilde{z} := t_1^{1/q} t_2^{-1/p}$  we assume that  $\tilde{z} \in \mathcal{J}_p$ , as otherwise our claim trivially holds by Remark 3. We then have that  $t \in \partial D_{\tilde{z}}$ , thus, by Lemma 5,  $\Lambda_p^*(t_1, t_2) > \Lambda_p^*(\tilde{z}^q, 1) = \mathcal{I}_{\|Z\|}(\tilde{z})$ . We know that  $\Lambda_p^*$  is a convex function with a root in the expectation  $(m_{p,q}^q, 1)$  of the  $V_i^{(n)}$  from (6), since it is a rate function (apply arguments from [17, Lemma 2.2.5] in  $\mathbb{R}^2$ ). We also show in Lemma 21 ii) that  $\mathcal{H}_x \Lambda_p^*(x) = \mathfrak{H}_x^{-1}$ , and have argued in Section 2 for why  $\mathfrak{H}_x$  is positive definite on  $\mathcal{D}_p$ , hence the Hessian of  $\Lambda_p^*$  is also positive definite on  $\mathcal{J}_p$ , giving us the strict convexity of  $\Lambda_p^*$  and, thereby, strict convexity of  $\mathcal{I}_{\|Z\|}(z) = \Lambda_p^*(z^*)$  on  $\mathcal{J}_p$ . Hence we know that  $\mathcal{I}_{\|Z\|}(z)$  is strictly increasing in  $z$  for  $z > m_{p,q}$ . Thus, as  $\tilde{z} > z > m_{p,q}$ , it follows that

$$\Lambda_p^*(t_1, t_2) > \Lambda_p^*(\tilde{z}^q, 1) = \mathcal{I}_{\|Z\|}(\tilde{z}) > \mathcal{I}_{\|Z\|}(z) = \Lambda_p^*(z^q, 1) = \Lambda_p^*(z^*),$$

showing that  $z^* = (z^q, 1)$  minimizes  $\Lambda_p^*$  over  $\overline{D_z}$ . The proof of ii) is analogous, also using the strict monotonicity of the rate function. □

Suppose that the distributions of  $S^{(n)}$  and  $\mathcal{S}^{(n)}$  have respective densities  $h^{(n)}$  and  $\mathfrak{h}^{(n)}$ . Then we can formulate our probabilities of interest as

$$\mathbb{P}\left(n^{1/p-1/q} \|Z^{(n)}\|_q > z\right) = \mathbb{P}\left(S^{(n)} \in D_z\right) = \int_{D_z} h^{(n)}(x) dx \tag{19}$$

and

$$\mathbb{P}\left(n^{1/p-1/q} \|\mathcal{Z}^{(n)}\|_q > z\right) = \mathbb{P}\left(\mathcal{S}^{(n)} \in \mathcal{D}_z\right) = \int_{\mathcal{D}_z} \mathfrak{h}^{(n)}(x) dx. \tag{20}$$

The following section will be devoted to showing the existence of these densities  $h^{(n)}$  and  $\mathfrak{h}^{(n)}$  and presenting them explicitly, while Sections 6 and 7 will then approximate their integrals over their respective deviation areas  $D_z$  and  $\mathcal{D}_z$ .

### 5 Joint density estimate

Recalling the notation and definitions established in Section 2, we assume the same set-up as in Section 4 and can formulate the following local limit theorems for the densities  $h^{(n)}$  and  $\mathfrak{h}^{(n)}$  of our probabilistic representations  $S^{(n)}$  and  $\mathcal{S}^{(n)}$ .

**Proposition 18.** *For  $S^{(n)} = \frac{1}{n} \sum_{i=1}^n V_i^{(n)}$  with  $V_i^{(n)} = (|Y_i^{(n)}|^q, |Y_i^{(n)}|^p)$ ,  $Y_i^{(n)} \sim \mathbf{N}_p$  i.i.d., and  $x \in \mathcal{J}_p$ , it holds that for sufficiently large  $n \in \mathbb{N}$  the distribution of  $S^{(n)}$  has the Lebesgue density*

$$h^{(n)}(x) = \frac{n}{2\pi} (\det \mathfrak{H}_x)^{-1/2} e^{-n \Lambda_p^*(x)} (1 + o(1)),$$

where  $\mathfrak{H}_x := \mathcal{H}_\tau \Lambda_p(\tau(x))$  as in (10).

For the proof of this, we refer to the results of Borovkov and Rogozin [10] or their convenient reformulation in [1, Theorem 3.1]. Therein, a local density estimate is derived for a sum of i.i.d. random vectors in  $\mathbb{R}^d$  via the saddle point method. As discussed in Section 2.2, this means that one writes the density via the Fourier inversion theorem as a complex integral of its Fourier transform and then uses Cauchy’s theorem to deform the path of integration, so that it passes through a complex saddle point. For sufficiently large  $n \in \mathbb{N}$ , the mass of the integral then heavily concentrates around that saddle point and standard integral expansion methods can be used to great effect. Naturally, this requires the conditions of the Fourier inversion theorem to be met, that is, the Fourier transform of the density has to be integrable. In [1, Theorem 3.1] this follows from the assumption that all the i.i.d. random vectors have a common bounded density, though it is noted in [1, Remark 3.2], that this can be replaced by any argument ensuring that the Fourier inversion theorem can be applied. In our setting, the i.i.d. vectors are given by  $V_i^{(n)} := (|Y_i^{(n)}|^q, |Y_i^{(n)}|^p)$ , whose coordinates are highly dependent, thus such a density of the  $V_i^{(n)}$  is not available. However, one can write the Fourier transform of  $V_i^{(n)}$  with respect to the underlying distribution  $\mathbf{N}_p$  of the  $Y_i^{(n)}$ , and then infer integrability via the properties of its density  $f_p$  and the Hausdorff–Young inequality, as was done by Liao and Ramanan in [37, Lemma 6.1]. As the considered settings are quite similar, virtually the same arguments can be applied in our case, thereby making sure our referral to [1, Theorem 3.1] is indeed justified.

**Proposition 19.** *For  $\mathcal{S}^{(n)} = \frac{1}{n} \sum_{i=1}^n \mathcal{Z}_i^{(n)}$  with  $\mathcal{Z}_i^{(n)} = (|Y_i^{(n)}|^q, |Y_i^{(n)}|^p, U^{1/n})$ ,  $Y_i^{(n)} \sim \mathbf{N}_p$  i.i.d.,  $U$  uniformly distributed on  $[0, 1]$  independently of the  $Y_i^{(n)}$ , and  $x = (x_1, x_2) \in \mathcal{J}_p$ ,  $y \in (0, 1]$ , it holds that for sufficiently large  $n \in \mathbb{N}$  the distribution of  $\mathcal{S}^{(n)}$  has the Lebesgue density*

$$\mathfrak{h}^{(n)}(x_1, x_2, y) = \frac{n^2}{2\pi} y^{-1} (\det \mathfrak{H}_x)^{-1/2} e^{-n \mathcal{I}_\mathcal{S}(x_1, x_2, y)} (1 + o(1)),$$

where  $\mathcal{I}_\mathcal{S}(x_1, x_2, y) := [\Lambda_p^*(x) - \log(y)]$  and  $\mathfrak{H}_x := \mathcal{H}_\tau \Lambda_p(\tau(x))$  as in (10).

**Proof.** By direct calculation, we can see for  $y \in [0, 1]$  that  $\mathbb{P}(U^{1/n} \leq y) = \mathbb{P}(U \leq y^n) = y^n$ , hence the density of  $U^{1/n}$  is given by  $f_{U^{1/n}}(y) = n y^{n-1}$ . As

$U^{1/n}$  is independent of the  $Y_i^{(n)}$ , and thereby also of  $S^{(n)} = (|Y_i^{(n)}|^q, |Y_i^{(n)}|^p)$ , the density of  $\mathcal{S}^{(n)} = \frac{1}{n} \sum_{i=1}^n (|Y_i^{(n)}|^q, |Y_i^{(n)}|^p, U^{1/n})$  is given by the product of their densities, hence

$$\begin{aligned} \mathcal{K}^{(n)}(x_1, x_2, y) &= h^{(n)}(x_1, x_2) f_{U^{1/n}}(y) \\ &= \frac{n^2}{2\pi} y^{-1} (\det \mathfrak{H}_x)^{-1/2} e^{-n [\Lambda_p^*(x) - \log(y)]} (1 + o(1)). \end{aligned}$$

This completes our proof. □

### 6 Proof of the main result for $\ell_p^n$ -spheres

In (19) we have reformulated the deviation probability  $\mathbb{P}(n^{1/p-1/q} \|Z^{(n)}\|_q > z)$  as an integral of the density estimate  $h^{(n)}$  of the probabilistic representation  $S^{(n)}$  over the deviation area  $D_z$ . In Proposition 18 we gave  $h^{(n)}$  explicitly. For the proof of Theorem 11 it remains to calculate that integral. To do so, the integral will be split up into a neighbourhood  $B_z$  of the point  $z^*$ , that has been shown in Lemma 17 to be the infimum of  $\Lambda_p^*$  over  $\bar{D}_z$ , and its complement  $B_z^c$ . The LDP from Proposition 4 will be used to show the negligibility of the integral outside of the neighbourhood of  $z^*$ . Within the neighbourhood  $B_z$ , we use a result from Adriani and Baldi [1], which uses the Weingarten maps of the planar curves given by the boundary of  $D_z \cap B_z$  and the level set of  $\Lambda_p^*$  at  $z^*$ , to compute the integral. Following that, we will give these Weingarten maps explicitly, finishing our proof.

**Proof of Theorem 11.** We assume the set-up of Theorem 11 and use the reformulation (19) to proceed by considering  $\mathbb{P}(S^{(n)} \in D_z)$ . Let  $B_z \subset \mathbb{R}^2$  be an open neighbourhood around  $z^*$ , small enough that  $B_z \subset \mathcal{J}_p$ . Then

$$\begin{aligned} \mathbb{P}(S^{(n)} \in D_z) &= \int_{D_z} h^{(n)}(x) dx \\ &= \int_{D_z \cap B_z} h^{(n)}(x) dx + \int_{D_z \cap B_z^c} h^{(n)}(x) dx. \end{aligned} \tag{21}$$

Since  $z^* \notin B_z^c$ , by Lemma 17, there exists an  $\eta > 0$ , such that

$$\inf_{y \in D_z \cap B_z^c} \Lambda_p^*(y) > \Lambda_p^*(z^*) + \eta,$$

and thus, by the LDP in Proposition 4, it holds that

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(S^{(n)} \in D_z \cap B_z^c) \leq - \inf_{y \in D_z \cap B_z^c} \Lambda_p^*(y) \leq -\Lambda_p^*(z^*) - \eta.$$

This gives us that

$$\begin{aligned} \mathbb{P}(S^{(n)} \in D_z \cap B_z^c) &\leq e^{-n \Lambda_p^*(z^*) - n \eta} (1 + o(1)) \\ &= \frac{1}{e^{n \eta}} e^{-n \Lambda_p^*(z^*)} (1 + o(1)). \end{aligned} \tag{22}$$

Furthermore, by our density estimate in Proposition 18, it holds that

$$\int_{D_z \cap B_z} h^{(n)}(x) \, dx = \frac{n}{2\pi} \int_{D_z \cap B_z} (\det \mathfrak{H}_x)^{-1/2} e^{-n \Lambda_p^*(x)} \, dx (1 + o(1)). \tag{23}$$

To calculate this explicitly, we will rely on a technique established in [1, Proof of Theorem 4.4]. Therein, an asymptotic integral expansion of Bleistein and Handelsmann [9, Equation (8.3.63)] for Laplace integrals is reformulated via the Weingarten maps of the integration area and the level set of the exponential function at its minimum, both seen as hypersurfaces. We will present it as one concise result, similar to that formulated in [37, Lemma 4.6].

**Proposition 20.** *Let  $D \subset \mathbb{R}^d$  be a bounded domain such that  $\partial D$  is a differentiable hypersurface in  $\mathbb{R}^d$ . Furthermore, let  $g : \mathbb{R}^d \rightarrow \mathbb{R}$  be a differentiable function and  $\phi : D \rightarrow [0, \infty)$  be a nonnegative function that is twice differentiable and attains a unique infimum over  $\overline{D}$  at  $x^* \in \partial D$ . Define the hypersurfaces*

$$\mathcal{E}_D = \partial D \quad \text{and} \quad \mathcal{E}_\phi = \{x \in \mathbb{R}^d : \phi(x) = \phi(x^*)\},$$

and denote by  $L_D$  and  $L_\phi$  their respective Weingarten maps at  $x^*$ . Then, for sufficiently large  $n \in \mathbb{N}$ , it holds that

$$\begin{aligned} & \int_D g(x) e^{-n \phi(x)} \, dx \\ &= \frac{(2\pi)^{(d-1)/2} \det(L_\phi^{-1}(L_\phi - L_D))^{-1/2}}{n^{(d+1)/2} (\mathcal{H}_x \phi(x^*)^{-1} \nabla_x \phi(x^*), \nabla_x \phi(x^*))^{1/2}} g(x^*) e^{-n \phi(x^*)} (1 + o(1)). \end{aligned}$$

The proof of this is given by first applying the result from [9, Equation (8.3.63)] for Laplace-type integrals and then using the reformulation of the terms therein from [1, Equation (4.6)] with respect to the Weingarten map.

Let us now check that the above conditions hold for the integral in (23). We have that  $D_z \cap B_z$  is bounded, and for  $z > m_{p,q}$ , we can write  $\partial D_z$  as the graph of the infinitely differentiable function  $f : (0, \infty) \rightarrow (0, \infty)$  with  $f(t_1) = z^{-p} t_1^{p/q}$ , thus both  $\partial D_z$  and  $\partial(D_z \cap B_z)$  are differentiable planar curves. As discussed in Section 2, for  $x \in \mathcal{J}_p$ ,  $g_x(\tau) := \langle x, \tau \rangle - \Lambda_p(\tau)$  has a unique supremum  $\tau(x)$ , i.e.  $x - \nabla_\tau \Lambda_p(\tau) = 0$  has a unique solution in  $(x, \tau)$ . It was also discussed that  $\mathcal{H}_\tau \Lambda_p(\tau)$  is invertible for all  $\tau \in \mathcal{D}_p$ , thus, it follows from the implicit function theorem that  $x \mapsto \tau(x)$  is as differentiable in  $x$  as  $(x, \tau) \mapsto (x - \nabla_\tau \Lambda_p(\tau))$  is in  $\tau$ . As  $\Lambda_p$  is the logarithm of the joint m.g.f. of the  $V_i^{(n)}$  from (6), the components of its derivatives are themselves infinitely differentiable within  $\mathcal{D}_p$  by the standard properties of the m.g.f. (see, e.g., [15, Theorem 5.4]), yielding that  $\tau(x)$  is infinitely differentiable on  $\mathcal{J}_p$ . This overall gives us the infinite differentiability of the function  $\Lambda_p(\tau(x))$  and thereby also of  $\Lambda_p^*(x)$  on  $\mathcal{J}_p$ . For  $B_z$  chosen small enough, it follows that for any  $z > m_{p,q}$  with  $z^* \in \mathcal{J}_p$ ,  $\Lambda_p^*$  is twice differentiable on  $D_z \cap B_z$ . Nonnegativity of  $\Lambda_p^*$  follows directly by the standard properties of rate functions (apply arguments from, e.g., [17, Lemma 2.2.5] in  $\mathbb{R}^2$ ). By the infinite differentiability of  $\Lambda_p(\tau(x))$  in  $x$ , we get the differentiability of  $g(x) := (\det \mathfrak{H}_x)^{-1/2} = (\det \mathcal{H}_\tau \Lambda_p(\tau(x)))^{-1/2}$  in  $x$ . Lemma 17 gives us

the uniqueness of  $z^* = (z^q, 1) \in \partial(D_z \cap B_z)$  as an infimum on  $\overline{D_z}$  and  $\overline{D_z \cap B_z}$ . Thus, in view of the above, we can use Proposition 20 for  $D = D_z \cap B_z \subset \mathbb{R}^2$  with  $g(x) = (\det \mathfrak{H}_x)^{-1/2}$ ,  $\phi(x) = \Lambda_p^*(x)$ , and  $x^* = z^*$ , and get that

$$\begin{aligned} & \int_{D_z \cap B_z} h^{(n)}(x) \, dx \\ &= \frac{n}{2\pi} \frac{(2\pi)^{1/2} \det(L_\Lambda^{-1}(L_\Lambda - L_D))^{-1/2} (\det \mathfrak{H}_{z^*})^{-1/2} e^{-n \Lambda_p^*(z^*)}}{n^{3/2} \langle \mathcal{H}_x \Lambda_p^*(z^*)^{-1} \nabla_x \Lambda_p^*(z^*), \nabla_x \Lambda_p^*(z^*) \rangle^{1/2}} (1 + o(1)), \end{aligned} \tag{24}$$

for the respective Weingarten maps at  $z^*$  of the curves

$$\mathcal{E}_D = \partial(D_z \cap B_z) \quad \text{and} \quad \mathcal{E}_\Lambda = \{x \in \mathbb{R}^2 : \Lambda_p^*(x) = \Lambda_p^*(z^*)\}.$$

Let us present the following identities for some of the terms in the fraction above, resulting from the definition of  $\tau(x)$  and the properties of the Legendre–Fenchel transform:

**Lemma 21.** *It holds that*

$$i) \quad \nabla_x \Lambda_p^*(x) = \tau(x),$$

$$ii) \quad \mathcal{H}_x \Lambda_p^*(x) = \mathfrak{H}_x^{-1}.$$

**Proof.** We start by showing that  $\nabla_x \Lambda_p^*(x) = \tau(x)$ . We have defined  $\tau(x)$  as the supremum of  $\langle x, \tau \rangle - \Lambda_p(\tau)$  in  $\tau \in \mathbb{R}^2$  (see (9)), thus it follows that

$$\nabla_\tau \langle x, \tau \rangle - \Lambda_p(\tau) \Big|_{\tau=\tau(x)} = x - \nabla_\tau \Lambda_p(\tau(x)) = 0. \tag{25}$$

With this, it follows that

$$\begin{aligned} \nabla_x \Lambda_p^*(x) &= \nabla_x \langle x, \tau(x) \rangle - \Lambda_p(\tau(x)) \\ &= \tau(x) + J_x \tau(x) x - \nabla_x \Lambda_p(\tau(x)) \\ &= \tau(x) + J_x \tau(x) x - J_x \tau(x) \nabla_\tau \Lambda_p(\tau(x)) \\ &= \tau(x) + J_x \tau(x) [x - \nabla_\tau \Lambda_p(\tau(x))] \\ &= \tau(x). \end{aligned}$$

Let us now prove that  $\mathcal{H}_x \Lambda_p^*(x) = \mathfrak{H}_x^{-1}$ . On the one hand, it follows from the above that

$$\mathcal{H}_x \Lambda_p^*(x) = J_x \tau(x), \tag{26}$$

while on the other hand, it holds that

$$\begin{aligned} \mathcal{H}_x \Lambda_p^*(x) &= \mathcal{H}_x \langle x, \tau(x) \rangle - \Lambda_p(\tau(x)) \\ &= \mathcal{H}_x \langle x, \tau(x) \rangle - \mathcal{H}_x \Lambda_p(\tau(x)) \end{aligned}$$

$$\begin{aligned}
 &= J_x[\nabla_x \langle x, \tau(x) \rangle] - J_x[\nabla_x \Lambda_p(\tau(x))] \\
 &= J_x[\tau(x) + J_x \tau(x)x] - J_x[J_x \tau(x) \nabla_\tau \Lambda_p(\tau(x))] \\
 &= J_x \tau(x) + J_x[J_x \tau(x)x] - \mathcal{H}_x \tau(x) \nabla_\tau \Lambda_p(\tau(x)) \\
 &\quad - J_x \tau(x) J_x[\nabla_\tau \Lambda_p(\tau(x))] \\
 &= 2J_x \tau(x) + \mathcal{H}_x \tau(x)[x - \nabla_\tau \Lambda_p(\tau(x))] \\
 &\quad - J_x \tau(x) J_x \tau(x) \mathcal{H}_\tau \Lambda_p(\tau(x)) \\
 &= 2J_x \tau(x) - J_x \tau(x) J_x \tau(x) \mathcal{H}_\tau \Lambda_p(\tau(x)). \tag{27}
 \end{aligned}$$

Equating the terms (26) and (27) yields

$$\begin{aligned}
 J_x \tau(x) &= 2J_x \tau(x) - J_x \tau(x) J_x \tau(x) \mathcal{H}_\tau \Lambda_p(\tau(x)) \\
 \Leftrightarrow 0 &= J_x \tau(x) - J_x \tau(x) J_x \tau(x) \mathcal{H}_\tau \Lambda_p(\tau(x)) \\
 \Leftrightarrow 0 &= I_2 - J_x \tau(x) \mathcal{H}_\tau \Lambda_p(\tau(x)) \\
 \Leftrightarrow J_x \tau(x) &= \mathcal{H}_\tau \Lambda_p(\tau(x))^{-1},
 \end{aligned}$$

where  $I_2$  denotes the identity matrix in  $\mathbb{R}^2$ . Again using (26) on the above yields

$$\mathcal{H}_x \Lambda_p^*(x) = J_x \tau(x) = \mathcal{H}_\tau \Lambda_p(\tau(x))^{-1} = \mathfrak{H}_x^{-1},$$

and thereby finishes the proof. □

Via Lemma 21, we get

$$\left\langle \mathcal{H}_x \Lambda_p^*(z^*)^{-1} \nabla_x \Lambda_p^*(z^*), \nabla_x \Lambda_p^*(z^*) \right\rangle = \left\langle \mathfrak{H}_{z^*} \tau(z^*), \tau(z^*) \right\rangle.$$

With the definition of  $\xi(z)^2$  in (11) the integral in (24) hence simplifies as follows:

$$\begin{aligned}
 &\int_{D_z \cap B_z} h^{(n)}(x) dx \\
 &= \frac{1}{\sqrt{2\pi n} \xi(z)} \left( \det(L_\Lambda^{-1}(L_\Lambda - L_D)) \right)^{-1/2} e^{-n \Lambda_p^*(z^*)} (1 + o(1)). \tag{28}
 \end{aligned}$$

We see that it only remains to prove that  $\det(L_\Lambda^{-1}(L_\Lambda - L_D)) = \kappa(z)^2$ . We proceed to calculate the Weingarten maps of the curves  $\mathcal{C}_D$  and  $\mathcal{C}_\Lambda$  explicitly. As discussed in Section 2.5, the Weingarten map of a planar curve at a point  $x$  reduces to the absolute value of its curvature in  $x$ . As previously mentioned,  $\partial D_z$  is the graph of a function  $f : (0, \infty) \rightarrow (0, \infty)$  with  $f(t_1) = z^{-p} t_1^{p/q}$ . Thus, the same holds locally for  $\mathcal{C}_D = \partial(D_z \cap B_z)$  in a neighbourhood of  $z^*$ , so by the curvature formula for graphs of functions, as seen in Remark 10 ii), it holds that

$$L_D = \frac{|f''(z^q)|}{(1 + f'(z^q)^2)^{3/2}},$$

where

$$f'(t_1)^2 = \left( pq^{-1} z^{-p} t_1^{(p/q)-1} \right)^2 \Rightarrow f'(z^q)^2 = p^2 q^{-2} z^{-2q},$$

and

$$f''(t_1) = pq^{-1} \left( pq^{-1} - 1 \right) z^{-p} t_1^{(p/q)-2} \Rightarrow f''(z^q) = (p^2 - pq)q^{-2} z^{-2q}.$$

This yields

$$L_D = \frac{|(p^2 - pq)q^{-2} z^{-2q}|}{(1 + p^2 q^{-2} z^{-2q})^{3/2}} = \frac{|pq(p - q)z^q|}{(z^{2q} + p^2 q^{-2})^{3/2}}. \tag{29}$$

The curve  $\mathcal{C}_\Lambda$  is the zero set of the function  $F(x) := \Lambda_p^*(x) - \Lambda_p^*(z^*)$ . From Lemma 21 we know that

$$(F_{[1,0]}, F_{[0,1]}) = \left( \frac{\partial}{\partial x_1} \Lambda_p^*(z^*), \frac{\partial}{\partial x_2} \Lambda_p^*(z^*) \right) = \tau(z^*)$$

and

$$\begin{pmatrix} F_{[2,0]} & F_{[1,1]} \\ F_{[1,1]} & F_{[0,2]} \end{pmatrix} = \begin{pmatrix} \frac{\partial^2}{\partial^2 x_1} \Lambda_p^*(z^*) & \frac{\partial^2}{\partial x_2 \partial x_1} \Lambda_p^*(z^*) \\ \frac{\partial^2}{\partial x_1 \partial x_2} \Lambda_p^*(z^*) & \frac{\partial^2}{\partial^2 x_2} \Lambda_p^*(z^*) \end{pmatrix} = \mathfrak{H}_{z^*}^{-1},$$

for derivatives  $F_{[i,j]} = F_{[i,j]}(z^*)$  as in (1). Hence, by the curvature formula for implicit curves from Lemma 9 and Remark 10 i), we get

$$L_\Lambda = \frac{\left| \tau(z^*)_2^2 \left( \mathfrak{H}_{z^*}^{-1} \right)_{11} - 2\tau(z^*)_1 \tau(z^*)_2 \left( \mathfrak{H}_{z^*}^{-1} \right)_{12} + \tau(z^*)_1^2 \left( \mathfrak{H}_{z^*}^{-1} \right)_{22} \right|}{\left( \tau(z^*)_1^2 + \tau(z^*)_2^2 \right)^{3/2}}. \tag{30}$$

Since both  $L_D$  and  $L_\Lambda$  are one-dimensional, it follows from (29) and (30) that

$$\det(L_\Lambda^{-1}(L_\Lambda - L_D)) = L_\Lambda^{-1}(L_\Lambda - L_D) = 1 - \frac{L_D}{L_\Lambda} = \kappa(z)^2.$$

for  $\kappa(z)^2$  as in (12). It now follows with (28) that

$$\int_{D_z \cap B_z} h^{(n)}(x) dx = \frac{1}{\sqrt{2\pi n} \xi(z) \kappa(z)} e^{-n\Lambda_p^*(z^*)} (1 + o(1)). \tag{31}$$

Comparing (31) with the upper bound of the integral outside of  $B_z$  in (22), we can see that the integral over  $B_z^c$  is negligible for large  $n \in \mathbb{N}$ . Thus, combining (21), (22) and (31) finishes the proof of Theorem 11.  $\square$

### 7 Proof of the main result for $\ell_p^n$ -balls

We use the notation and definitions established in Sections 2 through 4. Let  $1 \leq q < p < \infty$  and  $z > m_{p,q}$  be such that  $z^* \in \mathcal{J}_p$ . We proceed similarly to the previous proof, using the reformulation of  $\mathbb{P}(n^{1/p-1/q} \|\mathcal{X}^{(n)}\|_q > z)$  from (20) in conjunction



with the density approximation from Proposition 19. The resulting integral over  $\mathcal{D}_z$  is again split into a neighbourhood of the minimum of  $\mathcal{I}_S$  over  $\mathcal{D}_z$  and its complement, which, according to Lemma 17, is attained at  $z^{**} = (z^q, 1, 1)$ . For the integral within that neighbourhood, we apply a result of Breitung and Hohenbichler [12], which yields an integral approximation under less restrictive differentiability conditions than those in Proposition 20. This result is again geometric in nature, as the behaviour of the density on  $\partial\mathcal{D}_z$  still heavily dictates the value of the overall approximation. However, since this result is formulated for a certain neighbourhood of the origin, we first need to construct a sufficient transformation, mapping our deviation area into such a neighbourhood. After that, we calculate the specific approximation in our setting.

**Proof of Theorem 12.** We assume the set-up of Theorem 12 and use the reformulation (20) to proceed by considering  $\mathbb{P}(\mathcal{S}^{(n)} \in \mathcal{D}_z)$ . Let  $\mathcal{B}_z \subset \mathbb{R}^3$  be an open neighbourhood around  $z^{**} = (z^q, 1, 1)$  small enough that the first two coordinates of points within  $\mathcal{B}_z$  lie in  $\mathcal{J}_p$  and the third is positive. Then it holds by Proposition 19 that

$$\begin{aligned} \mathbb{P}(\mathcal{S}^{(n)} \in \mathcal{D}_z) &= \int_{\mathcal{D}_z \cap \mathcal{B}_z} \mathfrak{h}^{(n)}(x_1, x_2, y) \, dx_1 dx_2 dy \\ &\quad + \int_{\mathcal{D}_z \cap \mathcal{B}_z^c} \mathfrak{h}^{(n)}(x_1, x_2, y) \, dx_1 dx_2 dy. \end{aligned} \tag{32}$$

As in the proof of Theorem 11, we can deduce from Lemma 17 ii) and the LDP in Proposition 6 that there is an  $\eta > 0$ , such that

$$\begin{aligned} \mathbb{P}(\mathcal{S}^{(n)} \in \mathcal{D}_z \cap \mathcal{B}_z^c) &\leq e^{-n\mathcal{I}_S(z^{**}) - n\eta}(1 + o(1)) \\ &= \frac{1}{e^{n\eta}} e^{-n\Lambda_p^*(z^*)}(1 + o(1)), \end{aligned} \tag{33}$$

with  $\mathcal{I}_S(t) = [\Lambda_p^*(t_1, t_2) - \log(t_3)]$ , as defined in Lemma 7. Let us now consider the first integral in (32). Since  $z^* \in \mathcal{J}_p$ , for sufficiently small  $\mathcal{B}_z$ , we have that  $x = (x_1, x_2) \in \mathcal{J}_p$  and  $y \in (0, 1]$ . By the density approximation from Proposition 19, it holds that

$$\begin{aligned} &\int_{\mathcal{D}_z \cap \mathcal{B}_z} \mathfrak{h}^{(n)}(x_1, x_2, y) \, dx_1 dx_2 dy \\ &= \frac{n^2}{2\pi} \int_{\mathcal{D}_z \cap \mathcal{B}_z} y^{-1} (\det \mathfrak{H}_x)^{-1/2} e^{-n\mathcal{I}_S(x_1, x_2, y)} \, dx_1 dx_2 dy (1 + o(1)). \end{aligned}$$

As we have seen in Lemma 17,  $\mathcal{I}_S$  attains its infimum on  $\overline{\mathcal{D}_z}$  at  $z^{**}$ . However, we cannot use the result of Adriani and Baldi from Proposition 20 here, since at  $z^{**}$  the boundary of  $\mathcal{D}_z \cap \mathcal{B}_z$  is not differentiable, and thereby not smooth. Hence, we use the following asymptotic integral approximation results based on Breitung and Hohenbichler [12], which gives a Laplace integral approximation very similar to that in Liao and Ramanan [37, Lemma 5.1], but under weaker conditions.

**Proposition 22.** *Let  $F \subset \mathbb{R}^3$  be a bounded closed set containing the origin in its interior. If*

- (a)  $f : F \rightarrow \mathbb{R}$  and  $g : F \rightarrow \mathbb{R}$  are continuous functions with  $g(\mathbf{0}) \neq 0$ , where  $\mathbf{0} := (0, 0, 0)$ ,
- (b)  $f(x) > f(\mathbf{0})$  for all  $x \in F \cap (\mathbb{R}_+^2 \times \mathbb{R}) \setminus \{0\}$ ,
- (c) there is a neighbourhood  $V \subset F$  of  $\mathbf{0}$  in which  $f$  is twice continuously differentiable,
- (d)  $f_{[1,0,0]} > 0$ ,  $f_{[0,1,0]} > 0$ , and  $f_{[0,0,2]} > 0$ , with derivatives  $f_{[i,j,k]} = f_{[i,j,k]}(\mathbf{0})$  as in (1),

then

$$\int_{F \cap (\mathbb{R}_+^2 \times \mathbb{R})} g(x) e^{-nf(x)} dx = \frac{\sqrt{2\pi}}{n^{5/2}} \frac{g(x^*)}{f_{[1,0,0]}f_{[0,1,0]}\sqrt{f_{[0,0,2]}}} e^{-nf(x^*)}(1 + o(1)).$$

**Remark 23.** This is the result from [12, Lemma 4] for  $n = 3$ ,  $k = 2$  and functions  $g$  and  $(-f)$  instead of  $h$  and  $f$ . The parameter  $\lambda$  in our setting is replaced by the integer  $n \in \mathbb{N}$ . Furthermore, a typo within the said result has been corrected, namely the sum in [12, Equation (11)] is replaced by a product (compare proof therein). This proposition is quite close to [37, Lemma 5.1], but does not require the same level of smoothness of  $f$  and  $g$ , and  $g$  does not depend on  $n \in \mathbb{N}$ .

To apply this, we use a transformation of  $\mathcal{D}_z \cap \mathcal{B}_z$ , mapping  $z^{**} = (z^q, 1, 1)$  to  $\mathbf{0}$ . Consider  $\mathfrak{J} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  with

$$\mathfrak{J}(x_1, x_2, y) = (y^q x_1 - z^q x_2^{q/p}, 1 - y, x_2 - 1) = (t_1, t_2, t_3).$$

It holds that  $\mathfrak{J}(z^{**}) = \mathbf{0}$  and  $\mathfrak{J}(\mathcal{D}_z) = \tilde{\mathcal{D}}_z := \{t \in \mathbb{R}^3 : t_1 > 0, t_2 \in [0, 1), t_3 > -1\}$ . Furthermore, in a neighbourhood of  $z^{**}$  small enough such that  $t_2 < 1$ ,  $\mathfrak{J}$  is invertible with

$$\mathfrak{J}^{-1}(t_1, t_2, t_3) = \left( \frac{t_1 + z^q(t_3 + 1)^{q/p}}{(1 - t_2)^q}, t_3 + 1, 1 - t_2 \right).$$

Let us calculate the Jacobian of  $\mathfrak{J}^{-1}$ :

$$J_t \mathfrak{J}^{-1}(t) = \begin{pmatrix} \frac{1}{(1-t_2)^q} & \frac{q(t_1+z^q(t_3+1)^{q/p})}{(1-t_2)^{q+1}} & \frac{z^q \frac{q}{p} (t_3+1)^{(q/p)-1}}{(1-t_2)^q} \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}. \tag{34}$$

Thus, we have that  $|\det J_t \mathfrak{J}^{-1}(t)| = (1 - t_2)^{-q}$ . We set

$$\mathcal{G}(x_1, x_2, y) := y^{-1}(\det \mathfrak{H}_x)^{-1/2},$$

as well as  $\tilde{\mathcal{B}}_z := \mathfrak{J}(\mathcal{B}_z)$ , and transform the area of integration via  $\mathfrak{J}^{-1}$ , yielding

$$\begin{aligned} &\mathbb{P} \left( \mathcal{S}^{(n)} \in \mathcal{D}_z \cap \mathcal{B}_z \right) \\ &= \int_{\mathcal{D}_z \cap \mathcal{B}_z} \mathfrak{H}^{(n)}(x_1, x_2, y) dx_1 dx_2 dy \end{aligned}$$

$$\begin{aligned} &= \frac{n^2}{2\pi} \int_{\mathcal{D}_z \cap \mathcal{B}_z} y^{-1} (\det \mathfrak{H}_x)^{-1/2} e^{-n [\Lambda_p^*(x_1, x_2) - \log(y)]} dx_1 dx_2 dy (1 + o(1)) \\ &= \frac{n^2}{2\pi} \int_{\mathcal{D}_z \cap \mathcal{B}_z} \mathcal{G}(x_1, x_2, y) e^{-n \mathcal{I}_S(x_1, x_2, y)} dx_1 dx_2 dy (1 + o(1)) \\ &= \frac{n^2}{2\pi} \int_{\tilde{\mathcal{D}}_z \cap \tilde{\mathcal{B}}_z} \mathcal{G} \circ \mathcal{J}^{-1}(t) e^{-n \mathcal{I}_S \circ \mathcal{J}^{-1}(t)} (1 - t_2)^{-q} dt (1 + o(1)). \end{aligned}$$

We now set  $\tilde{g}(t) := (1 - t_2)^{-q} \mathcal{G} \circ \mathcal{J}^{-1}(t)$  and  $\tilde{f}(t) := \mathcal{I}_S \circ \mathcal{J}^{-1}(t)$ , then

$$\mathbb{P} \left( \mathcal{S}^{(n)} \in \mathcal{D}_z \cap \mathcal{B}_z \right) = \frac{n^2}{2\pi} \int_{\tilde{\mathcal{D}}_z \cap \tilde{\mathcal{B}}_z} \tilde{g}(t) e^{-n \tilde{f}(t)} dt (1 + o(1)). \tag{35}$$

We intend to apply Proposition 22 to the integral in (35) for  $F = \tilde{\mathcal{B}}_z$ . It holds that  $\tilde{\mathcal{D}}_z \cap \tilde{\mathcal{B}}_z$  is bounded and since the value of the integral is the same if we integrate over the open set  $\tilde{\mathcal{B}}_z$  or its closure, we will continue to work with  $\tilde{\mathcal{B}}_z$ . Further, we have that  $\mathcal{B}_z$  contains the origin in its interior, as the interior point  $z^{**}$  of  $\mathcal{B}_z$  is again mapped by the continuous function  $\mathcal{J}$  onto an interior point, which is  $\mathcal{J}(z^{**}) = \mathbf{0}$ . Since we have chosen the neighbourhood  $\mathcal{B}_z$  of  $z^{**}$  small enough for  $\tilde{\mathcal{B}}_z$  to not contain  $(t_1, 1, t_3)$ , it holds that

$$\tilde{g}(t) = (1 - t_2)^{-q} \left[ (1 - t_2)^{-1} (\det \mathfrak{H}_{(\mathcal{J}^{-1}(t)_1, \mathcal{J}^{-1}(t)_3)})^{-1/2} \right]$$

is also differentiable on  $\tilde{\mathcal{D}}_z \cap \tilde{\mathcal{B}}_z$  as a composition of differentiable functions and thereby continuous on  $\tilde{\mathcal{B}}_z$ . The differentiability of  $\mathcal{J}^{-1}$ , together with that of  $\Lambda_p^*$  shown in the proof of Theorem 11, yields the differentiability (and thereby the continuity) of  $\tilde{f}(t) := \mathcal{I}_S \circ \mathcal{J}^{-1}(t)$  on  $\tilde{\mathcal{B}}_z$ . It holds furthermore that

$$\tilde{g}(\mathbf{0}) = (\det \mathfrak{H}_{z^*})^{-1/2}, \tag{36}$$

which is positive, since  $\mathfrak{H}_{z^*}$  is positive definite on  $\mathcal{J}_p$ , as discussed in Section 2. Again, for  $\mathcal{B}_z$  small enough, it also holds (up to a null set) that  $\tilde{\mathcal{B}}_z \cap (\mathbb{R}_+^2 \times \mathbb{R}) = \tilde{\mathcal{B}}_z \cap \tilde{\mathcal{D}}_z$ , on which we know from Lemma 7 and Lemma 17 that  $\mathbf{0} = \mathcal{J}(z^{**})$  is the unique infimum of  $\tilde{f}$  since

$$\tilde{f}(\mathbf{0}) = \mathcal{I}_S \circ \mathcal{J}^{-1}(\mathbf{0}) = \mathcal{I}_S(z^{**}) = \Lambda_p^*(z^*). \tag{37}$$

We can see from (34) that all partial derivatives of  $\mathcal{J}^{-1}$  are continuously differentiable in a sufficiently small neighbourhood of  $\mathbf{0}$ . Thereby,  $\mathcal{J}^{-1}$  is twice continuously differentiable in such a neighbourhood. The two-fold continuous differentiability of  $\Lambda_p^*$  has already been shown in the proof of Theorem 11. Finally, by Lemma 21 i), it holds that

$$\begin{aligned} \nabla_{(x_1, x_2, y)} \mathcal{I}_S(z^{**}) &= \left( \frac{\partial}{\partial x_1} \Lambda_p^*(x_1, x_2), \frac{\partial}{\partial x_2} \Lambda_p^*(x_1, x_2), -\frac{1}{y} \right) \Big|_{(x_1, x_2, y)=z^{**}} \\ &= \left( \tau(x)_1, \tau(x)_2, -\frac{1}{y} \right) \Big|_{(x_1, x_2, y)=z^{**}} \end{aligned}$$

$$= (\tau(z^*)_1, \tau(z^*)_2, -1),$$

from which we can deduce that

$$\begin{aligned} \nabla_t \tilde{f}(\mathbf{0}) &= \nabla_{(x_1, x_2, y)} \mathcal{I}_{\mathcal{S}}(z^{**}) J_t \mathcal{J}^{-1}(\mathbf{0}) \\ &= (\tau(z^*)_1, \tau(z^*)_2, -1) \begin{pmatrix} 1 & qz^q & z^q \frac{q}{p} \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \\ &= \left( \tau(z^*)_1, qz^q \tau(z^*)_1 + 1, z^q \frac{q}{p} \tau(z^*)_1 + \tau(z^*)_2 \right). \end{aligned} \tag{38}$$

It thereby follows that  $\nabla_t \tilde{f}(\mathbf{0}) \neq \mathbf{0}$ , as the first two components cannot be equal to zero simultaneously. But since  $\tilde{f}(t)$  attains its infimum on  $\tilde{\mathcal{B}}_z \cap (\mathbb{R}_+^2 \times \mathbb{R})$  in  $t = \mathbf{0}$ , it holds that  $\tilde{f}_{[1,0,0]} > 0$  and  $\tilde{f}_{[0,1,0]} > 0$ , as otherwise a step into either direction  $t_1, t_2$  would maintain or decrease the value of  $\tilde{f}$ , contradicting the unique infimum property of  $\mathbf{0}$ . On the other hand, by the same argument, it has to hold that  $\tilde{f}_{[0,0,1]} = 0$  and  $\tilde{f}_{[0,0,2]} > 0$ , as otherwise a step into either direction  $t_3, (-t_3)$  would maintain or decrease  $\tilde{f}$ , again contradicting the unique infimum property of  $\mathbf{0}$ . Hence, we have shown all conditions for Proposition 22, whereby it follows for the integral in (35) that

$$\begin{aligned} \mathbb{P}(\mathcal{S}^{(n)} \in \mathcal{D}_z \cap \mathcal{B}_z) &= \frac{n^2}{2\pi} \int_{\tilde{\mathcal{D}}_z \cap \tilde{\mathcal{B}}_z} \tilde{g}(t) e^{-n \tilde{f}(t)} dt (1 + o(1)) \\ &= \frac{n^2}{2\pi} \int_{\tilde{\mathcal{D}}_z \cap (\mathbb{R}_+^2 \times \mathbb{R})} \tilde{g}(t) e^{-n \tilde{f}(t)} dt (1 + o(1)) \\ &= \frac{1}{\sqrt{2\pi n}} \frac{\tilde{g}(\mathbf{0})}{\tilde{f}_{[1,0,0]} \tilde{f}_{[0,1,0]} \sqrt{\tilde{f}_{[0,0,2]}}} e^{-n \tilde{f}(\mathbf{0})} (1 + o(1)). \end{aligned} \tag{39}$$

The final term that remains to be calculated explicitly is  $\tilde{f}_{[0,0,2]}$ , as  $\tilde{f}_{[1,0,0]}$  and  $\tilde{f}_{[0,1,0]}$  are given in (38). We start by noting that

$$\begin{aligned} \frac{\partial}{\partial t_3} \mathcal{J}^{-1}(t) \Big|_{t=\mathbf{0}} &= \left( \frac{z^q \frac{q}{p} (t_3 + 1)^{(q/p)-1}}{(1 - t_2)^q}, 1, 0 \right) \Big|_{t=\mathbf{0}} \\ &= \left( z^q \frac{p}{q}, 1, 0 \right), \end{aligned}$$

and

$$\begin{aligned} \frac{\partial^2}{\partial^2 t_3} \mathcal{J}^{-1}(t) \Big|_{t=\mathbf{0}} &= \left( \frac{z^q \frac{q}{p} \left(\frac{q}{p} - 1\right) (t_3 + 1)^{(q/p)-2}}{(1 - t_2)^q}, 0, 0 \right) \Big|_{t=\mathbf{0}} \\ &= \left( \frac{z^q q^2}{p^2} - \frac{z^q q}{p}, 0, 0 \right). \end{aligned}$$

By Lemma 21 ii), we get that

$$\mathcal{H}_{(x_1, x_2, y)} \mathcal{I}_{\mathcal{S}}(z^{**}) = \begin{pmatrix} (\mathfrak{H}_{z^{**}}^{-1})_{11} & (\mathfrak{H}_{z^{**}}^{-1})_{12} & 0 \\ (\mathfrak{H}_{z^{**}}^{-1})_{21} & (\mathfrak{H}_{z^{**}}^{-1})_{22} & 0 \\ 0 & 0 & y^{-2} \end{pmatrix}.$$

It thereby follows that

$$\begin{aligned} \tilde{f}_{[0,0,2]} &= \frac{\partial^2}{\partial t_3^2} \mathcal{I}_{\mathcal{S}} \circ \mathcal{J}^{-1}(\mathbf{0}) \\ &= \frac{\partial}{\partial t_3} \left[ \nabla_{(x_1, x_2, y)} \mathcal{I}_{\mathcal{S}}(\mathcal{J}^{-1}(t)) \frac{\partial}{\partial t_3} \mathcal{J}^{-1}(t) \right] \Big|_{t=\mathbf{0}} \\ &= \frac{\partial}{\partial t_3} \left[ \nabla_{(x_1, x_2, y)} \mathcal{I}_{\mathcal{S}}(\mathcal{J}^{-1}(t)) \right] \Big|_{t=\mathbf{0}} \frac{\partial}{\partial t_3} \mathcal{J}^{-1}(\mathbf{0}) \\ &\quad + \nabla_{(x_1, x_2, y)} \mathcal{I}_{\mathcal{S}}(z^{**}) \frac{\partial^2}{\partial t_3^2} \mathcal{J}^{-1}(\mathbf{0}) \\ &= \left( z^q \frac{q}{p}, 1, 0 \right) \mathcal{H}_{(x_1, x_2, y)} \mathcal{I}_{\mathcal{S}}(z^{**}) \left( z^q \frac{q}{p}, 1, 0 \right) \\ &\quad + \left( \tau(z^*)_1, \tau(z^*)_2, -1 \right) \left( \frac{z^q q^2}{p^2} - \frac{z^q q}{p}, 0, 0 \right) \\ &= \left( \frac{z^q q}{p}, 1, 0 \right) \left( \frac{z^q q}{p} (\mathfrak{H}_{z^{**}}^{-1})_{11} + (\mathfrak{H}_{z^{**}}^{-1})_{12}, \frac{z^q q}{p} (\mathfrak{H}_{z^{**}}^{-1})_{21} + (\mathfrak{H}_{z^{**}}^{-1})_{22}, 0 \right) \\ &\quad + \tau(z^*)_1 \left( \frac{z^q q^2}{p^2} - \frac{z^q q}{p} \right) \\ \tilde{f}_{[0,0,2]} &= \frac{z^{2q} q^2}{p^2} (\mathfrak{H}_{z^{**}}^{-1})_{11} + \frac{2z^q q}{p} (\mathfrak{H}_{z^{**}}^{-1})_{12} + (\mathfrak{H}_{z^{**}}^{-1})_{22} \\ &\quad + \tau(z^*)_1 \left( \frac{z^q q^2}{p^2} - \frac{z^q q}{p} \right). \end{aligned} \tag{40}$$

Plugging the terms from (36), (38) and (40) into the fraction in (39), we get

$$\begin{aligned} &\frac{\tilde{g}(\mathbf{0})}{\tilde{f}_{[1,0,0]} \tilde{f}_{[0,1,0]} \sqrt{\tilde{f}_{[0,0,2]}}} \\ &= (\det \mathfrak{H}_{z^*})^{-1/2} (\tau(z^*)_1)^{-1} (qz^q \tau(z^*)_1 + 1)^{-1} \\ &\quad \times \left[ \frac{z^{2q} q^2}{p^2} (\mathfrak{H}_{z^{**}}^{-1})_{11} + \frac{2z^q q}{p} (\mathfrak{H}_{z^{**}}^{-1})_{12} + (\mathfrak{H}_{z^{**}}^{-1})_{22} + \tau(z^*)_1 \left( \frac{z^q q^2}{p^2} - \frac{z^q q}{p} \right) \right]^{-1/2} \end{aligned}$$

$$\begin{aligned}
 &= \left[ \det \mathfrak{H}_{z^*} (\tau(z^*)_1)^2 (qz^q \tau(z^*)_1 + 1)^2 \right. \\
 &\quad \times \left. \left( \frac{z^{2q} q^2}{p^2} (\mathfrak{H}_{z^*}^{-1})_{11} + \frac{2z^q q}{p} (\mathfrak{H}_{z^*}^{-1})_{12} + (\mathfrak{H}_{z^*}^{-1})_{22} + \tau(z^*)_1 \frac{z^q q (q-p)}{p^2} \right) \right]^{-1/2} \\
 &= \gamma(z)^{-1}, \tag{41}
 \end{aligned}$$

with  $\gamma(z)$  as in (13). Hence, it follows with (37), (39), and (41) that

$$\mathbb{P} \left( \mathcal{S}^{(n)} \in \mathcal{D}_z \cap \mathcal{B}_z \right) = \frac{1}{\sqrt{2\pi n} \gamma(z)} e^{-n \Lambda_p^*(z^*)} (1 + o(1)). \tag{42}$$

Combining the representation from (32) with the two integral estimates from (33) and (42) shows that the integral in the complement of  $\mathcal{B}_z$  can be neglected and we have that

$$\begin{aligned}
 \mathbb{P} \left( n^{1/p-1/q} \|\mathcal{X}^{(n)}\|_q > z \right) &= \mathbb{P} \left( \mathcal{S}^{(n)} \in \mathcal{D}_z \cap \mathcal{B}_z \right) \\
 &= \frac{1}{\sqrt{2\pi n} \gamma(z)} e^{-n \Lambda_p^*(z^*)} (1 + o(1)),
 \end{aligned}$$

which proves our second main result for  $\ell_p^n$ -balls. □

### Appendix

**Proof of Lemma 5.** Let  $z > m_{p,q}$  be such that  $z^* = (z^q, 1) \in \mathcal{J}_p$ . Then

$$\mathcal{I}_{\|Z\|}(z) = \inf_{\substack{t_1, t_2 > 0 \\ t_1^{1/q} t_2^{-1/p} = z}} \Lambda_p^*(t_1, t_2) = \inf_{\tilde{t}_1, \tilde{t}_2 > 0: \tilde{t}_1 = z \tilde{t}_2} \Lambda_p^*(\tilde{t}_1^q, \tilde{t}_2^p) = \inf_{\tilde{t}_2 > 0} \Lambda_p^*(z^q \tilde{t}_2^q, \tilde{t}_2^p).$$

We set  $t_z := (z^q \tilde{t}_2^q, \tilde{t}_2^p)$ , then with (9) it follows that

$$\mathcal{I}_{\|Z\|}(z) = \inf_{\tilde{t}_2 > 0} \sup_{s \in \mathbb{R}^2} (\langle s, t_z \rangle - \Lambda_p(s)) = \inf_{\tilde{t}_2 > 0} \left[ \langle \tau(t_z), t_z \rangle - \Lambda_p(\tau(t_z)) \right].$$

Our goal is to show that the infimum is attained at  $t_z^* := z^*$ , i.e. at  $\tilde{t}_2 = 1$ . Recall the definition  $g_t(s) := \langle s, t \rangle - \Lambda_p(s)$  for  $t \in \mathcal{J}_p$  from Section 2.4. By the definition of  $\tau(t_z)$  it holds that  $g_{t_z}(s)$  attains its supremum at  $\tau(t_z)$ , thus  $\nabla_s g_{t_z}(s)|_{s=\tau(t_z)} = t_z - \nabla_s \Lambda_p(s)|_{s=\tau(t_z)} = 0$ , which gives

$$t_z = (z^q \tilde{t}_2^q, \tilde{t}_2^p) = \left( \frac{\partial}{\partial s_1} \Lambda_p(s) \Big|_{s=\tau(t_z)}, \frac{\partial}{\partial s_2} \Lambda_p(s) \Big|_{s=\tau(t_z)} \right). \tag{43}$$

We now aim to write  $\frac{\partial}{\partial s_2} \Lambda_p(s)$  with respect to  $\frac{\partial}{\partial s_1} \Lambda_p(s)$  and then use the above equations. To do so, we firstly want to reformulate  $\Lambda_p$  along the lines of [20, Lemma 5.7]. It holds that

$$\Lambda_p(s) := \log \int_{\mathbb{R}} e^{s_1 |y|^q + s_2 |y|^p} f_p(y) dy$$

$$= \log \left( \frac{1}{2p^{1/p} \Gamma(1 + \frac{1}{p})} \int_{\mathbb{R}} e^{s_1|y|^q - \frac{1}{p}(1-ps_2)|y|^p} dy \right).$$

The change of variables  $x = (1 - ps_2)^{1/p} y$  then gives

$$\begin{aligned} \Lambda_p(s) &= \log \left( (1 - ps_2)^{-1/p} \int_{\mathbb{R}} e^{\frac{s_1}{(1-ps_2)^{q/p}}|x|^q} f_p(x) dx \right) \\ &= -\frac{1}{p} \log(1 - ps_2) + \log \varphi_{|X|^q} \left( \frac{s_1}{(1 - ps_2)^{q/p}} \right), \end{aligned}$$

where  $\varphi_{|X|^q}$  is the m.g.f. of a random variable  $|X|^q$  with  $X \sim \mathbf{N}_p$ . Hence,

$$\begin{aligned} \frac{\partial}{\partial s_1} \Lambda_p(s) &= \frac{\partial}{\partial s_1} \left[ \log \varphi_{|X|^q} \left( \frac{s_1}{(1 - ps_2)^{q/p}} \right) \right] \\ &= \varphi_{|X|^q} \left( \frac{s_1}{(1 - ps_2)^{q/p}} \right)^{-1} \frac{\partial}{\partial s_1} \left[ \varphi_{|X|^q} \left( \frac{s_1}{(1 - ps_2)^{q/p}} \right) \right] \\ &= \varphi_{|X|^q} \left( \frac{s_1}{(1 - ps_2)^{q/p}} \right)^{-1} \\ &\quad \times \int_{\mathbb{R}} (1 - ps_2)^{-q/p} |x|^q e^{\frac{s_1}{(1-ps_2)^{q/p}}|x|^q} f_p(x) dx \\ &= (1 - ps_2)^{-q/p} \varphi_{|X|^q} \left( \frac{s_1}{(1 - ps_2)^{q/p}} \right)^{-1} \varphi'_{|X|^q} \left( \frac{s_1}{(1 - ps_2)^{q/p}} \right), \end{aligned}$$

where  $\varphi'_{|X|^q} \left( \frac{s_1}{(1-ps_2)^{q/p}} \right) = \varphi'_{|X|^q}(t) \Big|_{t=\frac{s_1}{(1-ps_2)^{q/p}}}$ . Moreover, with the above we get that

$$\begin{aligned} \frac{\partial}{\partial s_2} \Lambda_p(s) &= (1 - ps_2)^{-1} + \frac{\partial}{\partial s_2} \left[ \log \varphi_{|X|^q} \left( \frac{s_1}{(1 - ps_2)^{q/p}} \right) \right] \\ &= (1 - ps_2)^{-1} + \varphi_{|X|^q} \left( \frac{s_1}{(1 - ps_2)^{q/p}} \right)^{-1} \frac{\partial}{\partial s_2} \left[ \varphi_{|X|^q} \left( \frac{s_1}{(1 - ps_2)^{q/p}} \right) \right] \\ &= (1 - ps_2)^{-1} \\ &\quad + \varphi_{|X|^q} \left( \frac{s_1}{(1 - ps_2)^{q/p}} \right)^{-1} \int_{\mathbb{R}} \frac{qs_1}{(1 - ps_2)^{(q+p)/p}} |x|^q e^{\frac{s_1}{(1-ps_2)^{q/p}}|x|^q} f_p(x) dx \\ &= (1 - ps_2)^{-1} \\ &\quad + \frac{qs_1}{(1 - ps_2)^{(q+p)/p}} \varphi_{|X|^q} \left( \frac{s_1}{(1 - ps_2)^{q/p}} \right)^{-1} \varphi'_{|X|^q} \left( \frac{s_1}{(1 - ps_2)^{q/p}} \right) \\ &= (1 - ps_2)^{-1} \end{aligned}$$

$$\begin{aligned}
& + \frac{qs_1}{1-ps_2} (1-ps_2)^{-q/p} \varphi_{|X|q} \left( \frac{s_1}{(1-ps_2)^{q/p}} \right)^{-1} \varphi'_{|X|q} \left( \frac{s_1}{(1-ps_2)^{q/p}} \right) \\
& = (1-ps_2)^{-1} + qs_1(1-ps_2)^{-1} \frac{\partial}{\partial s_1} \Lambda_p(s). \tag{44}
\end{aligned}$$

Plugging in the identities from (43) into (44) it follows for  $s = (\tau(t_z)_1, \tau(t_z)_2)$ :

$$\tilde{t}_2^p = (1-p\tau(t_z)_2)^{-1} + q\tau(t_z)_1(1-p\tau(t_z)_2)^{-1} z^q \tilde{t}_2^q. \tag{45}$$

Using this, we can calculate the derivative of  $\Lambda_p^*(t_z)$  in  $t$  (we write  $t$  instead of  $\tilde{t}_2$  for notational brevity), where  $\tau(t_z)$  is considered as a function in  $t$  as well. It holds that

$$\begin{aligned}
\frac{\partial}{\partial t} \Lambda_p^*(t_z) &= \frac{\partial}{\partial t} \Lambda_p^*(z^q t^q, t^p) \\
&= \frac{\partial}{\partial t} [(t_z, \tau(t_z)) - \Lambda_p(\tau(t_z))] \\
&= \frac{\partial}{\partial t} [z^q t^q \tau(t_z)_1 + t^p \tau(t_z)_2 - \Lambda_p(\tau(t_z))] \\
&= z^q q t^{q-1} \tau(t_z)_1 + z^q t^q \frac{\partial}{\partial t} \tau(t_z)_1 + p t^{p-1} \tau(t_z)_2 + t^p \frac{\partial}{\partial t} \tau(t_z)_2 \\
&\quad - \frac{\partial}{\partial t} \Lambda_p(\tau(t_z)) \\
&= z^q q t^{q-1} \tau(t_z)_1 + z^q t^q \frac{\partial}{\partial t} \tau(t_z)_1 + p t^{p-1} \tau(t_z)_2 + t^p \frac{\partial}{\partial t} \tau(t_z)_2 \\
&\quad - J_t(\tau(t_z)) \nabla_s \Lambda_p(s) \Big|_{s=\tau(t_z)} \\
&= z^q q t^{q-1} \tau(t_z)_1 + z^q t^q \frac{\partial}{\partial t} \tau(t_z)_1 + p t^{p-1} \tau(t_z)_2 + t^p \frac{\partial}{\partial t} \tau(t_z)_2 \\
&\quad - \frac{\partial}{\partial t} \tau(t_z)_1 \frac{\partial}{\partial s_1} \Lambda_p(s) \Big|_{s=\tau(t_z)} - \frac{\partial}{\partial t} \tau(t_z)_2 \frac{\partial}{\partial s_2} \Lambda_p(s) \Big|_{s=\tau(t_z)}.
\end{aligned}$$

We now use the identity from (43), which yields

$$\begin{aligned}
\frac{\partial}{\partial t} \Lambda_p^*(t_z) &= z^q q t^{q-1} \tau(t_z)_1 + z^q t^q \frac{\partial}{\partial t} \tau(t_z)_1 + p t^{p-1} \tau(t_z)_2 + t^p \frac{\partial}{\partial t} \tau(t_z)_2 \\
&\quad - \frac{\partial}{\partial t} \tau(t_z)_1 z^q t^q - \frac{\partial}{\partial t} \tau(t_z)_2 t^p \\
&= z^q q t^{q-1} \tau(t_z)_1 + p t^{p-1} \tau(t_z)_2. \tag{46}
\end{aligned}$$

Reformulating the identity in (45) yields



$$\begin{aligned}
 t^p &= (1 - p\tau(t_z)_2)^{-1} + q\tau(t_z)_1(1 - p\tau(t_z)_2)^{-1} z^q t^q \\
 \Leftrightarrow \quad &(1 - p\tau(t_z)_2)t^{p-1} - t^{-1} = z^q t^{q-1} q\tau(t_z)_1.
 \end{aligned} \tag{47}$$

Thus, if we set  $\frac{\partial}{\partial t} \Lambda_p^*(t_a) = 0$ , we get from (46) and (47) that

$$\begin{aligned}
 \frac{\partial}{\partial t} \Lambda_p^*(t_z) = 0 &\Leftrightarrow 0 = z^q q t^{q-1} \tau(t_z)_1 + p t^{p-1} \tau(t_z)_2 \\
 &\Leftrightarrow 0 = (1 - p\tau(t_z)_2)t^{p-1} - t^{-1} + p t^{p-1} \tau(t_z)_2 \\
 &\Leftrightarrow t = 1.
 \end{aligned}$$

Hence, the infimum of  $\Lambda_p^*$  over  $\partial D_z$  is attained at  $t_z^* = (z^q, 1) = z^*$ . Since  $\Lambda_p^*$  is strictly convex (see properties of the Legendre–Fenchel transform), this minimum is unique. Thereby, our claim is proven.  $\square$

**Proof of Lemma 7.** Let  $z > m_{p,q}$  be such that  $z^* = (z^q, 1) \in \mathcal{J}_p$ . Furthermore, set  $z^{**} := (z^q, 1, 1)$  and  $\mathcal{I}_S(t) := [\Lambda_p^*(t_1, t_2) - \log(t_3)]$ ,  $t \in \mathbb{R}^3$ . We use the definitions of  $\mathcal{I}_{\|\mathcal{X}\|}$  and  $\mathcal{I}_U$ , together with Lemma 5, to get that

$$\begin{aligned}
 \mathcal{I}_{\|\mathcal{X}\|}(z) &= \inf_{\substack{z = t_1^{1/q} t_2^{-1/p} t_3 \\ t_1, t_2 > 0, t_3 \in (0, 1)}} \mathcal{I}_S(t) \\
 &= \inf_{\substack{z = z_1 z_2 \\ z_1 > 0, z_2 \in (0, 1)}} \left[ \inf_{\substack{t_1, t_2 > 0 \\ t_1^{1/q} t_2^{-1/p} = z_1}} \Lambda_p^*(t_1, t_2) + \mathcal{I}_U(z_2) \right] \\
 &= \inf_{\substack{z = z_1 z_2 \\ z_1 > 0, z_2 \in (0, 1)}} \left[ \Lambda_p^*(z_1^q, 1) - \log(z_2) \right].
 \end{aligned}$$

By the same arguments as in the proof of Lemma 17, we know that  $\mathcal{I}_{\|\mathcal{X}\|}(z) = \Lambda_p^*(z^q, 1)$  is strictly convex in  $z$  on  $\mathcal{J}_p$  with a unique root in  $m_{p,q}$ . Hence, it follows that for  $z > m_{p,q}$  with  $z \in \mathcal{J}_p$  the function  $\mathcal{I}_{\|\mathcal{X}\|}(z) = \Lambda_p^*(z^q, 1)$  is strictly increasing in  $z$ . Since  $z_2 \leq 1$ ,  $z = z_1 z_2$ , and  $1 < q$ , we have  $z_1^q \geq z > m_{p,q}$ , meaning that  $\Lambda_p^*(z_1^q, 1)$  is strictly increasing in  $z_1$ . Furthermore, we can see that  $-\log(z_2)$  is strictly decreasing in  $z_2$ . Hence, rewriting  $z_1$  with respect to  $z_2$  then gives

$$\mathcal{I}_{\|\mathcal{X}\|}(z) = \inf_{\substack{z_1 = z/z_2 \\ z_2 \in (0, 1)}} \left[ \Lambda_p^* \left( \left( \frac{z}{z_2} \right)^q, 1 \right) - \log(z_2) \right],$$

which is strictly decreasing in  $z_2$ . Thus, choosing  $z_2 = 1$  gives  $z_1 = z$  and

$$\mathcal{I}_{\|\mathcal{X}\|}(z) = \mathcal{I}_S(z^{**}) = \Lambda_p^*(z^*),$$

finishing the proof.  $\square$

### Acknowledgments

The author would like to thank Kavita Ramanan and Joscha Prochno for the insightful exchanges on the topic of sharp large deviations in asymptotic geometric analysis.

Furthermore, the author would like to thank his supervisor Christoph Thäle for the helpful discussions, feedback and constructive criticism throughout the writing of this paper.

## References

- [1] Adriani, C., Baldi, P.: Sharp estimates of deviations of the sample mean in many dimensions. *Ann. Inst. Henri Poincaré Probab. Stat.* **33**(3), 371–385 (1997). [MR1457057](#)
- [2] Alonso-Gutiérrez, D., Prochno, J.: Thin-shell concentration for random vectors in Orlicz balls via moderate deviations and Gibbs measures. [arXiv:2011.07523](#) (2020)
- [3] Alonso-Gutiérrez, D., Prochno, J., Thäle, C.: Large deviations for high-dimensional random projections of  $\ell_p^n$ -balls. *Adv. Appl. Math.* **99**, 1–35 (2018). [MR3806754](#). <https://doi.org/10.1016/j.aam.2018.04.003>
- [4] Alonso-Gutiérrez, D., Prochno, J., Thäle, C.: Gaussian fluctuations for high-dimensional random projections of  $\ell_p^n$ -balls. *Bernoulli* **25**(4A), 3139–3174 (2019). [MR4003577](#). <https://doi.org/10.3150/18-BEJ1084>
- [5] Anttila, M., Ball, K., Perissinaki, I.: The central limit problem for convex bodies. *Trans. Am. Math. Soc.* **355**(12), 4723–4735 (2003). [MR1997580](#). <https://doi.org/10.1090/S0002-9947-03-03085-X>
- [6] Artstein-Avidan, S., Giannopoulos, A., Milman, V.D.: *Asymptotic Geometric Analysis. Part I. Mathematical Surveys and Monographs*, vol. 202, p. 451. American Mathematical Society, Providence, RI (2015). [MR3331351](#). <https://doi.org/10.1090/surv/202>
- [7] Bahadur, R.R., Ranga Rao, R.: On deviations of the sample mean. *Ann. Math. Stat.* **31**(4), 1015–1027 (1960). [MR0117775](#). <https://doi.org/10.1214/aoms/1177705674>
- [8] Barthe, F., Guédon, O., Mendelson, S., Naor, A.: A probabilistic approach to the geometry of the  $\ell_p^n$ -ball. *Ann. Probab.* **33**(2), 480–513 (2005). [MR2123199](#). <https://doi.org/10.1214/009117904000000874>
- [9] Bleistein, N., Handelsmann, R.A.: *Asymptotic Expansions of Integrals*. Dover Publications, Inc., New York (1975). [MR0863284](#)
- [10] Borovkov, A.A., Rogozin, B.A.: On the central limit theorem in the higher-dimensional case. *Teor. Veroâtn. Primen.* **10**, 61–69 (1965). [MR0173279](#)
- [11] Brazitikos, S., Giannopoulos, A., Valettas, P., Vritsiou, B.-H.: *Geometry of Isotropic Convex Bodies. Mathematical Surveys and Monographs*, vol. 196, p. 594. American Mathematical Society, Providence, RI (2014). [MR3185453](#). <https://doi.org/10.1090/surv/196>
- [12] Breitung, K., Hohenbichler, M.: Asymptotic approximations for multivariate integrals with an application to multinormal probabilities. *J. Multivar. Anal.* **30**(1), 80–97 (1989). [MR1003710](#). [https://doi.org/10.1016/0047-259X\(89\)90089-4](https://doi.org/10.1016/0047-259X(89)90089-4)
- [13] Cramér, H.: Sur un nouveau théorème-limite de la théorie des probabilités. *Actual. Sci. Ind.* **736**, 5–23 (1938)
- [14] Daniels, H.E.: Saddlepoint approximations in statistics. *Ann. Math. Stat.* **25**(4), 631–650 (1954). [MR0066602](#). <https://doi.org/10.1214/aoms/1177728652>
- [15] DasGupta, A.: *Fundamentals of Probability: a First Course*. Springer (2010). [MR2583199](#). <https://doi.org/10.1007/978-1-4419-5780-1>
- [16] Debye, P.: Nährungsformeln für die Zylinderfunktionen für große Werte des Arguments und unbeschränkt veränderliche Werte des Index. *Math. Ann.* **67**, 535–558 (1909). [MR1511547](#). <https://doi.org/10.1007/BF01450097>

- [17] Dembo, A., Zeitouni, O.: Large Deviations. Techniques and Applications. Stochastic Modelling and Applied Probability, vol. 38, p. 396. Springer (2010). [MR2571413](#). <https://doi.org/10.1007/978-3-642-03311-7>. Corrected reprint of the second (1998) edition.
- [18] den Hollander, F.: Large Deviations. Fields Institute Monographs, vol. 14, p. 143. American Mathematical Society, Providence, RI (2000). [MR1739680](#)
- [19] Esscher, F.: On the probability function in the collective theory of risk. *Scand. Actuar. J.* **1932**(3), 175–195 (1932). <https://doi.org/10.1080/03461238.1932.10405883>
- [20] Gantert, N., Kim, S.S., Ramanan, K.: Large deviations for random projections of  $\ell^p$  balls. *Ann. Probab.* **45**, 4419–4476 (2017). [MR3737915](#). <https://doi.org/10.1214/16-AOP1169>
- [21] Goldman, R.: Curvature formulas for implicit curves and surfaces. *Comput. Aided Geom. Des.* **22**(7), 632–658 (2005). *Geometric Modelling and Differential Geometry*. [MR2169053](#). <https://doi.org/10.1016/j.cagd.2005.06.005>
- [22] Gromov, M., Milman, V.D.: Generalization of the spherical isoperimetric inequality to uniformly convex banach spaces. *Compos. Math.* **62**(3), 263–282 (1987). [MR0901393](#)
- [23] Guédon, O.: Concentration phenomena in high dimensional geometry. In: *Journées MAS 2012*. ESAIM Proc., vol. 44, pp. 47–60. EDP Sci., Les Ulis (2014). [MR3178607](#). <https://doi.org/10.1051/proc/201444002>
- [24] Guédon, O., Nayar, P., Tkocz, T.: Concentration inequalities and geometry of convex bodies. In: *Analytical and Probabilistic Methods in the Geometry of Convex Bodies*. IMPAN Lect. Notes, vol. 2, pp. 9–86. Polish Acad. Sci. Inst. Math., Warsaw (2014). [MR3329056](#)
- [25] Hicks, N.J.: Notes on Differential Geometry. Van Nostrand Mathematical Studies, vol. 3. D. Van Nostrand Company, Inc., Princeton, N.J.-Toronto-New York-London (1965). VI, 183 p. [MR0179691](#)
- [26] Jensen, J.L.: Saddlepoint Approximations. Oxford Science Publications. Clarendon Press (1995). [MR1354837](#)
- [27] Kabluchko, Z., Prochno, J.: The maximum entropy principle and volumetric properties of Orlicz balls. *J. Math. Anal. Appl.* **495**(1), 124687 (2021). [MR4172842](#). <https://doi.org/10.1016/j.jmaa.2020.124687>
- [28] Kabluchko, Z., Prochno, J., Thäle, C.: High-dimensional limit theorems for random vectors in  $\ell_p^n$ -balls. *Commun. Contemp. Math.* **21**(1), 1750092, 30 p. (2019). [MR3904638](#). <https://doi.org/10.1142/S0219199717500924>
- [29] Kabluchko, Z., Prochno, J., Thäle, C.: Sanov-type large deviations in Schatten classes. *Ann. Inst. Henri Poincaré Probab. Stat.* **56**(2), 928–953 (2019). [MR4076771](#). <https://doi.org/10.1214/19-AIHP989>
- [30] Kabluchko, Z., Prochno, J., Thäle, C.: High-dimensional limit theorems for random vectors in  $\ell_p^n$ -balls. II. *Commun. Contemp. Math.* (to appear) (2019). [MR4216415](#). <https://doi.org/10.1142/S0219199719500731>
- [31] Kallenberg, O.: Foundations of Modern Probability, 2nd edn. Probability and its Applications, p. 638. Springer (2002). [MR1876169](#). <https://doi.org/10.1007/978-1-4757-4015-8>
- [32] Kim, S.S.: Problems at the interface of probability and convex geometry: Random projections and constrained processes. Ph.D. thesis, Brown University (2017)
- [33] Kim, S.S., Ramanan, K.: A conditional limit theorem for high-dimensional  $\ell^p$ -spheres. *J. Appl. Probab.* **55**, 1060–1077 (2018). [MR3899928](#). <https://doi.org/10.1017/jpr.2018.71>
- [34] Klartag, B.: A central limit theorem for convex sets. *Invent. Math.* **168**(1), 91–131 (2007). [MR2285748](#). <https://doi.org/10.1007/s00222-006-0028-8>

- [35] Klartag, B.: Power-law estimates for the central limit theorem for convex sets. *J. Funct. Anal.* **245**(1), 284–310 (2007). [MR2311626](#). <https://doi.org/10.1016/j.jfa.2006.12.005>
- [36] Klingenberg, W., Hoffman, D.: *A Course in Differential Geometry*. Graduate Texts in Mathematics. Springer (2013). [MR0474045](#)
- [37] Liao, Y.-T., Ramanan, K.: Geometric sharp large deviations for random projections of  $\ell_p^n$  spheres and balls. [arXiv:2001.04053v2](#) (2020).
- [38] Naor, A.: The surface measure and cone measure on the sphere of  $\ell_p^n$ . *Trans. Am. Math. Soc.* **359**(3), 1045–1079 (2007). [MR2262841](#). <https://doi.org/10.1090/S0002-9947-06-03939-0>
- [39] Petrov, V.V.: On the probabilities of large deviations for sums of independent random variables. *Teor. Veróatn. Primen.* **10**, 310–322 (1965). [MR0185645](#)
- [40] Prochno, J., Thäle, C., Turchi, N.: Geometry of  $\ell_p^n$ -balls: Classical results and recent developments. In: *High Dimensional Probability VIII*. Progress in Probability. N. Gozlan, R. Latala, K. Loucini, M. Madiman eds. Birkhäuser (2019). [MR4181365](#)
- [41] Rachev, S.T., Rüschendorf, L.: Approximate independence of distributions on spheres and their stability properties. *Ann. Probab.* **19**(3), 1311–1337 (1991). [MR1112418](#)
- [42] Richter, W.: Local limit theorems for large deviations. *Dokl. Akad. Nauk SSSR* **115**, 53–56 (1957). [MR0093816](#)
- [43] Richter, W.: Mehrdimensionale lokale Grenzwertsätze für grosse Abweichungen. *Teor. Veróatn. Primen.* **3**, 107–114 (1958). [MR0093814](#)
- [44] Rockafellar, R.T.: *Convex Analysis*, vol. 36. Princeton University Press (1970). [MR0274683](#)
- [45] Schechtman, G., Schmuckenschläger, M.: Another remark on the volume of the intersection of two  $L_p^n$  balls. In: *Geometric Aspects of Functional Analysis* (1989–90). Lecture Notes in Math., vol. 1469, pp. 174–178. Springer (1991). [MR1122622](#)
- [46] Schechtman, G., Zinn, J.: On the volume of the intersection of two  $L_p^n$  balls. *Proc. Am. Math. Soc.* **110**(1), 217–224 (1990). [MR1015684](#). <https://doi.org/10.2307/2048262>
- [47] Schechtman, G., Zinn, J.: Concentration on the  $l_p^n$  ball. In: *Geometric Aspects of Functional Analysis*. Lecture Notes in Math., vol. 1745, pp. 245–256. Springer (2000). [MR1796723](#). <https://doi.org/10.1007/BFb0107218>
- [48] Schmuckenschläger, M.: CLT and the volume of intersections of  $l_p^n$ -balls. *Geom. Dedic.* **85**(1-3), 189–195 (2001). [MR1845607](#). <https://doi.org/10.1023/A:1010353121014>
- [49] Vershynin, R.: *High-Dimensional Probability: An Introduction with Applications in Data Science*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press (2018). [MR3837109](#). <https://doi.org/10.1017/9781108231596>