LARGE DEVIATIONS FOR HIGH-DIMENSIONAL
RANDOM PROJECTIONS OF ℓ_p^n-BALLS

DAVID ALONSO-GUTIÉRREZ, JOSCHA PROCHNO, AND CHRISTOPH THÄLE

ABSTRACT. The paper provides a description of the large deviation behavior for the Euclidean norm of projections of ℓ_p^n-balls to high-dimensional random subspaces. More precisely, for each integer $n \geq 1$, let $k_n \in \{1, \ldots, n-1\}$, $E^{(n)}$ be a uniform random k_n-dimensional subspace of $\mathbb{R}^n$ and $X^{(n)}$ be a random point that is uniformly distributed in the ℓ_p^n-ball of $\mathbb{R}^n$ for some $p \in [1, \infty]$. Then the Euclidean norms $\|P_{E^{(n)}} X^{(n)}\|_2$ of the orthogonal projections are shown to satisfy a large deviation principle as the space dimension n tends to infinity. Its speed and rate function are identified, making thereby visible how they depend on p and the growth of the sequence of subspace dimensions k_n. As a key tool we prove a probabilistic representation of $\|P_{E^{(n)}} X^{(n)}\|_2$ which allows us to separate the influence of the parameter p and the subspace dimension k_n.

CONTENTS

1. Introduction

The geometry of convex bodies in high dimensions is a fascinating and vivid field at the core of what is known today as Asymptotic Geometric Analysis, a branch of mathematics at the crossroads between analysis, geometry and probability. In particular, it has been realized in the last decades that the presence of high dimensions forces certain regularity on the geometry of convex bodies that in many instances has a probabilistic flavor, compare with the surveys of Guédon [16, 17] and the monograph [7], for example. The arguably most prominent example is the central limit theorem, which is widely known in probability theory to capture the fluctuations of a sum of (independent) random variables (see, e.g., Chapter 5 in [18]). In the geometric context it roughly says that most k-dimensional marginals of a high-dimensional isotropic convex body are approximately Gaussian, provided that k is of smaller order than n^κ for some universal constant $\kappa \in (0,1)$, i.e., $k = o(n^\kappa)$. The central limit theorem for convex bodies was conjectured in [1] by Anttila, Ball and Perissinaki (for $k = 1$), who proved the conjecture for the case of uniform distributions on convex sets whose modulus of convexity and diameter satisfy some additional quantitative assumptions. Other contributions to different facets of the central limit problem for (special classes of) convex bodies are due to Bobkov and Koldobsky [5], Brehm, Hinow, Vogt and Voigt [8], E. Meckes [23, 24], E. and M. Meckes [25], E. Milman [27] or Paouris [29], just to mention a few. For general bodies, based on a principle going back to the work of Sudakov [34], and Diaconis and Freedman [12], a central limit theorem was proved by Klartag in [20, 21], who obtained that $\kappa \geq 1/15$. If in addition the convex body is 1-unconditional, that is, symmetric with respect to all coordinate hyperplanes, this has been extended by M. Meckes [26] to k-dimensional marginals with $k = o(n^{1/3})$. In particular, this class of convex bodies includes the ℓ_p^n-balls considered in the present text.

On the one hand the central limit theorem underlines the universal behavior of Gaussian fluctuations. On the other hand, it is widely known in probability theory that the so-called large deviation behavior, which considers fluctuations beyond the Gaussian scale, is much more sensitive to the distributions of the involved random variables. For example, Cramér's theorem (see, e.g., [10, Theorem 2.2.3] or [18, Theorem 27.5]) guarantees that if $X, X_1, X_2, \ldots$ are independent, identically distributed and centered random variables with cumulant generating function $\Lambda(u) := \log(\mathbb{E}e^{uX}) < \infty$ for all $u \in \mathbb{R}$, one has that

$$\lim_{n \to \infty} \frac{1}{n} \log \mathbb{P}(X_1 + \ldots + X_n \geq nt) = -\Lambda^*(t)$$

for all $t > \mathbb{E}X$, where Λ^* is the Legendre-Fenchel transform of Λ. Equivalently, this means that for any $\varepsilon > 0$ and any $t > \mathbb{E}X$ there exists some natural number n_0 so that for each $n \geq n_0$,

$$e^{-n(\Lambda^*(t)+\varepsilon)} \leq \mathbb{P}(X_1 + \ldots + X_n \geq nt) \leq e^{-n(\Lambda^*(t)-\varepsilon)}.$$

We emphasize that this function usually displays an entirely different behavior for random variables sharing the same properties on the scale of the central limit theorem. While large deviations have been investigated intensively in probability theory (see, for instance, [10, 11] and the references cited therein), they have – in sharp contrast to the central limit theorem – left almost no traces in Asymptotic Geometric Analysis so far. However and as already anticipated above, the study of large deviations of marginals of high-dimensional convex bodies might open new

perspectives and give access to non-universal features that allow to make transparent properties that distinguish between different convex bodies. In addition to the potential mentioned before, random projections of random vectors in high dimensions naturally appear in machine learning and information science, for instance, in linear regression [22] when searching for the best regression function and for the purpose of dimension reduction in information retrieval in text documents [4] to reduce the computational complexity.

It was only recently that Gantert, Kim and Ramanan [14, 13], and Kim and Ramanan [19] opened this field by deriving, in particular, a Large Deviation Principle (LDP) in the spirit of Donsker and Varadhan for 1-dimensional random projections of ℓ_p^n-balls in $\mathbb{R}^n$, as the space dimension n tends to infinity. More precisely, their results show that if for each $n \in \mathbb{N}$, $\Theta^{(n)} \in \mathbb{S}^{n-1}$ is a uniform random direction and $X^{(n)}$ is an independent random point uniformly distributed in the ℓ_p^n-ball of $\mathbb{R}^n$ for some fixed $p \in [1, \infty]$, then the sequence of rescaled random variables

$$n^{\frac{1}{p} - \frac{1}{2}} \langle X^{(n)}, \Theta^{(n)} \rangle$$

satisfies an LDP with speed $n^{\frac{2p}{2+p}}$ if $p \in [1, 2)$ and speed n if $p \in [2, \infty]$ and with a certain rate function that also depends on p (all notions and notation are explained in Section 2 below). In view of Klartag's multi-dimensional version of the central limit theorem for convex bodies (see [20, Theorem 1.3] and [21, Theorem 1.1]) it is also natural to consider projections onto higher-dimensional random subspaces as well. As a matter of fact, an understanding of the Grassmannian setting has been the basis for many deep results in the geometry of Banach spaces and asymptotic convex geometry (see, e.g., [7, 17]). The purpose of the present paper is to put the results from [13] (more precisely, the annealed framework) into a wider context and to provide a description of the large deviation behavior for the Euclidean norm of projections of ℓ_p^n-balls onto random subspaces in high dimensions. At the same time it helps to clarify the rôle of the involved parameters and contributes to a better understanding of large deviation principles and their potential in the theory of Asymptotic Geometric Analysis, which is of rising interest. The essential observation to such an extension to higher dimensions is a probabilistic representation of the Euclidean norm of a random projection (see Theorem 3.1) in the spirit of Schechtman and Zinn [32] that allows us to separate the influence of the parameter p and the subspace dimension. This representation might be of independent interest.

Let us explain our main results in more detail (again, we refer to Section 2 below for any unexplained notion or notation). We fix $p \in [1, \infty]$ and let for each $n \in \mathbb{N}$, $X^{(n)}$ be an independent random point that is uniformly distributed in the ℓ_p^n-ball of $\mathbb{R}^n$. Furthermore, we let $k_n \in \{1, \ldots, n-1\}$ be an integer and assume that $E^{(n)}$ is a random linear subspace distributed according to the Haar probability measure on the Grassmann manifold of k_n-dimensional subspaces in $\mathbb{R}^n$ which is independent of $X^{(n)}$. The sequence of random variables of interest to us are the Euclidean norms of the orthogonal projections $P_{E^{(n)}} X^{(n)}$ of $X^{(n)}$ onto $E^{(n)}$, that is,

$$\left\| P_{E^{(n)}} X^{(n)} \right\|_2, \qquad n \in \mathbb{N}.$$

We set $\| \boldsymbol{P_E X} \| := \left(n^{\frac{1}{p} - \frac{1}{2}} \| P_{E^{(n)}} X^{(n)} \|_2 \right)_{n \in \mathbb{N}}$ and note that if $k_n = 1$ for all $n \in \mathbb{N}$, this reduces to the sequence of random variables studied in [13]. We first consider

the case $p \in [2, \infty]$ and define for $p \in [2, \infty)$ the function

$$\mathcal{J}_p(y) := \inf_{\substack{x_1, x_2 > 0 \\ x_1^{1/2} x_2^{-1/p} = y}} \mathcal{I}_p^*(x_1, x_2), \qquad y \in \mathbb{R},$$

where $\mathcal{I}_p^*(x_1, x_2)$ is the Legendre-Fenchel transform of

$$\mathcal{I}_p(t_1, t_2) := \log \left(\int_{\mathbb{R}} e^{t_1 x^2 + t_2 |x|^p} f_p(x) \, \mathrm{d}x \right), \quad (t_1, t_2) \in \mathbb{R} \times \left(-\infty, \frac{1}{p} \right)$$

with $f_p(x) := (2p^{1/p} \Gamma(1 + 1/p))^{-1} e^{-|x|^p/p}$, $x \in \mathbb{R}$, being the density of a p-generalized Gaussian random variable. To handle the exceptional case $p = \infty$ simultaneously, we write $\mathcal{J}_\infty(y) := \mathcal{I}_\infty^*(y^2)$ with $\mathcal{I}_\infty^*$ being the Legendre-Fenchel transform of $\mathcal{I}_\infty(t) := \log \left(2 \int_0^1 e^{tx^2} \, \mathrm{d}x \right)$. Our first main result reads as follows.

Theorem 1.1. *Let $p \in [2, \infty]$ and assume that the limit $\lambda := \lim_{n \to \infty} \frac{k_n}{n}$ exists in $[0, 1]$. Then the sequence $\|\boldsymbol{P_E X}\|$ satisfies an LDP with speed n and rate function*

$$\mathcal{I}_{\|\boldsymbol{P_E X}\|}(y) := \begin{cases} \inf\limits_{x \geq y} \left[\frac{\lambda}{2} \log \left(\frac{\lambda x^2}{y^2} \right) + \frac{1-\lambda}{2} \log \left(\frac{1-\lambda}{1-y^2 x^{-2}} \right) + \mathcal{J}_p(x) \right] & : y > 0 \\ \mathcal{J}_p(0) & : y = 0, \lambda \in (0, 1] \\ \inf\limits_{x \geq 0} \mathcal{J}_p(x) & : y = 0, \lambda = 0 \\ +\infty & : y < 0, \end{cases}$$

where we understand the cases $\lambda \in \{0, 1\}$ as the corresponding limits.

We emphasize at this point that while the LDP in Theorem 1.1 shows a universal speed, its rate function depends in a subtle way on the underlying convex body via the parameter p.

Next, we shall discuss the special case $p = 2$, which corresponds to the Euclidean unit ball, in some more detail. First of all, in this situation the rate function can be made fully explicit and is given by

$$(1) \qquad \mathcal{I}_{\|\boldsymbol{P_E X}\|}(y) = \begin{cases} \frac{\lambda}{2} \log \left(\frac{\lambda}{y^2} \right) + \frac{1-\lambda}{2} \log \left(\frac{1-\lambda}{1-y^2} \right) & : y \in (0, 1) \\ +\infty & : \text{otherwise}, \end{cases}$$

where we understand the cases $\lambda \in \{0, 1\}$ as the corresponding limits and with 0 or 1 included in the effective domain of $\mathcal{I}_{\|\boldsymbol{P_E X}\|}$. In particular, if λ takes the value zero the rate function reduces to

$$\mathcal{I}_{\|\boldsymbol{P_E X}\|}(y) = \begin{cases} -\frac{1}{2} \log(1 - y^2) & : y \in [0, 1) \\ +\infty & : \text{otherwise} \end{cases}$$

and this is exactly the rate function that already appeared in the 1-dimensional LDP in [2, Theorem 3.4] or [13, Theorem 2.12]. In other words, this means that in the Euclidean case $p = 2$ the LDP does not 'feel' the random subspaces $E^{(n)}$ we project onto as long as their dimension is growing slowly with n, that is, if $k_n = o(n)$. The difference to the 1-dimensional projections becomes visible only in the 'truly' high-dimensional regime in which k_n is eventually proportional to n.

We now turn to the case $p \in [1, 2)$, which already for the 1-dimensional projections shows a large deviation behavior at different scales, but this time with a fully explicit rate function (see [13, Theorem 2.3]). Our next results shows that this continues to hold for high-dimensional random projections as well.

Theorem 1.2. *Let $p \in [1, 2)$ and assume that the limit $\lambda := \lim_{n \to \infty} \frac{k_n}{n}$ exists in $(0, 1]$. Then the sequence $\|\boldsymbol{P_E X}\|$ satisfies an LDP with speed $n^{p/2}$ and rate function*

$$\mathcal{I}_{\|\boldsymbol{P_E X}\|}(y) := \begin{cases} \frac{1}{p} \left(\frac{y^2}{\lambda} - m \right)^{\frac{p}{2}} & : y \geq \sqrt{\lambda m} \\ +\infty & : otherwise , \end{cases}$$

where $m = m_p := \frac{p^{p/2}}{3} \frac{\Gamma(1 + \frac{3}{p})}{\Gamma(1 + \frac{1}{p})}$.

We emphasize that for $p \in [1, 2)$ the LDP for random projections of ℓ_p^n-balls holds at a non-universal and p-dependent speed. Moreover, a comparison with [13, Theorem 2.3] shows that both, the speed and the rate function differ from those for the 1-dimensional random projections. In fact, in this situation (where $k_n = 1$ for all $n \in \mathbb{N}$) the sequence $\|\boldsymbol{P_E X}\|$ satisfies an LDP with speed $n^{\frac{2p}{2+p}}$ and rate function

$$\mathcal{I}_{\|\boldsymbol{P_E X}\|}(y) = \begin{cases} \frac{2+p}{2p} y^{\frac{2p}{2+p}} & : y \geq 0 \\ +\infty & : \text{otherwise} \end{cases}$$

Note that the rate function stated here slightly differs from the rate function in [13], since we are not dealing with signed distances in our set-up but rather with their absolute values. Note that our Theorem 1.2 leaves open the case where the subspace dimensions k_n are such that $\frac{k_n}{n} \to 0$, as $n \to \infty$. We *conjecture* that in this case the LDP for $\|\boldsymbol{P_E X}\|$ is the same as for $\|\boldsymbol{P_E X}\|$ with $k_n \equiv 1$ discussed above.

After having presented our main theorems, let us comment on the tools we are going to use in their proofs. They basically reflect a lively interplay between geometric arguments with techniques and methods from large deviation theory. As already anticipated above, the key to Theorem 1.1 and Theorem 1.2 is a new probabilistic representation of the random variables $\|P_{E^{(n)}} X^{(n)}\|_2$. Notably, in the special case that $k_n = 1$ for all $n \in \mathbb{N}$ this is different from the one that has been used in [13]. More precisely, for each $n \in \mathbb{N}$ we will identify $\|P_{E^{(n)}} X^{(n)}\|_2$ with the product of three independent random variables:

$$\|P_{E^{(n)}} X^{(n)}\|_2 = U^{1/n} \cdot Z^{(n)} \cdot G^{(n)} .$$

Here,

- U is uniformly distributed on $[0, 1]$,
- $Z^{(n)}$ is the quotient of the ℓ_2^n- and the ℓ_p^n-norm of an n-dimensional random vector consisting of independent p-generalized Gaussian random entries,
- $G^{(n)}$ is given by $(\sum_{i=1}^{k_n} g_i^2)^{1/2} / (\sum_{i=1}^{n} g_i^2)^{1/2}$ with standard Gaussian random variables $g_1, \dots, g_n$ that are independent.

The essential feature of this representation is that the parameter p influences only the random variables $Z^{(n)}$, while on the other hand the dimension parameter k_n shows up exclusively in the definition of $G^{(n)}$. This in turn allows us to study the different effects separately and paves the way to the higher-dimensional generalizations of the results in [13]. We emphasize that the representation of $\|P_{E^{(n)}} X^{(n)}\|_2$ as a product is well reflected by the rate function appearing in Theorem 1.1, which possesses the following probabilistic interpretation: while the radial part $U^{1/n}$ has no influence as already seen in the 1-dimensional case, the rate function is the infimum of the sum of two rate functions corresponding to LDPs for $Z^{(n)}$ and $G^{(n)}$.

Moreover, the latter corresponds to the rate function (1) appearing in the particular Euclidean case $p = 2$.

The rest of this paper is structured as follows. In Section 2 we introduce our notation, recall the necessary background material from large deviation theory and provide some preliminaries on the geometry of ℓ_p^n-balls. The aforementioned probabilistic representation of $\|P_{E^{(n)}} X^{(n)}\|_2$ is the content of Section 3. We prove some auxiliary LDPs in Section 4 and in the final Section 5 we eventually prove Theorem 1.1 and Theorem 1.2. Since we have in mind a broad readership, especially colleagues working in functional analysis, convex geometry and probability theory, we decided to include background material, tools and arguments from both Asymptotic Geometric Analysis and probability theory.

2. Preliminaries

2.1. **Notation.** In this paper we denote by $|A|$ the n-dimensional Lebesgue measure of a Lebesgue measurable set $A \subset \mathbb{R}^n$ and we write $\mathscr{L}(\mathbb{R}^n)$ for the σ-field of all Lebesgue measurable subsets of $\mathbb{R}^n$. The collection of Borel sets in $\mathbb{R}^n$ is denoted by $\mathscr{B}(\mathbb{R}^n)$. We supply the n-dimensional Euclidean space $\mathbb{R}^n$ with its standard inner product $\langle \cdot, \cdot \rangle$ and the Euclidean norm $\| \cdot \|_2$. The boundary, the interior and the closure of a set $A \subset \mathbb{R}^n$ are denoted by ∂A, A° and $\bar{A}$, respectively.

We write $\mathbb{B}_2^n := \{x \in \mathbb{R}^n : \|x\|_2 \leq 1\}$ for the Euclidean unit ball and $\mathbb{S}^{n-1} := \{x \in \mathbb{R}^n : \|x\|_2 = 1\}$ for the corresponding unit sphere in $\mathbb{R}^n$, and σ_{n-1} for the uniform probability measure on $\mathbb{S}^{n-1}$, that is, the normalized spherical Lebesgue measure. As subsets of $\mathbb{R}^n$ they carry natural Borel σ-fields that we denote by $\mathscr{B}(\mathbb{B}_2^n)$ and $\mathscr{B}(\mathbb{S}^{n-1})$, respectively. Moreover, we recall that

$$(2) \qquad |\mathbb{B}_2^n| = \frac{\pi^{n/2}}{\Gamma(1 + \frac{n}{2})},$$

where $\Gamma(\cdot)$ is the Gamma-function.

The group of $(n \times n)$-orthogonal matrices is denoted by $\mathcal{O}(n)$ and we let $\mathcal{SO}(n)$ be the subgroup of orthogonal $n \times n$ matrices with determinant 1. As subsets of $\mathbb{R}^{n^2}$, $\mathcal{O}(n)$ and $\mathcal{SO}(n)$ can be equipped with the trace σ-field of $\mathscr{B}(\mathbb{R}^{n^2})$. Moreover, both compact groups $\mathcal{O}(n)$ and $\mathcal{SO}(n)$ carry a unique Haar probability measure which we denote by ν and $\tilde{\nu}$, respectively. Since $\mathcal{O}(n)$ consists of two copies of $\mathcal{SO}(n)$, the measure ν can easily be derived from $\tilde{\nu}$ and vice versa.

Given $k \in \{0, 1, \ldots, n\}$, we use the symbol $\mathbb{G}_{n,k}$ to denote the Grassmannian of k-dimensional linear subspaces of $\mathbb{R}^n$. Denoting by $d_H(\cdot, \cdot)$ the Hausdorff distance we supply $\mathbb{G}_{n,k}$ with the metric $d(E, F) := d_H(B_E, B_F)$, $E, F \in \mathbb{G}_{n,k}$, where B_E and B_F stand for the Euclidean unit balls in E and F, respectively. The Borel σ-field on $\mathbb{G}_{n,k}$ induced by this metric is denoted by $\mathscr{B}(\mathbb{G}_{n,k})$ and we supply the arising measurable space $\mathbb{G}_{n,k}$ with the unique Haar probability measure $\nu_{n,k}$. It can be identified with the image measure of the Haar probability measure $\tilde{\nu}$ on $\mathcal{SO}(n)$ under the mapping $\mathcal{SO}(n) \to \mathbb{G}_{n,k}$, $T \mapsto TE_0$ with $E_0 := \text{span}(\{e_1, \ldots, e_k\})$. Here, we write $e_1 := (1, 0, \ldots, 0), e_2 := (0, 1, 0, \ldots, 0), \ldots, e_n := (0, \ldots, 0, 1) \in \mathbb{R}^n$ for the standard orthonormal basis in $\mathbb{R}^n$ and $\text{span}(\{e_1, \ldots, e_k\}) \in \mathbb{G}_{n,k}$, $k \in \{1, \ldots, n\}$, for the k-dimensional linear subspace spanned by the first k vectors of this basis.

2.2. **Large Deviation Principles.** The purpose of this section is to provide the necessary background material from large deviation theory, which may be found in

[10, 11, 18], for example. We directly start with the definition of what we understand by a full and a weak large deviation principle. We refrain from presenting these definitions in the most general possible framework and rather restrict to the set-up needed in this paper. For this reason, let $d \geq 1$ be a fixed integer and assume that the d-dimensional Euclidean space $\mathbb{R}^d$ is supplied with its standard topology. In this subsection we denote for clarity the space dimension by d instead of n in order to distinguish it from our index parameter n. Finally, we make the assumption that all random objects we are dealing with are defined on a common (and sufficiently rich) probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Definition 2.1. *Let* $\mathbf{X} := (X^{(n)})_{n \in \mathbb{N}}$ *be a sequence of random vectors taking values in* $\mathbb{R}^d$. *Further, let* $s : \mathbb{N} \to [0, \infty]$ *and* $\mathcal{I} : \mathbb{R}^d \to [0, \infty]$ *be a lower semi-continuous function with compact level sets* $\{x \in \mathbb{R}^d : \mathcal{I}(x) \leq \alpha\}$, $\alpha \in \mathbb{R}$. *We say that* $\mathbf{X}$ *satisfies a (full)* large deviation principle *with speed* $s(n)$ *and (good) rate function* $\mathcal{I}$ *if*

$$
\begin{aligned}
-\inf_{x \in A^\circ} \mathcal{I}(x) &\leq \liminf_{n \to \infty} \frac{1}{s(n)} \log(\mathbb{P}(X^{(n)} \in A)) \\
&\leq \limsup_{n \to \infty} \frac{1}{s(n)} \log(\mathbb{P}(X^{(n)} \in A)) \leq -\inf_{x \in \overline{A}} \mathcal{I}(x)
\end{aligned}
\tag{3}
$$

for all $A \in \mathscr{L}(\mathbb{R}^d)$. *Moreover, we say that* $\mathbf{X}$ *satisfies a* weak large deviation principle *with speed* $s(n)$ *and rate function* $\mathcal{I}$ *if the lower bound in* (3) *holds as stated, while the upper bound is valid only for compact sets* $A \subset \mathbb{R}^n$.

We notice that on the class of all $\mathcal{I}$-continuity sets, that is, on the class of sets $A \in \mathscr{L}(\mathbb{R}^d)$ for which $\mathcal{I}(A^\circ) = \mathcal{I}(\overline{A})$ with $\mathcal{I}(A) := \inf\{\mathcal{I}(x) : x \in A\}$, one has the exact limit relation

$$
\lim_{n \to \infty} \frac{1}{s(n)} \log(\mathbb{P}(X^{(n)} \in A)) = -\mathcal{I}(A).
$$

In our paper we use the convention that the rate function in an LDP for a sequence of random vectors $\mathbf{X}$ is denoted by $\mathcal{I}_{\mathbf{X}}$.

What separates a weak from a full LDP is the so-called exponential tightness of the sequence of random variables (see, for instance, [10, Lemma 1.2.18] and [18, Lemma 27.9]).

Proposition 2.2. *Let* $\mathbf{X} := (X^{(n)})_{n \in \mathbb{N}}$ *be a sequence of random vectors taking values in* $\mathbb{R}^d$. *Suppose that* $\mathbf{X}$ *satisfies a weak LDP with speed* $s(n)$ *and rate function* $\mathcal{I}_{\mathbf{X}}$. *Then* $\mathbf{X}$ *satisfies a full LDP if and only if* $\mathbf{X}$ *is* exponentially tight, *that is, if and only if*

$$
\inf_K \limsup_{n \to \infty} \frac{1}{s(n)} \log(\mathbb{P}(X^{(n)} \notin K)) = -\infty,
$$

where the infimum is running over all compact sets $K \subset \mathbb{R}^d$.

The following proposition (see, for instance, [10, Theorem 4.1.11]) shows that it is sufficient to prove a weak LDP for a sequence of random variables solely for sets in a basis of the underlying topological space.

Proposition 2.3. *Let* $d \in \mathbb{N}$ *and* $\mathcal{A}$ *be basis of the standard topology in* $\mathbb{R}^d$. *Let* $\mathbf{X} = (X^{(n)})_{n \in \mathbb{N}}$ *be a sequence of* $\mathbb{R}^d$-valued random vectors. *For every* $A \in \mathcal{A}$, *define*

$$
\mathcal{I}_{\mathbf{X}}^{(A)} := -\liminf_{n \to \infty} \frac{1}{s(n)} \log(\mathbb{P}(X^{(n)} \in A))
$$

and for $x \in \mathbb{R}^d$ set $\mathcal{I}_{\mathbf{X}}(x) := \sup\{\mathcal{I}_{\mathbf{X}}^{(A)} : A \in \mathcal{A}, x \in A\}$. Suppose that for all $x \in \mathbb{R}^d$,

$$\mathcal{I}_{\mathbf{X}}(x) = \sup_{\substack{A \in \mathcal{A} \\ x \in A}} \left[-\limsup_{n \to \infty} \frac{1}{s(n)} \log\left(\mathbb{P}\left(X \in A\right)\right) \right].$$

Then $\mathbf{X}$ satisfies a weak LDP with speed $s(n)$ and rate function $\mathcal{I}_{\mathbf{X}}$.

Let $d \geq 1$ be a fixed integer and let X be an $\mathbb{R}^d$-valued random vector. We write

$$\Lambda(u) = \Lambda_X(u) := \log(\mathbb{E}e^{\langle X, u \rangle}), \qquad u \in \mathbb{R}^d,$$

for the cumulant generating function of X. Moreover, we define the (effective) domain of Λ to be the set $D_\Lambda := \{u \in \mathbb{R}^d : \Lambda(u) < \infty\} \subset \mathbb{R}^d$.

Definition 2.4. *The Legendre-Fenchel transform of a convex function $\Lambda : \mathbb{R}^d \to (-\infty, +\infty]$ is defined as*

$$\Lambda^*(x) := \sup_{u \in \mathbb{R}^d} \left[\langle u, x \rangle - \Lambda(u)\right], \qquad x \in \mathbb{R}^d.$$

The Legendre-Fenchel transform of the cumulant generating function plays a crucial rôle in the following result, usually referred to as Cramér's theorem, (see, e.g., [10, Theorem 2.2.30, Theorem 6.1.3, Corollary 6.1.6] or [18, Theorem 27.5]).

Proposition 2.5 (Cramér's theorem). *Let $X, X_1, X_2, \dots$ be independent and identically distributed random vectors taking values in $\mathbb{R}^d$. Assume that the origin is an interior point of D_Λ, where Λ stands for the cumulant generating function of X. Then the partial sums $\frac{1}{n} \sum_{i=1}^{n} X_i$, $n \in \mathbb{N}$ satisfy an LDP with speed n and good rate function Λ^*.*

It will be rather important for us to deduce from an already existing large deviation principle a new one by applying various transformations. We first consider the large deviation behavior under the formation of vectors. For this, assume that d_1 and d_2 are integers and that $\mathbf{X} = (X^{(n)})_{n \in \mathbb{N}}$ is a sequence of $\mathbb{R}^{d_1}$-valued random vectors and that $\mathbf{Y} = (Y^{(n)})_{n \in \mathbb{N}}$ is a sequence of $\mathbb{R}^{d_2}$-random vectors. Assuming that $\mathbf{X}$ and $\mathbf{Y}$ satisfy large deviation principles, does then also the sequence $\mathbf{Z} := ((X^{(n)}, Y^{(n)}))_{n \in \mathbb{N}}$ of $\mathbb{R}^{d_1 + d_2}$-valued random vectors satisfy a large deviation principle and, if so, what is its rate function? The following result is only implicit in [10]. For the sake of completeness we present a self-contained proof in the appendix, since we were not able to precisely locate it in the existing literature.

Proposition 2.6. *Assume that $\mathbf{X}$ satisfies an LDP with speed $s(n)$ and good rate function $\mathcal{I}_{\mathbf{X}}$ and that $\mathbf{Y}$ satisfies an LDP with speed $s(n)$ and good rate function $\mathcal{I}_{\mathbf{Y}}$. Then, if $X^{(n)}$ and $Y^{(n)}$ are independent for every $n \in \mathbb{N}$, $\mathbf{Z} = ((X^{(n)}, Y^{(n)}))_{n \in \mathbb{N}}$ satisfies an LDP with speed $s(n)$ and rate function $\mathcal{I}_{\mathbf{Z}}$, where $\mathcal{I}_{\mathbf{Z}}(x) = \mathcal{I}_{\mathbf{X}}(x_1) + \mathcal{I}_{\mathbf{Y}}(x_2)$ for all $x = (x_1, x_2) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$.*

Next, assume that a sequence $\mathbf{X} = (X^{(n)})_{n \in \mathbb{N}}$ of random variables satisfies an LDP with speed n and rate function $\mathcal{I}_{\mathbf{X}}$. Suppose now that $\mathbf{Y} = (Y^{(n)})_{n \in \mathbb{N}}$ is a sequence of random variables that are 'close' to the ones from $\mathbf{X}$. Our aim is to transfer in such a situation the LDP from $\mathbf{X}$ to $\mathbf{Y}$. The conditions under which such an approach is working are the content of the next result, which we took from [10, Theorem 4.2.13] or [18, Lemma 27.13].

Proposition 2.7. *Let* $\mathbf{X} = (X^{(n)})_{n\in\mathbb{N}}$ *and* $\mathbf{Y} = (Y^{(n)})_{n\in\mathbb{N}}$ *be two sequence of* $\mathbb{R}^d$*-valued random vectors and assume that* $\mathbf{X}$ *satisfies an LDP with speed* $s(n)$ *and rate function* $\mathcal{I}_{\mathbf{X}}$*. Further, suppose that* $\mathbf{X}$ *and* $\mathbf{Y}$ *are* exponentially equivalent*, i.e.,*

$$\limsup_{n\to\infty} \frac{1}{s(n)} \log(\mathbb{P}(\|X^{(n)} - Y^{(n)}\|_2 > \delta)) = -\infty$$

for any $\delta > 0$*. Then* $\mathbf{Y}$ *satisfies an LDP with the same speed and the same rate function.*

Remark 2.8. *If the dimension* $d \in \mathbb{N}$ *is fixed, then, since all norms are equivalent, we may consider the* ℓ_1*-norm instead of the* ℓ_2*-norm in the definition of exponential equivalence.*

Finally, we consider the possibility to 'transport' a large deviation principle to another one by means of a continuous function. This device is known as the contraction principle and we refer to [10, Theorem 4.2.1] or [18, Theorem 27.11(i)].

Proposition 2.9 (Contraction principle)**.** *Let* $d_1, d_2 \in \mathbb{N}$ *and let* $F : \mathbb{R}^{d_1} \to \mathbb{R}^{d_2}$ *be a continuous function. Further, let* $\mathbf{X} = (X^{(n)})_{n\in\mathbb{N}}$ *be a sequence of* $\mathbb{R}^{d_1}$*-valued random vectors that satisfies an LDP with speed* $s(n)$ *and rate function* $\mathcal{I}_{\mathbf{X}}$*. Then the sequence* $\mathbf{Y} := (F(X^{(n)}))_{n\in\mathbb{N}}$ *of* $\mathbb{R}^{d_2}$*-valued random vectors satisfies an LDP with the same speed and with rate function* $\mathcal{I}_{\mathbf{Y}} = \mathcal{I}_{\mathbf{X}} \circ F^{-1}$*, i.e.,* $\mathcal{I}_{\mathbf{Y}}(y) := \inf\{\mathcal{I}_{\mathbf{X}}(x) : F(x) = y\}$*,* $y \in \mathbb{R}^{d_2}$*, with the convention that* $\mathcal{I}_{\mathbf{Y}}(y) = +\infty$ *if* $F^{-1}(\{y\}) = \emptyset$*.*

While this form of the contraction principle was sufficient to analyse the large deviation behavior for 1-dimensional random projections of ℓ_p^n-balls, we will need a refinement to treat the higher-dimensional cases. More precisely, to handle this situation we need to allow the continuous function to depend on n. The following result can be found in [10, Corollary 4.2.21].

Proposition 2.10. *Let* $d_1, d_2 \in \mathbb{N}$ *and let* $F : \mathbb{R}^{d_1} \to \mathbb{R}^{d_2}$ *be a continuous function. Suppose that* $\mathbf{X} = (X^{(n)})_{n\in\mathbb{N}}$ *is a sequence of* $\mathbb{R}^{d_1}$*-valued random variables that satisfies an LDP with speed* $s(n)$ *and rate function* $\mathcal{I}_{\mathbf{X}}$*. Further, suppose that for each* $n \in \mathbb{N}$*,* $F_n : \mathbb{R}^{d_1} \to \mathbb{R}^{d_2}$ *is a measurable function such that for all* $\delta > 0$*,* $\Gamma_{n,\delta} := \{x \in \mathbb{R}^{d_1} : \|F_n(x) - F(x)\|_2 > \delta\} \in \mathscr{L}(\mathbb{R}^{d_1})$ *and*

$$\limsup_{n\to\infty} \frac{1}{s(n)} \log(\mathbb{P}(X^{(n)} \in \Gamma_{n,\delta})) = -\infty\,.$$

Then the sequence of $\mathbb{R}^{d_2}$*-valued random variables* $(F_n(X^{(n)}))_{n\in\mathbb{N}}$ *satisfies an LDP with the same speed and with rate function* $\mathcal{I}_{\mathbf{X}} \circ F^{-1}$*.*

Finally, we recall from [11, Theorem V.6] a version of what is known as the Gärtner-Ellis theorem in large deviation theory. It can be regarded as a generalization of Cramér's theorem in different directions.

Proposition 2.11 (Gärtner-Ellis theorem)**.** *Let* $\mathbf{X} = (X^{(n)})_{n\in\mathbb{N}}$ *be a sequence of random variables with cumulant generating functions* Λ_n*. We assume that for each* $y \in \mathbb{R}$ *the limit* $\Lambda(y) := \lim_{n\to\infty} \frac{1}{n}\Lambda_n(ny)$ *exists in* $[-\infty, \infty]$ *and that* $0 \in D_\Lambda^\circ$*. We also assume that the function* Λ *is lower semi-continuous and differentiable on* $\mathbb{R}$ *and that* $D_\Lambda = \mathbb{R}$*. Then* $\mathbf{X}$ *satisfies a LDP with speed* n *and rate function* Λ^**.*

2.3. Geometry of ℓ_p^n-balls.

Let $n \geq 1$ be an integer and consider the n-dimensional Euclidean space $\mathbb{R}^n$. For any $p \in [1, \infty]$ the ℓ_p^n-norm, $\|x\|_p$, of $x = (x_1, \ldots, n) \in \mathbb{R}^n$ is given by

$$\|x\|_p := \begin{cases} \left(\sum\limits_{i=1}^n |x_i|^p \right)^{1/p} & : p < \infty \\ \max\{|x_1|, \ldots, |x_n|\} & : p = \infty \,. \end{cases}$$

Although $\|x\|_p$ depends on the space dimension n, we decided to suppress this dependency in our notation for simplicity, since n will always be clear from the context.

For any n and p let us denote by $\mathbb{B}_p^n := \{x \in \mathbb{R}^n : \|x\|_p \leq 1\}$ the ℓ_p^n-ball in $\mathbb{R}^n$ and denote by $\mathbb{S}_p^{n-1} := \{x \in \mathbb{R}^n : \|x\|_p = 1\}$ the corresponding unit sphere. The restriction of the Lebesgue measure to $\mathbb{B}_p^n$ provides a natural volume measure on $\mathbb{B}_p^n$. Although one could supply $\mathbb{S}_p^{n-1}$ with the $(n-1)$-dimensional Hausdorff measure, the so-called cone measure turns out to be more useful as explained later (see [28] for the relation between these two measures).

Definition 2.12. *For a set $A \in \mathscr{B}(\mathbb{S}_p^{n-1})$ we define*

$$\mu_p(A) := \frac{|\{rx : x \in A, r \in [0,1]\}|}{|\mathbb{B}_p^n|} \,.$$

The measure μ_p is called the cone (probability) measure of $\mathbb{B}_p^n$.

We remark that the cone measure μ_p coincides with the $(n-1)$-dimensional Hausdorff probability measure on $\mathbb{S}_p^{n-1}$ if and only if $p = 1$, $p = 2$ or $p = +\infty$. In particular, μ_2 is the same as σ_{n-1}, the normalized spherical Lebesgue measure.

The proofs of our results heavily rely on the following probabilistic representations for the volume and the cone probability measure of $\mathbb{B}_p^n$ for $p \in [1, \infty)$, which are taken from [30] and [32] (we also refer to [3] for a different representation).

Proposition 2.13. *Let $n \in \mathbb{N}$ and $p \in [1, \infty)$. Suppose that $Z_1, \ldots, Z_n$ are independent p-generalized Gaussian random variables whose distribution has density*

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

with respect to the Lebesgue measure on $\mathbb{R}$. Consider the random vector $Z := (Z_1, \ldots, Z_n) \in \mathbb{R}^n$ and define $X := Z/\|Z\|_p$. Furthermore, let U be a uniformly distributed random variable on $[0,1]$, which is independent of the Z_i's, and let us write $Y := U^{1/n}X$. Then,

- (i) *the random vector $X \in \mathbb{S}_p^{n-1}$ is independent of $\|Z\|_p$ and is distributed according to μ_p,*
- (ii) *the random vector $Y \in \mathbb{B}_p^n$ is uniformly distributed in $\mathbb{B}_p^n$.*

In the rest of this paper $(g_i)_{i \in \mathbb{N}}$ will always denote a sequence of independent real-valued standard Gaussians, U will denote an independent random variable uniformly distributed on $[0,1]$ and, for $p \geq 1$, $(Z_i)_{i \in \mathbb{N}}$ will denote a sequence of independent p-generalized Gaussian random variables with density f_p. All these random variables are assumed to be independent.

For further probabilistic aspects pertaining the geometry of ℓ_p^n-balls we refer to [3, 31, 32, 33] as well as the references cited therein.

3. A PROBABILISTIC REPRESENTATION FOR $\|P_E X\|_2$

In this section the dimension of the space n will be fixed. Thus, for simplicity in the notation, we will omit the indices that will refer to the dimension n. Fix $p \in [1, \infty)$, let X be a point chosen according to the uniform distribution on $\mathbb{B}_p^n$ and let $E \in \mathbb{G}_{n,k}$ be an independent random subspace with distribution $\nu_{n,k}$ for some $k \in \{1, \dots, n\}$. In this section we will develop the already announced probabilistic representation for $\|P_E X\|_2$, which will turn out to be crucial in the proofs of Theorems 1.1 and 1.2. The key feature of this representation is that it will allow us to identify $\|P_E X\|_2$ with a continuous function of two random variables $V_1^{(n)}$ and $W^{(n)}$. These random variables in turn can be written as functions of sums of independent identically distributed random variables. Besides, only one of them will depend on p, while the other one will depend only on the dimension k of the random subspace E. These properties, together with Cramér's theorem and the contraction principle will give us the LDPs in the main theorems.

Theorem 3.1. *For any $n \in \mathbb{N}$ and $k \in \{1, \dots, n\}$ let X be a random vector uniformly distributed in $\mathbb{B}_p^n$ for some $p \in [1, \infty)$ and let $E \in \mathbb{G}_{n,k}$ be a random subspace distributed according to $\nu_{n,k}$. Then the random variable $\|P_E X\|_2$ has the same distribution as the random variable*

$$
U^{1/n} \frac{\left(\sum_{i=1}^n Z_i^2\right)^{1/2}}{\left(\sum_{i=1}^n |Z_i|^p\right)^{1/p}} \frac{\left(\sum_{i=1}^k g_i^2\right)^{1/2}}{\left(\sum_{i=1}^n g_i^2\right)^{1/2}} .
$$

Proof. Let $x \in \mathbb{R}^n$ be a fixed vector. By construction of the Haar measure $\nu_{n,k}$ on $\mathbb{G}_{n,k}$ and uniqueness of the Haar measure ν on $\mathcal{O}(n)$, we have that, for any $t \in \mathbb{R}$,

$$
\nu_{n,k} \left(E \in \mathbb{G}_{n,k} : \|P_E x\|_2 \geq t \right) = \nu \left(T \in \mathcal{O}(n) : \|P_{T E_0} x\|_2 \geq t \right)
$$
$$
= \nu \left(T \in \mathcal{O}(n) : \|P_{E_0} T x\|_2 \geq t \right)
$$
$$
= \nu \left(T \in \mathcal{O}(n) : \|x\|_2 \left\| P_{E_0} T \frac{x}{\|x\|_2} \right\|_2 \geq t \right),
$$

where $E_0 := \operatorname{span}(\{e_1, \dots, e_k\})$. Again, by the uniqueness of the Haar measure σ_{n-1} on $\mathbb{S}^{n-1}$, $T(x/\|x\|_2)$ is a random vector uniformly distributed on $\mathbb{S}^{n-1}$ according to σ_{n-1}, provided that $T \in \mathcal{O}(n)$ has distribution ν. Thus,

$$
\nu \left(T \in \mathcal{O}(n) : \|x\|_2 \left\| P_{E_0} T \frac{x}{\|x\|_2} \right\|_2 \geq t \right) = \sigma_{n-1} \left(u \in \mathbb{S}^{n-1} : \|x\|_2 \|P_{E_0} u\|_2 \geq t \right) .
$$

Since $G = (g_1, \dots, g_n)$ is a standard Gaussian random vector in $\mathbb{R}^n$, by Proposition 2.13, the random vector $\frac{G}{\|G\|_2}$ is distributed on $\mathbb{S}^{n-1}$ according to σ_{n-1}. Thus,

$$
\sigma_{n-1} \left(u \in \mathbb{S}^{n-1} : \|x\|_2 \|P_{E_0} u\|_2 \geq t \right) = \mathbb{P} \left(\|x\|_2 \frac{\|P_{E_0} G\|_2}{\|G\|_2} \geq t \right) .
$$

Consequently, if X is a random vector uniformly distributed on $\mathbb{B}_p^n$, $E \in \mathbb{G}_{n,k}$ is a random subspace independent of X having distribution $\nu_{n,k}$, and G is a standard Gaussian random vector in $\mathbb{R}^n$ that is independent of X and E, we have that

$$
\mathbb{P}^{(X,E)} \left((x, F) \in \mathbb{B}_p^n \times \mathbb{G}_{n,k} : \|P_F x\|_2 \geq t \right)
$$
$$
= \frac{1}{|\mathbb{B}_p^n|} \int_{\mathbb{B}_p^n} \nu_{n,k} \left(E \in \mathbb{G}_{n,k} : \|P_E x\|_2 \geq t \right) \, \mathrm{d}x
$$

$$= \frac{1}{|\mathbb{B}_p^n|} \int_{\mathbb{B}_p^n} \mathbb{P}\left(\|x\|_2 \frac{\|P_{E_0}G\|_2}{\|G\|_2} \geq t\right) \mathrm{d}x$$

$$= \mathbb{P}^{(X,G)}\left((x,g) \in \mathbb{B}_p^n \times \mathbb{R}^n : \|x\|_2 \frac{\|P_{E_0}g\|_2}{\|g\|_2} \geq t\right).$$

Here, $\mathbb{P}^{(X,E)}$ denotes the joint distribution of the random vector $(X,E) \in \mathbb{B}_p^n \times \mathbb{G}_{n,k}$, while $\mathbb{P}^{(X,G)}$ stands for that of $(X,G) \in \mathbb{B}_p^n \times \mathbb{R}^n$. Now, let $Z = (Z_1, \ldots, Z_n)$ be a random vector having independent p-generalized Gaussian random entries. Then, by Proposition 2.13, the random vector $U^{1/n}\frac{Z}{\|Z\|_p}$ is uniformly distributed in $\mathbb{B}_p^n$. Therefore,

$$\mathbb{P}^{(X,G)}\left((x,g) \in \mathbb{B}_p^n \times \mathbb{R}^n : \|x\|_2 \frac{\|P_{E_0}g\|_2}{\|g\|_2} \geq t\right)$$

$$= \mathbb{P}^{(U,Z,G)}\left((u,z,g) \in [0,1] \times \mathbb{R}^n \times \mathbb{R}^n : u^{1/n}\frac{\|z\|_2}{\|z\|_p}\frac{\|P_{E_0}g\|_2}{\|g\|_2} \geq t\right)$$

with $\mathbb{P}^{(U,Z,G)}$ being the joint distribution of the random vector $(U,Z,G) \in [0,1] \times \mathbb{R}^n \times \mathbb{R}^n$. Consequently, we conclude that the two random variables

$$\|P_E X\|_2 \qquad \text{and} \qquad U^{1/n}\frac{\left(\sum_{i=1}^n Z_i^2\right)^{1/2}}{\left(\sum_{i=1}^n |Z_i|^p\right)^{1/p}}\frac{\left(\sum_{i=1}^k g_i^2\right)^{1/2}}{\left(\sum_{i=1}^n g_i^2\right)^{1/2}}$$

have the same distribution. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\Box$

4. PROOF OF AUXILIARY LDPS

The purpose of this section is to derive a number of auxiliary LDPs for the factors appearing in the probabilistic representation for $\|P_{E^{(n)}}X^{(n)}\|_2$ in Theorem 3.1. These results can be seen as intermediate steps in the proof of Theorem 1.1. Recall the set-up and the notation introduced above, define for each $n \in \mathbb{N}$ the random variables

- $V^{(n)} := \dfrac{\left(\sum_{i=1}^{k_n} g_i^2\right)^{1/2}}{\left(\sum_{i=1}^n g_i^2\right)^{1/2}}$,

- $V_1^{(n)} := U^{1/n}V^{(n)}$,

- $W^{(n)} := n^{\frac{1}{p}-\frac{1}{2}}\dfrac{\left(\sum_{i=1}^n Z_i^2\right)^{1/2}}{\left(\sum_{i=1}^n |Z_i|^p\right)^{1/p}}$,

and the sequences $\mathbf{V} := (V^{(n)})_{n\in\mathbb{N}}$, $\mathbf{V}_1 := (V_1^{(n)})_{n\in\mathbb{N}}$ and $\mathbf{W} := (W^{(n)})_{n\in\mathbb{N}}$. Using these definitions we notice that $n^{\frac{1}{p}-\frac{1}{2}}\|P_{E^{(n)}}X^{(n)}\|_2$ has the same distribution as $V_1^{(n)}W^{(n)}$.

For technical reasons, we will have to split the LDPs for the sequences $\mathbf{V}$, and $\mathbf{V}_1$ into the three different cases

- $\lambda \in (0,1)$,
- $\lambda = 0$,
- $\lambda = 1$,

where, $\lambda = \lim\limits_{n\to\infty} \frac{k_n}{n}$. Note that the LDPs for the random sequences $\mathbf{V}$ and $\mathbf{V}_1$ will be unaffected by the choice of the value p. The latter enters only in the LDP for the random sequence $\mathbf{W}$ and causes the different large deviation behavior of $\|P_E X\|$ displayed Theorem 1.1 and Theorem 1.2.

4.1. **LDP for the random sequence V.** We prepare the proof for the LDP for the sequence $\mathbf{V}$ with the following general lemma. For this we recall that a random variable is gamma distributed with parameters $a > 0$ and $b > 0$ provided that its Lebesgue density is

$$x \mapsto \frac{b^a}{\Gamma(a)}\, x^{a-1} e^{-bx}, \qquad x \geq 0.$$

Similarly, a random variable is beta distributed with parameters $a > 0$ and $b > 0$ if its Lebesgue density is given by

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

where $B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$ is Euler's beta function.

Lemma 4.1. *Let $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ be positive sequences such that the limits $\lim\limits_{n\to\infty} \frac{a_n}{n} = a \in [0,\infty)$ and $\lim\limits_{n\to\infty} \frac{b_n}{n} = b \in [0,\infty)$ exist are not equal to zero at the same time. For each $n \in \mathbb{N}$ let $X^{(n)}$ be a beta distributed random variable with parameters a_n and b_n. Then the sequence $\mathbf{X} = (X^{(n)})_{n\in\mathbb{N}}$ satisfies a LDP with speed n and with the rate function $\mathcal{I}_\mathbf{X}$ given by*

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

Proof. For $t \in \mathbb{R}$ let us write

$$\Lambda(t) := \lim_{n\to\infty} \frac{1}{n}\log \mathbb{E}e^{tnX^{(n)}}$$

$$= \lim_{n\to\infty} \frac{1}{n}\log \frac{1}{B(a_n,b_n)} \int_0^1 x^{a_n-1}(1-x)^{b_n-1}\, e^{tnx}\, \mathrm{d}x$$

$$= \lim_{n\to\infty} \frac{1}{n}\log \frac{1}{B(a_n,b_n)} + \lim_{n\to\infty} \frac{1}{n}\log \int_0^1 e^{n(tx + \frac{a_n-1}{n}\log x + \frac{b_n-1}{n}\log(1-x))}\, \mathrm{d}x.$$

According to Laplace's principle, the second term is equal to

$$\Psi^*(t) = \sup_{x\in(0,1)}\, [xt - \Psi(x)] \qquad \text{with} \qquad \Psi(x) = -a\log x - b\log(1-x),$$

where here and below we interpret the formal expression $0\log 0$ as 0. Using the explicit form of Ψ and the fact that a and b cannot both be equal to zero one easily computes that, for all $t \in \mathbb{R}$,

$$\Psi^*(t) = \frac{1}{2}\Big(t - (a+b) + \sqrt{(a+b-t)^2 - 4at}\Big)$$

(4)
$$+ a\log\Big(\frac{t - (a+b) + \sqrt{(a+b-t)^2 - 4at}}{2t}\Big)$$

$$+ b\log\Big(\frac{t + a + b + \sqrt{(a+b-t)^2 - 4at}}{2t}\Big).$$

To treat the first term, which involves the beta function, we distinguish different cases. First assume that $a_n \to \infty$ and $b_n \to \infty$. Then Stirling's formula yields that asymptotically, as $n \to \infty$, $B(a_n, b_n)$ behaves like

$$\sqrt{2\pi}\, \frac{a_n^{a_n-\frac{1}{2}} b_n^{b_n-\frac{1}{2}}}{(a_n+b_n)^{a_n+b_n-\frac{1}{2}}}\,.$$

Thus,

$$\lim_{n\to\infty} \frac{1}{n} \log \frac{1}{B(a_n, b_n)}$$
$$= -\lim_{n\to\infty} \Big[\frac{\log\sqrt{2\pi}}{n} + \frac{a_n-\frac{1}{2}}{n}\Big(\log n + \log \frac{a_n}{n}\Big)$$
$$+ \frac{b_n-\frac{1}{2}}{n}\Big(\log n + \log \frac{b_n}{n}\Big) - \frac{a_n+b_n-\frac{1}{2}}{n}\Big(\log n + \log \frac{a_n+b_n}{n}\Big) \Big]$$
$$= -a\log a - b\log b + (a+b)\log(a+b)\,.$$

If otherwise a_n stays bounded and $b_n \to \infty$ then $B(a_n, b_n)$ behaves like $\Gamma(a_n)b_n^{-a_n}$, which implies that in this case

$$\lim_{n\to\infty} \frac{1}{n} \log \frac{1}{B(a_n, b_n)} = -\lim_{n\to\infty} \frac{1}{n} \log B(a_n, b_n) = 0$$

and the same is true if $a_n \to \infty$ and b_n is bounded. Putting things together yields, for $y \in (0,1)$ and any (permitted) choice of a_n and b_n, the limiting relation

$$\Lambda(t) = \lim_{n\to\infty} \frac{1}{n} \log \mathbb{E} e^{tnX^{(n)}} = c_{a,b} + \Psi^*(t)$$

with the constant $c_{a,b}$ given by

$$c_{a,b} := -a\log a - b\log b + (a+b)\log(a+b)\,.$$

A glance at (4) shows that the function Λ is clearly lower semi-continuous and even infinitely differentiable on $\mathbb{R}$. In addition, $D_\Lambda = \mathbb{R}$. We can thus apply the Gärtner-Ellis theorem in Proposition 2.11 to deduce that the sequence $\mathbf{X}$ satisfies a LDP with speed n and rate function

$$\Lambda^*(y) = \sup_{t\in\mathbb{R}}[ty - \Lambda(t)] = \sup_{t\in\mathbb{R}}[ty - \Psi^*(t) - c_{a,b}]$$
$$= \sup_{t\in\mathbb{R}}[ty - \Psi^*(t)] - c_{a,b} = \Psi^{**}(y) - c_{a,b} = \Psi(y) - c_{a,b}\,,$$

where in the last step we used the fact that the Legendre-Fenchel transform is an involution (at least in our set-up, where Ψ is differentiable on $\mathbb{R}$). If $a > 0$ and $b > 0$ we have, for $y \in (0,1)$,

$$\Lambda^*(y) = -a\log y - b\log(1-y) + a\log a + b\log b - (a+b)\log(a+b)$$
$$= -a\log\frac{y}{a} - b\log\frac{1-y}{b} - (a+b)\log(a+b)\,.$$

If $a = 0$ and $b > 0$ we obtain, for $y \in [0,1)$, the representation

$$\Lambda(y)^* = -b\log(1-y)$$

and if $a > 0$ and $b = 0$ we get, for $y \in (0,1]$,

$$\Lambda(y)^* = -a\log y\,.$$

This completes the proof.   □

We shall now apply the previous result to deduce a LDP for the sequence $\mathbf{V}$.

Lemma 4.2. *For each $n \in \mathbb{N}$ let $k_n \in \mathbb{N}$ with $k_n \in \{1, \ldots, n-1\}$ and assume that $\lim_{n \to \infty} \frac{k_n}{n} = \lambda \in [0,1]$. Then the sequence $\mathbf{V}$ satisfies a LDP with speed n and rate function*

$$\mathcal{I}_{\mathbf{V}}^{(\lambda)}(y) = \begin{cases} -\frac{\lambda}{2} \log \frac{y^2}{\lambda} - \frac{1-\lambda}{2} \log \frac{1-y^2}{1-\lambda} & : y \in (0,1) \ and \ \lambda \in (0,1) \\ -\frac{1}{2} \log y^2 & : y \in (0,1] \ and \ \lambda = 1 \\ -\frac{1}{2} \log(1-y^2) & : y \in [0,1) \ and \ \lambda = 0 \\ +\infty & : otherwise. \end{cases}$$

Proof. Let us recall that

$$X^{(n)} := (V^{(n)})^2 = \frac{\sum_{i=1}^{k_n} g_i^2}{\sum_{i=1}^{k_n} g_i^2 + \sum_{i=k_n+1}^{n} g_i^2} = \frac{\Gamma_1}{\Gamma_1 + \Gamma_2}$$

with independent random variables Γ_1 and Γ_2, where Γ_1 is gamma distributed with parameters $k_n/2$ and $1/2$, while Γ_2 is gamma distributed with parameters $(n-k_n)/2$ and $1/2$. It is well known that in this situation the ratio $\frac{\Gamma_1}{\Gamma_1+\Gamma_2}$ is beta distributed with parameters $k_n/2$ and $(n-k_n)/2$. We can now apply Lemma 4.1 with $a_n = k_n/2$, $b_n = (n-k_n)/2$, $a = \lambda/2$ and $b = (1-\lambda)/2$ (note that $a+b = 1/2$) to deduce a LDP for the sequence $((V^{(n)})^2)_{n \in \mathbb{N}}$ with speed n and rate function $\mathcal{I}_{\mathbf{X}}^{(\lambda)}$ given as follows. If $\lambda \in (0,1)$ and $y \in (0,1)$, then

$$\begin{aligned} \mathcal{I}_{\mathbf{X}}^{(\lambda)}(y) &= -\frac{\lambda}{2} \log \frac{2y}{\lambda} - \frac{1-\lambda}{2} \log \frac{2(1-y)}{1-\lambda} - \frac{1}{2} \log \frac{1}{2} \\ &= -\frac{\lambda}{2} \log \frac{y}{\lambda} - \frac{1-\lambda}{2} \log \frac{1-y}{1-\lambda} - \frac{\lambda}{2} \log 2 - \frac{1-\lambda}{2} \log 2 + \frac{1}{2} \log 2 \\ &= -\frac{\lambda}{2} \log \frac{y}{\lambda} - \frac{1-\lambda}{2} \log \frac{1-y}{1-\lambda}, \end{aligned}$$

if $\lambda = 1$ and $y \in (0,1]$, then

$$\mathcal{I}_{\mathbf{X}}^{(\lambda)}(y) = \frac{1}{2} \log y$$

and if $\lambda = 0$ and $y \in [0,1)$, then

$$\mathcal{I}_{\mathbf{X}}^{(\lambda)}(y) = \frac{1}{2} \log(1-y)$$

and $\mathcal{I}_{\mathbf{X}}^{(\lambda)}(y) = +\infty$ in all other cases. Finally, noting that $V^{(n)} = F(X_n)$ with the continuous function $F(y) = \sqrt{y}$, $y > 0$, we apply the contraction principle in Proposition 2.9 to deduce the LDP with speed n and rate function

$$\mathcal{I}_{\mathbf{V}}^{(\lambda)}(y) = (\mathcal{I}_{\mathbf{X}}^{(\lambda)} \circ F^{-1})(y) = \mathcal{I}_{\mathbf{X}}^{(\lambda)}(y^2)$$

for the sequence $\mathbf{V}$. $\square$

4.2. LDP for the random sequence $\mathbf{V}_1$. In this subsection we will prove LDPs for the sequence $\mathbf{V}_1$, again in the three different cases $\lambda \in (0,1)$, $\lambda = 1$ and $\lambda = 0$. As we will see below, the radial part is in fact negligible and so the rate functions for $\mathbf{V}_1$ coincide with the corresponding ones for $\mathbf{V}$ obtained in Subsection 4.1.

In all three cases, we will use the following result proved in [13, Lemma 3.3], which is also a direct consequence of Lemma 4.1.

Lemma 4.3. *The sequence* $\mathbf{U} = (U^{1/n})_{n\in\mathbb{N}}$ *satisfies an LDP with speed* n *and rate function*

$$\mathcal{I}_{\mathbf{U}}(y) := \begin{cases} -\log(y) & : y \in (0,1] \\ +\infty & : otherwise. \end{cases}$$

Proof. Note that, for each $t \in [0,1]$,

$$\mathbb{P}(U^{1/n} \le t) = P(U \le t^n) = t^n,$$

which shows that $U^{1/n}$ is beta distributed with parameters n and 1. We can thus apply Lemma 4.1 with $a_n = n$, $b_n = 1$, $a = 1$ and $b = 0$ to deduce the LDP for the sequence $\mathbf{U}$. $\square$

We start with $\lambda \in (0,1)$.

Corollary 4.4. *For each* $n \in \mathbb{N}$ *let* $k_n \in \mathbb{N}$ *with* $k_n \in \{1, \ldots, n-1\}$ *and assume that*

$$\lim_{n\to\infty} \frac{k_n}{n} = \lambda \in (0,1).$$

Then the sequence $\mathbf{V}_1$ *satisfies an LDP with speed* n *and rate function*

$$\mathcal{I}_{\mathbf{V}_1}^{(\lambda)}(y) := \begin{cases} \frac{\lambda}{2}\log\left(\frac{\lambda}{y^2}\right) + \frac{1-\lambda}{2}\log\left(\frac{1-\lambda}{1-y^2}\right) & : y \in (0,1) \\ +\infty & : otherwise. \end{cases}$$

Proof. By Proposition 2.6, Lemma 4.2 and Lemma 4.3 the sequence of random vectors $\left((U^{1/n}, V^{(n)})\right)_{n\in\mathbb{N}}$ satisfies an LDP with speed n and rate function

$$\mathcal{I}(x_1, x_2) = \begin{cases} -\log(x_1) + \frac{\lambda}{2}\log\left(\frac{\lambda}{x_2^2}\right) + \frac{1-\lambda}{2}\log\left(\frac{1-\lambda}{1-x_2^2}\right) & : x_1 \in (0,1] \text{ and } x_2 \in (0,1) \\ +\infty & : \text{otherwise}, \end{cases}$$

$(x_1, x_2) \in \mathbb{R}^2$. Defining $F(x_1, x_2) := x_1 x_2$ and applying the contraction principle (see Proposition 2.9), we deduce that $\mathbf{V}_1 = \left(F(U^{1/n}, V^{(n)})\right)_{n\in\mathbb{N}}$ satisfies an LDP with speed n and rate function

$$\mathcal{I}_{\mathbf{V}_1}^{(\lambda)}(y) = \inf_{\substack{(x_1,x_2)\in\mathbb{R}^2 \\ F(x_1,x_2)=y}} \mathcal{I}(x_1, x_2), \qquad y \in \mathbb{R}.$$

If $y \in (0,1)$, then

$$\mathcal{I}_{\mathbf{V}_1}^{(\lambda)}(y)$$

$$= \inf_{x_1 x_2 = y}\left[-\log(x_1) + \frac{\lambda}{2}\log\left(\frac{\lambda}{x_2^2}\right) + \frac{1-\lambda}{2}\log\left(\frac{1-\lambda}{1-x_2^2}\right)\right]$$

$$= \inf_{x_1 x_2 = y}\left[-\frac{\lambda}{2}\log(x_1^2) - \frac{1-\lambda}{2}\log(x_1^2) + \frac{\lambda}{2}\log\left(\frac{\lambda}{x_2^2}\right) + \frac{1-\lambda}{2}\log\left(\frac{1-\lambda}{1-x_2^2}\right)\right]$$

$$= \inf_{x_1 x_2 = y}\left[\frac{\lambda}{2}\log\left(\frac{\lambda}{(x_1 x_2)^2}\right) + \frac{1-\lambda}{2}\log\left(\frac{1-\lambda}{x_1^2 - (x_1 x_2)^2}\right)\right]$$

$$= \inf_{x_1 x_2 = y}\left[\frac{\lambda}{2}\log\left(\frac{\lambda}{y^2}\right) + \frac{1-\lambda}{2}\log\left(\frac{1-\lambda}{x_1^2 - y^2}\right)\right]$$

$$= \frac{\lambda}{2}\log\left(\frac{\lambda}{y^2}\right) + \frac{1-\lambda}{2}\log\left(\frac{1-\lambda}{1-y^2}\right),$$

since the infimum is attained at $x_1 = 1$ and $x_2 = y$. If $y \notin (0,1)$, for every $(x_1, x_2) \in \mathbb{R}^2$ such that $x_1 x_2 = y$, we have $\mathcal{I}(x_1, x_2) = +\infty$. $\square$

In the same way we obtained Corollary 4.4, we also treat the LDP for $\mathbf{V}_1$ if $\lambda = 0$.

Corollary 4.5. *For each $n \in \mathbb{N}$ let $k_n \in \mathbb{N}$ with $k_n \in \{1, \ldots, n-1\}$ be such that*

$$\lim_{n \to \infty} \frac{k_n}{n} = 0 \,.$$

Then, the sequence of random variables $\mathbf{V}_1$ satisfies an LDP with speed n and rate function

$$\mathcal{I}_{\mathbf{V}_1}^{(0)}(y) := \begin{cases} -\frac{1}{2}\log(1-y^2) & : y \in [0,1) \\ +\infty & : otherwise \,. \end{cases}$$

Proof. By Proposition 2.6, Lemma 4.2 and Lemma 4.3 the sequence of random vectors $\left((U^{1/n}, V^{(n)})\right)_{n \in \mathbb{N}}$ satisfies an LDP with speed n and rate function

$$\mathcal{I}(x_1, x_2) = \begin{cases} -\log(x_1) - \frac{1}{2}\log(1-x_2^2) & : x_1 \in (0,1] \text{ and } x_2 \in [0,1) \\ +\infty & : \text{otherwise} \,, \end{cases}$$

$(x_1, x_2) \in \mathbb{R}^2$. Defining $F(x_1, x_2) := x_1 x_2$ and applying the contraction principle (see Proposition 2.9), we deduce that $\mathbf{V}_1 = \left(F(U^{1/n}, V^{(n)})\right)_{n \in \mathbb{N}}$ satisfies an LDP with speed n and rate function

$$\mathcal{I}_{\mathbf{V}_1}^{(0)}(y) = \inf_{\substack{(x_1,x_2) \in \mathbb{R}^2 \\ F(x_1,x_2)=y}} \mathcal{I}(x_1, x_2)\,, \qquad y \in \mathbb{R} \,.$$

If $y \in [0, 1)$, then

$$\begin{aligned} \mathcal{I}_{\mathbf{V}_1}^{(0)}(y) &= \inf_{x_1 x_2 = y} \left[-\log(x_1) - \frac{1}{2}\log(1-x_2^2) \right] \\ &= \inf_{x_1 x_2 = y} \left[-\frac{1}{2}\log(x_1^2) - \frac{1}{2}\log(1-x_2^2) \right] \\ &= \inf_{x_1 x_2 = y} \left[-\frac{1}{2}\log\left(x_1^2 - x_1^2 x_2^2\right) \right] \\ &= \inf_{x_1 x_2 = y} \left[-\frac{1}{2}\log\left(x_1^2 - y^2\right) \right] \\ &= -\frac{1}{2}\log(1-y^2)\,, \end{aligned}$$

since the infimum is attained at $x_1 = 1$ and $x_2 = y$. If $y \notin [0,1)$, for every $(x_1, x_2) \in \mathbb{R}^2$ such that $x_1 x_2 = y$, we have $\mathcal{I}(x_1, x_2) = +\infty$. $\qquad \square$

Finally, we consider the case $\lambda = 1$.

Corollary 4.6. *For each $n \in \mathbb{N}$ let $k_n \in \mathbb{N}$ with $k_n \in \{1, \ldots, n-1\}$ be such that*

$$\lim_{n \to \infty} \frac{k_n}{n} = 1 \,.$$

Then, the sequence of random variables $\mathbf{V}_1$ satisfies an LDP with speed n and rate function

$$\mathcal{I}_{\mathbf{V}_1}^{(1)}(y) := \begin{cases} -\frac{1}{2}\log(y^2) & : y \in (0,1] \\ +\infty & : otherwise \,. \end{cases}$$

Proof. By Proposition 2.6, Lemma 4.2 and Lemma 4.3 the sequence of random vectors $\big((U^{1/n}, V^{(n)})\big)_{n\in\mathbb{N}}$ satisfies an LDP with speed n and rate function

$$\mathcal{I}(x_1, x_2) = \begin{cases} -\log(x_1) - \frac{1}{2}\log(x_2^2) & : x_1 \in (0,1] \text{ and } x_2 \in (0,1] \\ +\infty & : \text{otherwise}, \end{cases}$$

$(x_1, x_2) \in \mathbb{R}^2$. Defining $F(x_1, x_2) := x_1 x_2$ and applying the contraction principle (see Proposition 2.9), we deduce that $\mathbf{V}_1 = \big(F(U^{1/n}, V^{(n)})\big)_{n\in\mathbb{N}}$ satisfies an LDP with speed n and rate function

$$\mathcal{I}^{(1)}_{\mathbf{V}_1}(y) = \inf_{\substack{(x_1, x_2) \in \mathbb{R}^2 \\ F(x_1, x_2) = y}} I(x_1, x_2), \qquad y \in \mathbb{R}.$$

If $y \in (0,1]$, then

$$\begin{aligned} \mathcal{I}^{(1)}_{\mathbf{V}_1}(y) &= \inf_{x_1 x_2 = y} \left[-\log(x_1) - \frac{1}{2}\log(x_2^2) \right] \\ &= \inf_{x_1 x_2 = y} \left[-\frac{1}{2}\log(x_1^2) - \frac{1}{2}\log(x_2^2) \right] \\ &= \inf_{x_1 x_2 = y} \left[-\frac{1}{2}\log(x_1^2 x_2^2) \right] \\ &= \inf_{x_1 x_2 = y} \left[-\frac{1}{2}\log(y^2) \right] \\ &= -\frac{1}{2}\log(y^2). \end{aligned}$$

If $y \notin (0,1]$, for every $(x_1, x_2) \in \mathbb{R}^2$ such that $x_1 x_2 = y$, we have $\mathcal{I}(x_1, x_2) = +\infty$. $\qquad\square$

4.3. **LDP for the random sequence W.** In this subsection we prove LDPs for the sequence $\mathbf{W} = (W^{(n)})_{n\in\mathbb{N}}$. We will only consider the case $p \in [2, \infty)$ (the special situation in which $p = \infty$ is treated directly in the proof of Theorem 1.1 in Section 5), where the result follows from Cramér's theorem and the contraction principle.

Lemma 4.7. *Let $p \in [2, \infty)$. Then $\mathbf{W}$ satisfies an LDP with speed n and rate function*

$$\mathcal{I}^{(p)}_{\mathbf{W}}(y) := \begin{cases} \inf_{\substack{x_1 \geq 0, x_2 > 0 \\ x_1^{1/2} x_2^{-1/p} = y}} \mathcal{I}^*(x_1, x_2) & : y \geq 0 \\ +\infty & : y < 0, \end{cases}$$

where $\mathcal{I}^*$ is the Legendre-Fenchel transform of

$$\mathcal{I}(t_1, t_2) := \log\left(\int_{\mathbb{R}} e^{t_1 x^2 + t_2 |x|^p} \frac{e^{-\frac{|x|^p}{p}}}{2p^{1/p}\Gamma(1 + \frac{1}{p})} \, dx \right)$$

with effective domain $\mathbb{R} \times (-\infty, 1/p)$ if $p > 2$ and $\{(t_1, t_2) \in \mathbb{R}^2 : t_1 + t_2 < \frac{1}{2}\}$ if $p = 2$.

Proof. We set

$$S^{(n)} := \frac{1}{n} \sum_{i=1}^{n} (Z_i^2, |Z_i|^p), \qquad n \in \mathbb{N}.$$

Let $t = (t_1, t_2) \in \mathbb{R}^2$ and define

$$\mathcal{I}(t_1, t_2) := \log\left(\mathbb{E}\, e^{\langle t, (Z_1^2, |Z_1|^p)\rangle}\right)$$

$$= \log\left(\int_{\mathbb{R}} e^{t_1 x^2 + t_2 |x|^p}\, \frac{e^{-\frac{|x|^p}{p}}}{2p^{1/p}\Gamma(1 + \frac{1}{p})}\, \mathrm{d}x\right)$$

$$= \log\left(\int_0^\infty \frac{e^{\frac{1}{p}(pt_1 x^2 - (1 - pt_2)x^p)}}{p^{1/p}\Gamma(1 + \frac{1}{p})}\, \mathrm{d}x\right),$$

which is finite in $\mathbb{R} \times \left(-\infty, \frac{1}{p}\right)$ if $p > 2$ and if $t_1 + t_2 < \frac{1}{2}$ for $p = 2$. Since $(0,0)$ is in the interior of the effective domain of $\mathcal{I}$, by Cramér's theorem (see Proposition 2.5), $(S^{(n)})_{n \in \mathbb{N}}$ satisfies an LDP with speed n and rate function $\mathcal{I}^*$. Notice that the effective domain of $\mathcal{I}^*$ is contained in $[0, \infty) \times [0, \infty)$. Moreover, $(W^{(n)})_{n \in \mathbb{N}} = (F(S^{(n)}))_{n \in \mathbb{N}}$, with $F : \mathbb{R}^2 \to \mathbb{R}$ being the function given by

$$F(x_1, x_2) = x_1^{1/2} x_2^{-1/p}.$$

Note that this function is continuous on $[0, \infty) \times (0, \infty)$. Hence, by the contraction principle (see Proposition 2.9), $\mathbf{W}$ satisfies an LDP with speed n and rate function

$$\mathcal{I}_{\mathbf{W}}^{(p)}(y) = \inf_{\substack{x_1 \geq 0, x_2 > 0 \\ x_1^{1/2} x_2^{-1/p} = y}} \mathcal{I}^*(x_1, x_2),$$

if $y \geq 0$ and $\mathcal{I}_{\mathbf{W}}^{(p)}(y) = +\infty$ if $y < 0$, because $F^{-1}(\{y\}) = \emptyset$. $\qquad\square$

Remark 4.8. *Note that if $p = 2$, the random variables $W^{(n)}$, $n \geq 1$ are constantly equal to 1. This means that for any $A \in \mathscr{B}(\mathbb{R})$,*

$$\lim_{n \to \infty} \frac{\log\left(\mathbb{P}(W^{(n)} \in A)\right)}{n} = \begin{cases} 0 & : 1 \in A \\ -\infty & : 1 \notin A. \end{cases}$$

Therefore,

$$\mathcal{I}_{\mathbf{W}}^{(2)}(y) = \begin{cases} 0 & : y = 1 \\ +\infty & : y \neq 1. \end{cases}$$

5. Proof of the main results

After these preparations, we can now present the proofs of our main results, Theorem 1.1 and Theorem 1.2.

5.1. **Proof of Theorem 1.1.** First, let $p \in [2, \infty)$. According to Theorem 3.1, for each $n \in \mathbb{N}$, the random variable $n^{\frac{1}{p} - \frac{1}{2}} \|P_{E^{(n)}} X^{(n)}\|_2$ has the same distribution as $V_1^{(n)} W^{(n)}$. By Proposition 2.6, if for $\lambda \in [0, 1]$, $\mathcal{I}_{\mathbf{V}_1}^{(\lambda)}$ and $\mathcal{I}_{\mathbf{W}}^{(p)}$ are the rate functions defined in Corollaries 4.4, 4.5 and 4.6, and Lemma 4.7 the sequence of random vector $\left((V_1^{(n)}, W^{(n)})\right)_{n \in \mathbb{N}}$ satisfies an LDP with speed n and rate function

$$\mathcal{I}_{\mathbf{V}_1}^{(\lambda)}(x_1) + \mathcal{I}_{\mathbf{W}}^{(p)}(x_2)$$
$$= \begin{cases} \frac{\lambda}{2}\log\left(\frac{\lambda}{x_1^2}\right) + \frac{1-\lambda}{2}\log\left(\frac{1-\lambda}{1-x_1^2}\right) + \mathcal{I}_{\mathbf{W}}^{(p)}(x_2) & : x_1 \in (0, 1) \text{ and } x_2 \in D_{\mathcal{I}_{\mathbf{W}}^{(p)}} \\ +\infty & : \text{otherwise}, \end{cases}$$

$(x_1, x_2) \in \mathbb{R}^2$. By the contraction principle (see Proposition 2.9) applied to the function $F : \mathbb{R}^2 \to \mathbb{R}$, $F(x_1, x_2) = x_1 x_2$, we conclude that the sequence of random variables $(F(V_1^{(n)}, W^{(n)}))_{n \in \mathbb{N}}$ satisfies an LDP with speed n and rate function

$$\mathcal{I}_{\|P_E X\|}(y) = \inf_{x_1 x_2 = y} \left[\mathcal{I}_{\mathbf{V}_1}(x_1) + \mathcal{I}_{\mathbf{W}}^{(p)}(x_2) \right], \qquad y \in \mathbb{R}.$$

If $y < 0$, then, for any $x_1, x_2 \in \mathbb{R}$ such that $x_1 x_2 = y$, either x_1 or x_2 is negative and so $\mathcal{I}_{\|P_E X\|}(y) = +\infty$. If $y = 0$, then, for any $x_1, x_2 \in \mathbb{R}$ with $x_1 x_2 = y$, either $x_1 = 0$, or $x_2 = 0$. If $x_1 = 0$ then $\mathcal{I}_{\mathbf{V}_1}^{(\lambda)}(0) = +\infty$ if $\lambda \neq 0$. If $\lambda = 0$ then $\mathcal{I}_{\mathbf{V}_1}^{(0)}(0) = 0$. Thus, if $\lambda \neq 0$, we see that

$$\mathcal{I}_{\|P_E X\|}(0) = \inf_{x_1 \in (0,1)} \left[\mathcal{I}_{\mathbf{V}_1}(x_1) + \mathcal{I}_{\mathbf{W}}^{(p)}(0) \right].$$

Since $\inf_{x_1 \in (0,1)} \mathcal{I}_{\mathbf{V}_1}(x_1)$ is attained when $x_1 = \sqrt{\lambda}$ and $\mathcal{I}_{\mathbf{V}_1}^{(\lambda)}(\sqrt{\lambda}) = 0$, we obtain $\mathcal{I}_{\|P_E X\|}(0) = \mathcal{I}_{\mathbf{W}}^{(p)}(0)$. If $\lambda = 0$, then

$$\mathcal{I}_{\|P_E X\|}(0) = \min\{\mathcal{I}_{\mathbf{W}}^{(p)}(0), \inf_{x \geq 0} \mathcal{I}_{\mathbf{W}}^{(p)}(x)\} = \inf_{x \geq 0} \mathcal{I}_{\mathbf{W}}^{(p)}(x).$$

If $y > 0$, then, for any $x_1, x_2 \in \mathbb{R}$ such that $x_1 x_2 = y$, we can write $x_1 = \frac{y}{x_2}$. Since if $0 < x_2 < y$, we have that $x_1 > 1$, in such a case $\mathcal{I}_{\mathbf{V}_1}^{(\lambda)}(x_1) = +\infty$. If $x_2 < 0$ then $\mathcal{I}_{\mathbf{W}}^{(p)}(x_2) = +\infty$. Thus,

$$\mathcal{I}_{\|P_E X\|}(y) = \inf_{x_2 \geq y} \left[\frac{\lambda}{2 x_2^2} \log \left(\frac{\lambda}{y^2} \right) + \frac{1 - \lambda}{2} \log \left(\frac{1 - \lambda}{1 - \frac{y^2}{x_2^2}} \right) + \mathcal{I}_{\mathbf{W}}^{(p)}(x_2) \right],$$

which is the function in the statement of the theorem.

Finally, we consider the case $p = +\infty$ and notice that $\frac{1}{\sqrt{n}} \|P_E X\|_2$ has the same distribution as the product $(\overline{W}^{(n)})^{1/2} V^{(n)}$ with $\overline{W}^{(n)} := \frac{1}{n} \sum_{i=1}^{n} X_i^2$, where $X = (X_1, \ldots, X_n)$ is a random vector whose entries are independent and uniformly distributed on $[-1, 1]$. By Cramér's theorem (see Proposition 2.5) it follows that $\overline{W}^{(n)}$ satisfies an LDP with speed n and rate function $\mathcal{I}_\infty^*$, the Legendre-Fenchel transform of $\mathcal{I}_\infty(t) = \log \left(2 \int_0^1 e^{t x^2} \, \mathrm{d}t \right)$. Thus, according to Lemma 4.2 and the contraction principle (see Proposition 2.9), the sequence $(\frac{1}{\sqrt{n}} \|P_E X\|_2)_{n \in \mathbb{N}}$ satisfies an LDP with speed n and rate function

$$\mathcal{I}_{\|P_E X\|}(y) = \inf_{\substack{x_1, x_2 \in \mathbb{R} \\ x_1^{1/2} x_2 = y}} \left[\mathcal{I}_\infty(x_1) + \mathcal{I}_{\mathbf{V}}^{(\lambda)}(x_2) \right]$$

$$= \begin{cases} \inf_{x \geq y} \left[\frac{\lambda}{2} \log \left(\frac{\lambda x^2}{y^2} \right) + \frac{1 - \lambda}{2} \log \left(\frac{1 - \lambda}{1 - y^2 x^{-2}} \right) \right] & : y > 0 \\ \inf_{x > 0} \mathcal{I}_\infty(x) & : y = 0 \text{ and } \lambda = 0 \\ +\infty & : \text{otherwise}. \end{cases}$$

This complete the proof of the theorem. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

Remark 5.1. *In the special case $p = 2$ we find that*

$$\mathcal{I}_{\|P_E X\|}(y) = \inf_{x_1 x_2 = y} \left[\mathcal{I}_{\mathbf{V}_1}^{(\lambda)}(x_1) + \mathcal{I}_{\mathbf{W}}^{(2)}(x_2) \right] = \mathcal{I}_{\mathbf{V}_1}^{(\lambda)}(y) + \mathcal{I}_{\mathbf{W}}^{(2)}(1) = \mathcal{I}_{\mathbf{V}_1}(y),$$

since $\mathcal{I}_{\mathbf{W}}^{(2)}(1) = 0$, see Remark 4.8. This proves relation (1) *in the introduction.*

5.2. **Proof of Theorem 1.2.** The proof of Theorem 1.2 requires different tools. In particular, it relies on a large deviation result for sums of so-called stretched exponential random variables taken from a paper of Gantert, Ramanan and Rembart [15]. We start by computing the variance of a p-generalized Gaussian random variable.

Lemma 5.2. *Let Z be a p-generalized Gaussian random variable for some $p \in [1, \infty)$. Then*

$$\mathbb{E}Z^2 = \frac{p^{\frac{p}{2}}}{3} \frac{\Gamma\left(1 + \frac{3}{p}\right)}{\Gamma\left(1 + \frac{1}{p}\right)}.$$

Proof. Recalling the definition of the density f_p of Z from Proposition 2.13 and applying the change of variables $u = x^p/p$, we see that

$$\mathbb{E}Z^2 = \int_{-\infty}^{\infty} x^2 \, f_p(x) \, \mathrm{d}x = \frac{1}{p^{1-\frac{2}{p}} \Gamma\left(1 + \frac{1}{p}\right)} \int_0^{\infty} u^{\frac{3}{p}-1} \, e^{-u} \, \mathrm{d}u = \frac{p^{\frac{p}{2}}}{3} \frac{\Gamma\left(1 + \frac{3}{p}\right)}{\Gamma\left(1 + \frac{1}{p}\right)}$$

and the proof is complete. $\qquad\square$

The next lemma provides bounds for the tails of the random variable Z^2. A function $f : (0, \infty) \to \mathbb{R}$ is said to be slowly varying (at infinity) provided that $\lim\limits_{t\to\infty} \frac{f(at)}{f(t)} = 1$ for any $a > 0$.

Lemma 5.3. *For $p \in [1, 2)$ let Z be a p-generalized Gaussian random variable and for $t > 0$ define functions*

$$b(t) := \frac{1}{p} + \frac{p-1}{2} t^{-\frac{p}{2}} \log t \,, \qquad c_1(t) := \frac{t^{\frac{p}{2}}}{t^{\frac{p}{2}} + 1} \qquad \text{and} \qquad c_2 = c_2(t) := 2.$$

These functions are slowly varying and, for all $t > 0$, one has that

$$c_1(t) \, e^{-b(t)\, t^{\frac{p}{2}}} \leq \mathbb{P}(Z^2 \geq t) \leq c_2 \, e^{-b(t)\, t^{\frac{p}{2}}} \,.$$

Proof. Let us first check that b, c_1 and c_2 are slowly varying. For c_2 this is trivial, while for b and c_1 we have that, for all $a > 0$,

$$\lim_{t\to\infty} \frac{b(at)}{b(t)} = \lim_{t\to\infty} \frac{\frac{1}{p} + \frac{p-1}{2}(at)^{-\frac{p}{2}} \log(at)}{\frac{1}{p} + \frac{p-1}{2}t^{-\frac{p}{2}} \log t} = 1 \,,$$

$$\lim_{t\to\infty} \frac{c_1(at)}{c_1(t)} = \lim_{t\to\infty} \frac{(at)^{\frac{p}{2}}}{(at)^{\frac{p}{2}} + 1} \frac{t^{\frac{p}{2}} + 1}{t^{\frac{p}{2}}} = 1 \,.$$

It is well known and easily shown that, for $t > 0$,

$$\frac{t}{t^p + 1} \, e^{-t^p/p} \leq \int_t^{\infty} e^{-s^p/p} \, \mathrm{d}s \leq \frac{1}{t^{p-1}} \, e^{-t^p/p} \,.$$

This readily implies the upper bound

$$\mathbb{P}(Z^2 \geq t) \leq \frac{2}{t^{\frac{p-1}{2}}} \, e^{-t^{\frac{p}{2}}/p} = 2 \, e^{-t^{\frac{p}{2}}\left(\frac{1}{p} + \frac{p-1}{2}t^{-\frac{p}{2}} \log t\right)} = c_2 \, e^{-b(t)\, t^{\frac{p}{2}}}$$

as well as the lower bound by writing

$$\mathbb{P}(Z^2 \geq t) \geq \frac{2\sqrt{t}}{t^{\frac{p}{2}} + 1} \, e^{-t^{\frac{p}{2}}/p} = \frac{t^{\frac{p}{2}}}{t^{\frac{p}{2}} + 1} \, e^{-t^{\frac{p}{2}}\left(\frac{1}{p} + \frac{p-1}{2}t^{-\frac{p}{2}} \log t\right)} = c_1(t) \, e^{-b(t)\, t^{\frac{p}{2}}} \,.$$

The argument is thus complete. $\qquad\square$

In the arguments that follow we also need the following two auxiliary LDP's.

Lemma 5.4.   (a) *Suppose that $k_n \to \infty$, as $n \to \infty$. Then the sequence $\mathbf{G} :=$ $(\frac{1}{k_n} \sum_{i=1}^{k_n} g_i^2)_{i \in \mathbb{N}}$ satisfies an LDP with speed k_n and rate function*

$$\mathcal{I}_{\mathbf{G}}(y) = \begin{cases} \frac{y-1}{2} - \frac{1}{2} \log y & : y > 0 \\ +\infty & : otherwise. \end{cases}$$

(b) *Let $p \in [1, \infty)$. Then the sequence $\mathbf{Z}_p := (\frac{1}{n} \sum_{i=1}^n |Z_i|^p)_{i \in \mathbb{N}}$ satisfies an LDP with speed n and rate function*

$$\mathcal{I}_{\mathbf{Z}_p}(y) = \begin{cases} \frac{1}{p} y - y^{\frac{1}{p+1}} \left(1 + \frac{1}{p}\right) & : y > 0 \\ +\infty & : otherwise. \end{cases}$$

Proof. Part (a) has already been verified in the proof of Lemma 4.2. To prove the statement in (b) we apply Cramér's theorem (Proposition 2.5). Indeed, the moment generating function of $|Z|^p$, where Z has a p-generalized Gaussian distribution, is given by

$$\mathcal{I}(y) = \int_{-\infty}^{\infty} e^{|x|^p y} f_p(x) \, \mathrm{d}x = \frac{1}{(1 - yp)^{\frac{1}{p}}}, \qquad y < \frac{1}{p}.$$

In particular, zero is an interior point of the effective domain of $\mathcal{I}$. As a consequence, $\mathbf{Z}_p$ satisfies an LDP with speed n and rate function given by the Legendre-Fenchel transform $\mathcal{I}^*$ of $\mathcal{I}$. The latter is given by

$$\mathcal{I}^*(y) = \sup_{x \in \mathbb{R}} [xy - \mathcal{I}(x)] = \frac{1}{p} y - y^{\frac{1}{p+1}} \left(1 + \frac{1}{p}\right)$$

if $y > 0$ and $\mathcal{I}^*(y) = +\infty$ otherwise. $\square$

After these preparations, we can now present the proof of Theorem 1.2. From now on we shall assume that we are dealing with a fixed parameter $p \in [1, 2)$.

Proof of Theorem 1.2. Applying [15, Theorem 1] (with $a_j(n) = \frac{1}{n}$, $s = s_1 = 1$) together with Lemma 5.3 implies that for $y \geq m$, with $m = \mathbb{E}Z^2$ from Lemma 5.2,

$$\lim_{n \to \infty} \frac{1}{b(n)n^{\frac{p}{2}}} \log \mathbb{P}\left(\frac{1}{n} \sum_{i=1}^n Z_i^2 \geq y\right) = -(y - m)^{\frac{p}{2}}.$$

Since the sequences $(b(n)n^{p/2})_{n \in \mathbb{N}}$ and $(\frac{1}{p} n^{p/2})_{n \in \mathbb{N}}$ are asymptotically equivalent, the pre-factor $1/(b(n)n^{p/2})$ can be replaced by $1/n^{p/2}$. Moreover, since the random variables Z_i^2 are non-negative, this result can be lifted to an LDP, see also [15, Remark 3.2]. Thus, the sequence of random variables $\mathbf{Z} := (\frac{1}{n} \sum_{i=1}^n Z_i^2)_{n \in \mathbb{N}}$ satisfies an LDP with speed $n^{p/2}$ and rate function

$$\mathcal{I}_{\mathbf{Z}}(y) = \begin{cases} \frac{1}{p} (y - m)^{\frac{p}{2}} & : y \geq m \\ +\infty & : \text{otherwise}. \end{cases}$$

Next, we apply the contraction principle to the function F given by $F : (0, \infty) \to (0, \infty)$, $F(x) = \sqrt{x}$. This yields an LDP for the sequence $\sqrt{\mathbf{Z}} := ((\frac{1}{n} \sum_{i=1}^n Z_i^2)^{\frac{1}{2}})_{n \in \mathbb{N}}$ with speed $n^{p/2}$ and rate function

$$\mathcal{I}_{\sqrt{\mathbf{Z}}}(y) = \begin{cases} \frac{1}{p} (y^2 - m)^{\frac{p}{2}} & : y \geq \sqrt{m} \\ +\infty & : \text{otherwise}. \end{cases}$$

In a next step, we apply Proposition 2.10 to the functions $F_n : \mathbb{R} \to \mathbb{R}$, $F_n(x) = \sqrt{\frac{k_n}{n}}\,x$ and $F : \mathbb{R} \to \mathbb{R}$, $F(x) = \sqrt{\lambda}\,x$, where $\lambda > 0$ by assumption (the technical condition in Proposition 2.10 is easily seen to be satisfied in this situation). This leads to an LDP for the sequence $\widetilde{\mathbf{Z}} := (\sqrt{\frac{k_n}{n}}(\frac{1}{n}\sum_{i=1}^{n} Z_i^2))_{n \in \mathbb{N}}$ with speed $n^{p/2}$ and rate function

$$(5) \qquad \mathcal{I}_{\widetilde{\mathbf{Z}}}(y) = \inf_{F(x)=y} \mathcal{I}_{\sqrt{\mathbf{Z}}}(x) = \begin{cases} \frac{1}{p}\left(\frac{y^2}{\lambda} - m\right)^{\frac{p}{2}} & : y \geq \sqrt{\lambda m} \\ +\infty & : \text{otherwise}, \end{cases}$$

which coincides with the function $\mathcal{I}_{\|\boldsymbol{P_E X}\|}(y)$ in the statement of Theorem 1.2.

In the remaining part of the proof we shall argue that the two random sequences $\widetilde{\mathbf{Z}}$ and $\|\boldsymbol{P_E X}\|$ are exponentially equivalent and thus satisfy the same LDP. For this, we observe that according to Theorem 3.1, for each $n \in \mathbb{N}$, the random variable $\|n^{\frac{1}{p}-\frac{1}{2}} P_{E^{(n)}} X^{(n)}\|$ has the same distribution as

$$U^{1/n} \frac{\left(\frac{1}{n}\sum\limits_{i=1}^{n} Z_i^2\right)^{1/2}}{\left(\frac{1}{n}\sum\limits_{i=1}^{n} |Z_i|^p\right)^{1/p}} \frac{\left(\frac{1}{k_n}\sum\limits_{i=1}^{k_n} g_i^2\right)^{1/2}}{\left(\frac{1}{n}\sum\limits_{i=1}^{n} g_i^2\right)^{1/2}} \sqrt{\frac{k_n}{n}}\,.$$

Now, fix $\delta, \varepsilon > 0$ and note that

$$\mathbb{P}\left(\left| \sqrt{\frac{k_n}{n}}\left(\frac{1}{n}\sum_{i=1}^{n} Z_i^2\right)^{1/2} - U^{1/n} \frac{\left(\frac{1}{n}\sum\limits_{i=1}^{n} Z_i^2\right)^{1/2}}{\left(\frac{1}{n}\sum\limits_{i=1}^{n} |Z_i|^p\right)^{1/p}} \frac{\left(\frac{1}{k_n}\sum\limits_{i=1}^{k_n} g_i^2\right)^{1/2}}{\left(\frac{1}{n}\sum\limits_{i=1}^{n} g_i^2\right)^{1/2}} \sqrt{\frac{k_n}{n}} \right| > \delta\right)$$

$$\leq \mathbb{P}\left(\sqrt{\frac{k_n}{n}}\left(\frac{1}{n}\sum_{i=1}^{n} Z_i^2\right)^{1/2} > \frac{\delta}{\varepsilon}\right)$$

$$+ \mathbb{P}\left(1 - U^{1/n} \frac{1}{\left(\frac{1}{n}\sum\limits_{i=1}^{n} |Z_i|^p\right)^{1/p}} \frac{\left(\frac{1}{k_n}\sum\limits_{i=1}^{k_n} g_i^2\right)^{1/2}}{\left(\frac{1}{n}\sum\limits_{i=1}^{n} g_i^2\right)^{1/2}} > \varepsilon\right)$$

$$+ \mathbb{P}\left(1 - U^{1/n} \frac{1}{\left(\frac{1}{n}\sum\limits_{i=1}^{n} |Z_i|^p\right)^{1/p}} \frac{\left(\frac{1}{k_n}\sum\limits_{i=1}^{k_n} g_i^2\right)^{1/2}}{\left(\frac{1}{n}\sum\limits_{i=1}^{n} g_i^2\right)^{1/2}} < -\varepsilon\right)$$

$$=: T_1 + T_2 + T_3\,.$$

We further estimate T_2 by

$$T_2 \leq \mathbb{P}(U^{1/n} < (1-\varepsilon)^{1/4}) + \mathbb{P}\left(\frac{1}{k_n}\sum_{i=1}^{k_n} g_i^2 < (1-\varepsilon)^{1/2}\right)$$

$$+ \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} |Z_i|^p > (1-\varepsilon)^{-p/4}\right) + \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} g_i^2 > (1-\varepsilon)^{-1/2}\right)$$

$$=: T_{2,1} + T_{2,2} + T_{2,3} + T_{2,4}\,.$$

According to Lemma 4.3 and Lemma 5.4, the terms $T_{2,1}$, $T_{2,3}$ and $T_{2,4}$ decay exponentially with speed n. Indeed, this follows from the fact that the rate functions of

the corresponding LDP's do not vanish at $(1-\varepsilon)^{1/4}$, $(1-\varepsilon)^{-p/4}$ and $(1-\varepsilon)^{-1/2}$, respectively. In addition and again by Lemma 5.4, the term $T_{2,2}$ decays exponentially with speed k_n and again the rate function in the corresponding LDP does vanish at $(1-\varepsilon)^{-p/4}$.

Similarly, for T_3 we have the bound

$$
T_3 \leq \mathbb{P}(U^{1/n} > (1+\varepsilon)^{1/4}) + \mathbb{P}\Big(\frac{1}{k_n}\sum_{i=1}^{k_n} g_i^2 > (1+\varepsilon)^{1/2}\Big)
$$

$$
+ \mathbb{P}\Big(\frac{1}{n}\sum_{i=1}^{n} |Z_i|^p < (1+\varepsilon)^{-p/4}\Big) + \mathbb{P}\Big(\frac{1}{n}\sum_{i=1}^{n} g_i^2 < (1+\varepsilon)^{-1/2}\Big)
$$

$$
= \mathbb{P}\Big(\frac{1}{k_n}\sum_{i=1}^{k_n} g_i^2 > (1+\varepsilon)^{1/2}\Big)
$$

$$
+ \mathbb{P}\Big(\frac{1}{n}\sum_{i=1}^{n} |Z_i|^p < (1+\varepsilon)^{-p/4}\Big) + \mathbb{P}\Big(\frac{1}{n}\sum_{i=1}^{n} g_i^2 < (1+\varepsilon)^{-1/2}\Big).
$$

As for T_2 discussed above, these terms decay exponentially with speed n and since $\frac{k_n}{n} \to \lambda > 0$, as $n \to \infty$, we conclude that

$$
\limsup_{n\to\infty} \frac{1}{n^{p/2}} \log T_2 + \limsup_{n\to\infty} \frac{1}{n^{p/2}} \log T_3 = -\infty.
$$

Hence,

$$
\limsup_{n\to\infty} \frac{1}{n^{p/2}} \log \mathbb{P}\Big(\Big|\sqrt{\frac{k_n}{n}}\Big(\frac{1}{n}\sum_{i=1}^{n} Z_i^2\Big)^{1/2} - \|n^{\frac{1}{p}-\frac{1}{2}} P_{E^{(n)}} X^{(n)}\|\Big| > \delta\Big)
$$

$$
\leq \limsup_{n\to\infty} \frac{1}{n^{p/2}} \log T_1.
$$

Sending $\varepsilon \to 0$, the above LDP for the sequence $\widetilde{\mathbf{Z}}$ (recall (5)) shows that this limit exists and is equal to $-\infty$. We have thus proved that the random sequences $\widetilde{\mathbf{Z}}$ and $\|\boldsymbol{P_E X}\|$ are exponentially equivalent. So, by Proposition 2.7 they satisfy the same LDP. This completes the proof of Theorem 1.2. □

6. APPENDIX

Let us present the proof of Proposition 2.6.

Proof. For every $m \in \mathbb{N}$, any $\delta > 0$ and any $x \in \mathbb{R}^m$, let us denote by

$$
B_\infty(x,\delta) := \{y \in \mathbb{R}^m : \|x-y\|_\infty < \delta\},
$$

the cube in $\mathbb{R}^m$ centered at x with side length 2δ. Let $d_1, d_2 \in \mathbb{N}$.

Lower bound. Let $A \in \mathscr{L}(\mathbb{R}^{d_1} \times \mathbb{R}^{d_2})$ with non-empty interior. Let $z = (x,y) \in A^\circ$ and $\delta > 0$ such that $B_\infty(z,\delta) = B_\infty(x,\delta) \times B_\infty(y,\delta) \subset A$. From the independence of $X^{(n)}$ and $Y^{(n)}$, we conclude that

$$
\frac{1}{s(n)} \log \Big(\mathbb{P}\big(Z^{(n)} \in A^\circ\big)\Big) \geq \frac{1}{s(n)} \log \Big(\mathbb{P}\big(Z^{(n)} \in B_\infty(x,\delta) \times B_\infty(y,\delta)\big)\Big)
$$

$$
= \frac{1}{s(n)} \log \Big(\mathbb{P}\big(X^{(n)} \in B_\infty(x,\delta)\big)\Big)
$$

$$+ \frac{1}{s(n)} \log \left(\mathbb{P}\big(Y^{(n)} \in B_\infty(y, \delta)\big) \right).$$

Consequently, for every $z = (x, y) \in A^\circ$ and $\delta > 0$ such that $B_\infty(z, \delta) \subset A^\circ$,

$$\liminf_{n \to \infty} \frac{1}{s(n)} \log \left(\mathbb{P}\big(Z^{(n)} \in A^\circ\big) \right) \geq - \inf_{y_1 \in B_\infty(x, \delta)} \mathcal{I}_{\mathbf{X}}(y_1) - \inf_{y_2 \in B_\infty(y, \delta)} \mathcal{I}_{\mathbf{Y}}(y_2)$$

and, since this inequality is true for every such x and $\delta > 0$,

$$\liminf_{n \to \infty} \frac{1}{s(n)} \log \left(\mathbb{P}\big(Z^{(n)} \in A^\circ\big) \right) \geq - \inf_{(x,y) \in A^\circ} \left[\mathcal{I}_{\mathbf{X}}(x) + \mathcal{I}_{\mathbf{Y}}(y) \right].$$

Upper bound. Since we are considering the σ-algebra $\mathscr{L}(\mathbb{R}^{d_1} \times \mathbb{R}^{d_2})$, by Proposition 2.2 it is enough to prove the upper bound for subsets $A \subset \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ such that $\bar{A}$ is compact together with the exponential tightness for $\mathbf{Z}$. In such a case, for every cover of $\bar{A}$ by subsets of the form $B_\infty(z_i, \delta_i)$, $z_i = (x_i, y_i) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ we can extract a finite number of sets such that $\bar{A} \subset \bigcup_{i=1}^{N} B_\infty(z_i, \delta_i)$ and for every such finite cover we have

$$\frac{1}{s(n)} \log \left(\mathbb{P}\big(Z^{(n)} \in \bar{A}\big) \right) \leq \frac{1}{s(n)} \log \left(\sum_{i=1}^{N} \mathbb{P}\big(Z^{(n)} \in B_\infty(x_i, \delta_i) \times B_\infty(y_i, \delta_i)\big) \right)$$

$$\leq \frac{1}{s(n)} \log \left(N \max_{1 \leq i \leq N} \mathbb{P}\big(Z^{(n)} \in B_\infty(x_i, \delta_i) \times B_\infty(y_i, \delta_i)\big) \right)$$

$$= \frac{\log N}{s(n)} + \max_{1 \leq i \leq N} \left\{ \frac{1}{s(n)} \log \left(\mathbb{P}\big(X^{(n)} \in B_\infty(x_i, \delta_i)\big) \right) \right.$$

$$\left. + \frac{1}{s(n)} \log \left(\mathbb{P}\big(Y^{(n)} \in B_\infty(y_i, \delta_i)\big) \right) \right\}.$$

Now, taking the $\limsup_{n \to \infty}$, since N is a fixed number depending on the cover of $\bar{A}$ we extracted, the first term tends to 0 and, since the $\limsup_{n \to \infty}$ of the maximum equals the maximum of the $\limsup_{n \to \infty}$, we have

$$\limsup_{n \to \infty} \frac{1}{s(n)} \log \left(\mathbb{P}\big(Z^{(n)} \in \bar{A}\big) \right)$$

$$\leq \max_{1 \leq i \leq N} \left\{ - \inf_{x \in B_\infty(x_i, \delta_i)} \mathcal{I}_{\mathbf{X}}(x) - \inf_{y \in B_\infty(y_i, \delta_i)} \mathcal{I}_{\mathbf{Y}}(y) \right\}$$

$$- \min_{1 \leq i \leq N} \left\{ \inf_{x \in B_\infty(x_i, \delta_i)} \mathcal{I}_{\mathbf{X}}(x) + \inf_{y \in B_\infty(y_i, \delta_i)} \mathcal{I}_{\mathbf{Y}}(y) \right\}.$$

Since this is true for any cover of $\bar{A}$ and every finite cover we extract from it, we obtain

$$\limsup_{n \to \infty} \frac{1}{s(n)} \log \left(\mathbb{P}\big(Z^{(n)} \in \bar{A}\big) \right) \leq - \inf_{(x,y) \in \bar{A}} \left[\mathcal{I}_{\mathbf{X}}(x) + \mathcal{I}_{\mathbf{Y}}(y) \right],$$

which proves the upper bound for subsets of $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ such that $\bar{A}$ is compact.

Exponential tightness. To show the exponential tightness of $\mathbf{Z}$, let $\alpha > 0$ be any positive number. Since $\mathbf{X}$ and $\mathbf{Y}$ are assumed to satisfy a (full) LDP, by Proposition 2.2, $\mathbf{X}$ and $\mathbf{Y}$ are exponentially tight and, thus, there exist compact sets $K_{1,\alpha} \subset \mathbb{R}^{d_1}$ and $K_{2,\alpha} \subset \mathbb{R}^{d_2}$ such that

$$\limsup_{n \to \infty} \frac{1}{s(n)} \log \mathbb{P}(X^{(n)} \notin K_{1,\alpha}) < -\frac{\alpha}{2}$$

and

$$\limsup_{n \to \infty} \frac{1}{s(n)} \log \mathbb{P}(Y^{(n)} \notin K_{2,\alpha}) < -\frac{\alpha}{2}.$$

Thus, for any $\alpha > 0$, taking $K_\alpha := K_{1,\alpha} \times K_{2,\alpha} \subset \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, we obtain that

$$\limsup_{n \to \infty} \frac{1}{s(n)} \log \mathbb{P}(Z^{(n)} \notin K_\alpha) < -\alpha.$$

Consequently, $\mathbf{Z}$ is exponentially tight and the proof is thus complete. $\qquad\square$

Acknowledgement

References

[1] M. Anttila, K. Ball, and I. Perissinaki. The central limit problem for convex bodies. *Trans. Am. Math. Soc.*, 355(12):4723–4735, 2003.

[2] F. Barthe, F. Gamboa, L. Lozada-Chang, and A. Rouault. Generalized Dirichlet distributions on the ball and moments. *ALEA Lat. Am. J. Probab. Math. Stat.*, 7:319–340, 2010.

[3] F. Barthe, O. Guédon, S. Mendelson, and A. Naor. A probabilistic approach to the geometry of the ℓ_p^n-ball. *Ann. Probab.*, 33(2):480–513, 2005.

[4] E. Bingham and H. Mannila. Random projection in dimensionality reduction: Applications to image and text data. In *Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '01, pages 245–250, New York, NY, USA, 2001. ACM.

[5] S.G. Bobkov and A. Koldobsky. On the central limit property of convex bodies. In *Geometric Aspects of Functional Analysis*, volume 1807 of *Lecture Notes in Math.*, pages 44–52. Springer, Berlin, 2003.

[6] A.A. Borovkov. *Probability Theory*. Universitext. Springer, London, 2013. Translated from the 2009 Russian fifth edition by O. B. Borovkova and P. S. Ruzankin, Edited by K. A. Borovkov.

[7] S. Brazitikos, A. Giannopoulos, P. Valettas, and B.-H. Vritsiou. *Geometry of Iso tropic Convex Bodies*, volume 196 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2014.

[8] U. Brehm, P. Hinow, H. Vogt, and J. Voigt. Moment inequalities and central limit properties of isotropic convex bodies. *Math. Z.*, 240(1):37–51, 2002.

[9] N.R. Chaganty and J. Sethuraman. Strong large deviation and local limit theorems. *Ann. Probab.*, 21(3):1671–1690, 1993.

[10] A. Dembo and O. Zeitouni. *Large Deviations. Techniques and Applications*, volume 38 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2010. Corrected reprint of the second (1998) edition.

[11] F. den Hollander. *Large Deviations*, volume 14 of *Fields Institute Monographs*. American Mathematical Society, Providence, RI, 2000.

[12] P. Diaconis and D. Freedman. Asymptotics of graphical projection pursuit. *Ann. Stat.*, 12:793–815, 1984.

[13] N. Gantert, S.S. Kim, and K. Ramanan. Large deviations for random projections of ℓ^p balls. *ArXiv: 1512.04988*, December 2015.

[14] N. Gantert, S.S. Kim, and K. Ramanan. *Cramér's Theorem is Atypical*, pages 253–270. Springer International Publishing, Cham, 2016.

[15] N. Gantert, K. Ramanan, and F. Rembart. Large deviations for weighted sums of stretched exponential random variables. *Electron. Commun. Probab.*, 19:no. 41, 14, 2014.

[16] O. Guédon. Concentration phenomena in high dimensional geometry. In *Journées MAS 2012*, volume 44 of *ESAIM Proc.*, pages 47–60. EDP Sci., Les Ulis, 2014.

[17] O. Guédon, P. Nayar, and T. Tkocz. Concentration inequalities and geometry of convex bodies. In *Analytical and probabilistic methods in the geometry of convex bodies*, volume 2 of *IMPAN Lect. Notes*, pages 9–86. Polish Acad. Sci. Inst. Math., Warsaw, 2014.

[18] O. Kallenberg. *Foundations of Modern Probability*. Probability and its Applications. Springer-Verlag, New York, second edition, 2002.

[19] S.S. Kim and K. Ramanan. A Sanov-type theorem for empirical measures associated with the surface and cone measures on ℓ^p spheres. *ArXiv: 1509.05442*, September 2015.

[20] B. Klartag. A central limit theorem for convex sets. *Invent. Math.*, 168(1):91–131, 2007.

[21] B. Klartag. Power-law estimates for the central limit theorem for convex sets. *J. Funct. Anal.*, 245(1):284–310, 2007.

[22] O.-A. Maillard and R. Munos. Linear regression with random projections. *Journal of Machine Learning Research*, 13:2735–2772, 2012.

[23] E.S. Meckes. Approximation of projections of random vectors. *J. Theoret. Probab.*, 25(2):333–352, 2012.

[24] E.S. Meckes. Projections of probability distributions: a measure-theoretic Dvoretzky theorem. In *Geometric Aspects of Functional Analysis*, volume 2050 of *Lecture Notes in Math.*, pages 317–326. Springer, Heidelberg, 2012.

[25] E.S. Meckes and M.W. Meckes. The central limit problem for random vectors with symmetries. *J. Theoret. Probab.*, 20(4):697–720, 2007.

[26] M.W. Meckes. Gaussian marginals of convex bodies with symmetries. *Beiträge Algebra Geom.*, 50(1):101–118, 2009.

[27] E. Milman. On Gaussian marginals of uniformly convex bodies. *J. Theoret. Probab.*, 22(1):256–278, 2009.

[28] A. Naor. The surface measure and cone measure on the sphere of ℓ_p^n. *Trans. Amer. Math. Soc.*, 359(3):1045–1079 (electronic), 2007.

[29] G. Paouris. Concentration of mass and central limit properties of isotropic convex bodies. *Proc. Amer. Math. Soc.*, 133(2):565–575, 2005.

[30] S.T. Rachev and L. Rüschendorf. Approximate independence of distributions on spheres and their stability properties. *Ann. Probab.*, 19(3):1311–1337, 1991.

[31] G. Schechtman and M. Schmuckenschläger. Another remark on the volume of the intersection of two L_p^n balls. In *Geometric aspects of functional analysis (1989–90)*, volume 1469 of *Lecture Notes in Math.*, pages 174–178. Springer, Berlin, 1991.

[32] G. Schechtman and J. Zinn. On the volume of the intersection of two L_p^n balls. *Proc. Amer. Math. Soc.*, 110(1):217–224, 1990.

[33] G. Schechtman and J. Zinn. Concentration on the l_p^n ball. In *Geometric aspects of functional analysis*, volume 1745 of *Lecture Notes in Math.*, pages 245–256. Springer, Berlin, 2000.

[34] V.N. Sudakov. Typical distributions of linear functionals in finite-dimensional spaces of high dimension. *Dokl. Akad. Nauk SSSR*, 243(6):1402–1405, 1978.

Departamento de Matemáticas, Universidad de Zaragoza, Spain
E-mail address: alonsod@unizar.es

School of Mathematics & Physical Sciences, University of Hull, United Kingdom
E-mail address: j.prochno@hull.ac.uk

Faculty of Mathematics, Ruhr University Bochum, Germany
E-mail address: christoph.thaele@rub.de