LARGE DEVIATIONS FOR UNIFORM PROJECTIONS
OF p-RADIAL DISTRIBUTIONS ON ℓ_p^n-BALLS

BY

TOM KAUFMANN (Bochum), HOLGER SAMBALE (Bochum), and
CHRISTOPH THÄLE (Bochum)

Abstract. We consider products of uniform random variables from the Stiefel manifold of orthonormal k-frames in $\mathbb{R}^n$, $k \leqslant n$, and random vectors from the n-dimensional ℓ_p^n-ball $\mathbb{B}_p^n$ with certain p-radial distributions, $p \in [1, \infty)$. The distribution of this product geometrically corresponds to the projection of the p-radial distribution on $\mathbb{B}_p^n$ onto a random k-dimensional subspace. We derive large deviation principles (LDPs) on the space of probability measures on $\mathbb{R}^k$ for sequences of such projections.

2020 Mathematics Subject Classification: Primary 52A23; Secondary 60F10.

Key words and phrases: large deviation principle, ℓ_p^n-ball, random projection, Stiefel manifold.

1. INTRODUCTION

The study of high-dimensional convexity goes back to studying infinite-dimensional normed spaces via their local structures, such as their unit balls, but has since become of substantial interest in its own right. The concentration phenomena exhibited by convex objects in high dimensions, analyzed in the language of probability, are of great use in applications such as compressed sensing, information theory and approximation theory (see [3, 5, 7, 8, 18]). Analogues of many well known limit results from probability have been found in high-dimensional convexity, such as the central limit theorem (CLT) of Klartag [16, 17], and recently, starting with the work of Gantert, Kim, and Ramanan [6], large deviations began to be considered with increasing interest as well. For the sake of brevity, we refer the reader to the classic literature on large deviations theory for more details on basic definitions and results [4].

Especially the n-dimensional ℓ_p^n-ball $\mathbb{B}_p^n := \{x \in \mathbb{R}^n : \|x\|_p \leqslant 1\}$, for $p \in [1, \infty)$ and $\|x\|_p := (\sum_{i=1}^n |x_i|^p)^{1/p}$, has been extensively considered in this regard (see, e.g., [1, 12, 11, 13, 10]) due to both its relevance in geometry and func-

tional analysis, and its accessibility via useful probabilistic representation results, which will be outlined below. For an overview of classical and more recent results about ℓ_p^n-balls, we refer to the survey article [19].

The setting of the present work is a generalization of the one initiated by Kabluchko and Prochno [10] and can be described as follows. For $k \leqslant n$, the Stiefel manifold $\mathbb{V}_{n,k}$ is the set of all orthonormal k-frames in $\mathbb{R}^n$, i.e., the set of all k-tuples of orthonormal vectors $v_1, \ldots, v_k$ in $\mathbb{R}^n$. Arranging these vectors into a $k \times n$ matrix V with rows $v_1^T, \ldots, v_k^T$, we have the identification

$$\mathbb{V}_{n,k} = \{V \in \mathbb{R}^{k \times n} \colon VV^T = I_k\},$$

where I_k denotes the $k \times k$ identity matrix. We equip $\mathbb{V}_{n,k}$ with the uniform distribution (i.e., the invariant Haar probability measure) $\mu_{n,k}$, writing $V_{n,k}$ for the corresponding random variable. Recall that $V_{n,k}$ is characterized by the following invariance property: for any orthogonal matrices $O \in \mathbb{R}^{k \times k}$ and $O' \in \mathbb{R}^{n \times n}$, $OV_{n,k}O'$ has the same distribution as $V_{n,k}$.

In particular, for random vectors $X^{(n)}$ taking values in $\mathbb{R}^n$, we may regard $V \in \mathbb{V}_{n,k}$ as a linear map $V \colon \mathbb{R}^n \to \mathbb{R}^k$ and study the distribution of the vectors $VX^{(n)} \in \mathbb{R}^k$, which we denote by

$$\mu_{VX^{(n)}}(A) := \mathbb{P}(VX^{(n)} \in A)$$

for any Borel set $A \subseteq \mathbb{R}^k$. In addition, we may also choose $V_{n,k} \in \mathbb{V}_{n,k}$ at random according to the uniform distribution $\mu_{n,k}$ on $\mathbb{V}_{n,k}$. In this case, the distribution of $V_{n,k}X^{(n)}$, which we denote by

$$(1.1) \qquad\qquad \mu_{V_{n,k}X^{(n)}}(A) := \mathbb{P}(V_{n,k}X^{(n)} \in A),$$

is a random probability measure on $\mathbb{R}^k$, that is, a random variable taking values in the space $\mathcal{M}_1(\mathbb{R}^k)$ of probability measures on $\mathbb{R}^k$, which we equip with the topology of weak convergence. This can be geometrically interpreted as the projection of the distribution of the random vector $X^{(n)}$ onto a uniform random k-dimensional subspace.

We are interested in large deviation principles (LDPs) for the random probability measures $\mu_{V_{n,k}X^{(n)}}$, where $X^{(n)} \in \mathbb{B}_p^n$ with a distribution taken from the class of p-radial distributions introduced by Barthe, Guédon, Mendelson and Naor [2]. Kabluchko and Prochno [10] gave very general LDPs for random matrices in the orthogonal group and the Stiefel manifold, and showed an LDP for k-dimensional projections of the special case of the *uniform distribution* on $\mathbb{B}_p^n$ as an application. Based on [10], our work substantially extends the set of projected distributions for which such an LDP is shown from the uniform distribution to the aforementioned class of p-radial distributions. In particular, we will see that the large deviation behaviour observed by Kabluchko and Prochno [10] is universal for a large class of

probability measures on $\mathbb{B}_p^n$. Moreover, we shall describe geometrically motivated distributions on $\mathbb{B}_p^n$ for which the LDP needs a suitable modification we also provide. We should also delineate the present work from the results shown by Kim and Ramanan [15, Theorems 2.4 & 2.6], who have shown, among other results, LDPs for uniform random projections of uniform random vectors in $\mathbb{B}_p^n$ onto k-dimensional subspaces. By the same arguments as put forth in [10], we note that while the settings are quite similar, the key difference is in the object of study, which in [15] is the projection point itself, hence yielding an LDP on $\mathbb{R}^k$, whereas in both [10] and this work it is the projected distribution on $\mathbb{R}^k$, thus the main result yields an LDP on the space $\mathcal{M}_1(\mathbb{R}^k)$ of probability measures on $\mathbb{R}^k$.

In the next section we briefly list the notation and background material we will need to formulate our theorems, which are presented in Section 3. Section 4 will then contain their respective proofs.

2. PRELIMINARIES

Let us first define the objects and distributions needed for the main results. We write $\mathcal{B}(\mathbb{R}^n)$ for the Borel σ-algebra, $\langle \cdot, \cdot \rangle_2$ for the Euclidean scalar product, and $\mathrm{vol}_n(\cdot)$ for the Lebesgue measure on $\mathbb{R}^n$. As already mentioned in the introduction, for $p \in [1,\infty)$, $n \in \mathbb{N}$, and $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ we denote by $\|x\|_p := (\sum_{i=1}^n |x_i|^p)^{1/p}$ the ℓ_p^n-norm of x and by $\mathbb{B}_p^n := \{x \in \mathbb{R}^n : \|x\|_p \leqslant 1\}$ and $\mathbb{S}_p^{n-1} := \{x \in \mathbb{R}^n : \|x\|_p = 1\}$ unit ℓ_p^n-ball and unit ℓ_p^n-sphere, respectively. These definitions extend to $p \in (0, 1)$, even if $\| \cdot \|_p$ is no longer a norm but a quasi-norm in this case. We define the uniform distribution on $\mathbb{B}_p^n$ and the cone probability measure on $\mathbb{S}_p^{n-1}$ as

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

Following [2], for a sequence $(\mathbf{W}_n)_{n \in \mathbb{N}}$ of Borel probability measures on $[0, \infty)$ we define the sequence of distributions

$$(2.1) \qquad \mathbf{P}_{n,p,\mathbf{W}_n} := \mathbf{W}_n(\{0\})\mathbf{C}_{n,p} + \Psi_n \mathbf{U}_{n,p}$$

on $\mathbb{B}_p^n$, where $\Psi_n(x) = \psi(\|x\|_p)$, $x \in \mathbb{B}_p^n$, is the p-radial density given by

$$\psi_n(s) = \frac{1}{p^{n/p}\Gamma\left(\frac{n}{p}+1\right)} \frac{1}{(1-s^p)^{\frac{n}{p}+1}} \left[\int_{(0,\infty)} w^{\frac{n}{p}} e^{-\frac{1}{p}\left(\frac{s^p}{1-s^p}\right)w} \, \mathbf{W}_n(\mathrm{d}w) \right]$$

with $0 \leqslant s \leqslant 1$. One can think of $\mathbf{W}_n$ as indicating how probability mass is distributed p-radially within $\mathbb{B}_p^n$. The motivation behind this class of distributions is twofold. First, they encompass many relevant distributions on $\mathbb{B}_p^n$. For instance, choosing $\mathbf{W}_n \equiv \delta_0$ to be the Dirac measure at 0, we have $\mathbf{P}_{n,p,\mathbf{W}_n} \equiv \mathbf{C}_{n,p}$ and

for $\mathbf{W}_n \equiv \mathrm{Exp}(1)$, we have $\mathbf{P}_{n,p,\mathbf{W}_n} \equiv \mathbf{U}_{n,p}$. For $m \in \mathbb{N}$ choosing $\mathbf{W}_n = \gamma(m/p, 1/p)$, i.e., a gamma distribution with shape m/p and rate $1/p$, it can be shown that $\mathbf{P}_{n,p,\mathbf{W}_n}$ corresponds to the projection of $\mathbf{C}_{n+m,p}$ onto its first n co-ordinates. An analogous correspondence holds for $\mathbf{W}_n = \gamma(1 + m/p, 1/p)$ and the projection of $\mathbf{U}_{n+m,p}$ onto its first n coordinates [2]. The second reason we consider the class of distributions $\mathbf{P}_{n,p,\mathbf{W}_n}$ is a useful probabilistic representation result. For any $p \in (0, \infty)$ we say a real-valued random variable X has a *p-generalized Gaussian distribution*, denoted as $X \sim \mathbf{N}_p$, if its distribution has Lebesgue density

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

For $X \sim \mathbf{N}_p$ and $r > 0$ the rth absolute moment of X is given by

$$(2.2) \qquad \mathbb{E}[|X|^r] = p^{r/p}\Gamma\left(\frac{1+r}{p}\right)\Gamma\left(\frac{1}{p}\right)^{-1}.$$

In particular, $\mathbb{E}[|X|^r] < \infty$ for all $r > 0$. Using this p-generalized Gaussian distribution, the following results from [2, Theorem 3] gives a way to represent a random vector $X^{(n)} \sim \mathbf{P}_{n,p,\mathbf{W}_n}$ in $\mathbb{B}_p^n$ as a vector with i.i.d. p-generalized Gaussian coordinates, normalized by its norm and via $\mathbf{W}_n$.

PROPOSITION 2.1. *Let $n \in \mathbb{N}$ and $p \in (0, \infty)$. Let $Z^{(n)} = (Z_1, \dots, Z_n)$ be a random vector, where $Z_1, \dots, Z_n$ are i.i.d. with $Z_i \sim \mathbf{N}_p$, and W_n a random variable with distribution $\mathbf{W}_n$ on $[0, \infty)$, independent of $Z^{(n)}$. Then the random vector*

$$\frac{Z^{(n)}}{(\|Z^{(n)}\|_p^p + W_n)^{1/p}}$$

has distribution $\mathbf{P}_{n,p,\mathbf{W}_n}$ as in (2.1).

The final result presented in this section is one of the main results from [10], namely Theorem D therein. It provides an LDP for random projections of product measures, which we will use in conjunction with the representation from Proposition 2.1 to prove our theorems. In what follows we shall write $\mathcal{D}(X)$ for the distribution of a random variable X. Moreover, let

$$\mathcal{R}_2^{k\times\infty} := \{A = (A_{ij})_{i,j=1}^{k,\infty} \colon (A_{ij})_{j\in\mathbb{N}} \in \ell_2,\ i = 1, \dots, k\}$$

be the set of all matrices $A \in \mathbb{R}^{k\times\infty}$ with square-summable rows. For $A \in \mathbb{R}^{k\times\infty}$ we denote by $\|AA^T\|_{\mathrm{op}}$ the operator norm of the matrix $AA^T \in \mathbb{R}^{k\times k}$, where the condition $A \in \mathcal{R}_2^{k\times\infty}$ guarantees that AA^T is well-defined.

PROPOSITION 2.2. *Fix $k \in \mathbb{N}$. For each $n \in \mathbb{N}$ let $Z^{(n)} = (Z_1, \dots, Z_n)$ be an n-dimensional random vector, where $Z_1, Z_2, \dots$ are i.i.d. non-Gaussian random variables with symmetric distribution and finite moments of all orders. Let $\sigma^2 := \mathbb{E}[Z_1^2] > 0$ be the variance of Z_1. Then the sequence of random probability*

measures $\mu_{V_{n,k}Z^{(n)}}$, $n \geqslant k$, as in (1.1) satisfies an LDP on $\mathcal{M}_1(\mathbb{R}^k)$ with speed n and good rate function $I \colon \mathcal{M}_1(\mathbb{R}^k) \to [0, \infty]$ given by

$$I(\nu) = -\tfrac{1}{2} \log \det(I_k - AA^T)$$

if ν admits a representation of the form

$$\nu = \mathcal{D}\Big(\sum_{j=1}^{\infty} A_{\bullet,j} Z_j + \sigma(I_k - AA^T)^{1/2} N_k \Big)$$

for some matrix $A \in \mathcal{R}_2^{k \times \infty}$ with columns $A_{\bullet,1}, A_{\bullet,2}, \dots$ such that $\|AA^T\|_{\mathrm{op}} < 1$, where N_k is a k-dimensional standard Gaussian random vector independent of $Z_1, Z_2, \dots$. If ν does not admit a representation of this form, then $I(\nu) = \infty$.

Note that the specific distribution of the Z_i has a rather subtle influence on the rate function of the LDP via the matrix A used in the representation of a given measure $\nu \in \mathcal{M}_1(\mathbb{R}^k)$. As a side remark, note that in [10, Theorem D], the case of $\sigma^2 = 0$ actually has to be excluded. We have amended the result accordingly.

3. MAIN RESULTS

We are now in a position to present the first of our main results for the projections of p-radial distributions $\mathbf{P}_{n,p,\mathbf{W}_n}$ on ℓ_p^n-balls. In what follows we shall write $\overset{\mathcal{D}}{=}$ for equality in distribution.

THEOREM 3.1. *Fix $p \in (0, \infty)$, $p \neq 2$, and $k \in \mathbb{N}$. Moreover, let $(\mathbf{W}_n)_{n \in \mathbb{N}}$ be a sequence of Borel probability measures on $[0, \infty)$ and $(W_n)_{n \in \mathbb{N}}$ a sequence of random variables with $W_n \sim \mathbf{W}_n$, such that $W_n/n \to \alpha \in [0, \infty)$ in probability. Finally, let $X^{(n)}, Y^{(n)}$ be random vectors in $\mathbb{B}_p^n$ with $Y^{(n)} \sim \mathbf{P}_{n,p,\mathbf{W}_n}$ and $X^{(n)} \overset{\mathcal{D}}{=} n^{1/p} Y^{(n)}$. Then the sequence of random probability measures $\mu_{V_{n,k}X^{(n)}}$, $n \geqslant k$, as in (1.1) satisfies an LDP on $\mathcal{M}_1(\mathbb{R}^k)$ with speed n and good rate function $I \colon \mathcal{M}_1(\mathbb{R}^k) \to [0, \infty]$ given by*

$$I(\nu) = -\tfrac{1}{2} \log \det(I_k - AA^T)$$

if ν admits a representation of the form

$$\nu = \mathcal{D}\Big(\Big(\frac{1}{1+\alpha} \Big)^{1/p} \sum_{j=1}^{\infty} A_{\bullet,j} Z_j + \sigma_{p,\alpha}(I_k - AA^T)^{1/2} N_k \Big)$$

for some matrix $A \in \mathcal{R}_2^{k \times \infty}$ with columns $A_{\bullet,1}, A_{\bullet,2}, \dots$ such that $\|AA^T\|_{\mathrm{op}} < 1$, where $Z_1, Z_2, \dots$ are i.i.d. with $Z_i \sim \mathbf{N}_p$,

$$\sigma_{p,\alpha}^2 := \Big(\frac{p}{1+\alpha} \Big)^{2/p} \frac{\Gamma(3/p)}{\Gamma(1/p)},$$

and N_k is an independent k-dimensional standard Gaussian random vector. If ν does not admit a representation of this form, then $I(\nu) = \infty$.

As discussed earlier, choosing $\mathbf{W}_n \equiv \delta_0$ gives $\mathbf{P}_{n,p,\mathbf{W}_n} = \mathbf{C}_{n,p}$, and $\mathbf{W}_n \equiv$ Exp(1) yields $\mathbf{P}_{n,p,\mathbf{W}_n} = \mathbf{U}_{n,p}$; in both cases for $W_n \sim \mathbf{W}_n$ we have $W_n/n \to 0$ in probability and we recover [10, Theorem C]. Hence, we can see that $\mathbf{C}_{n,p}$ and $\mathbf{U}_{n,p}$ share the same LDP behaviour in high dimensions, which is in line with similar observations made for other functionals (see, e.g., [1, 12, 14]). Moreover, the result even implies a certain universality of the rate function, since despite the expected sensitivity of LDPs to the underlying distributions, the rate function is the same for all sequences $(\mathbf{W}_n)_{n \in \mathbb{N}}$ that share the same limiting behaviour.

Note that as in [10, Theorem C] (cf. also the non-Gaussianity assumption in Proposition 2.2 above), the case $p = 2$ has to be excluded in Theorem 3.1 since in this case, the matrix $A \in \mathcal{R}_2^{k \times \infty}$ in the representation of ν is not uniquely determined (and hence the rate function I is not well-defined). For further discussion, we refer to [10, the annotations following Theorems C and D, and Section 6].

Given the setting of Theorem 3.1, if we consider the case $W_n/n \to \infty$ in probability (formally corresponding to the choice $\alpha = \infty$), by the representation result in Proposition 2.1 one can see that this corresponds to each component of $X^{(n)}$ converging to 0 in probability, that is, we arrive at a trivial limit. To avoid this, we may choose a different scaling as in the following theorem.

THEOREM 3.2. *Fix $p \in (0, \infty)$, $p \neq 2$, and $k \in \mathbb{N}$. Moreover, let $(\mathbf{W}_n)_{n \in \mathbb{N}}$ be a sequence of Borel probability measures on $[0, \infty)$ and $(W_n)_{n \in \mathbb{N}}$ a sequence of random variables with $W_n \sim \mathbf{W}_n$ and $W_n/n^\kappa \to \beta \in (0, \infty)$ in probability for some $\kappa > 1$, and assume that the sequence of random variables $(W_n/n^\kappa)^{-2/p}$ is uniformly integrable. Finally, let $X^{(n)}, Y^{(n)}$ be random vectors in $\mathbb{B}_p^n$ with $Y^{(n)} \sim \mathbf{P}_{n,p,\mathbf{W}_n}$ and $X^{(n)} \overset{\mathcal{D}}{=} n^{\kappa/p} Y^{(n)}$. Then the sequence of random probability measures $\mu_{V_{n,k} X^{(n)}}$, $n \geqslant k$, as in (1.1) satisfies an LDP on $\mathcal{M}_1(\mathbb{R}^k)$ with speed n and good rate function $I \colon \mathcal{M}_1(\mathbb{R}^k) \to [0, \infty]$ given by*

$$I(\nu) = -\tfrac{1}{2} \log \det(I_k - AA^T)$$

if ν admits a representation of the form

$$\nu = \mathcal{D}\left(\left(\frac{1}{\beta} \right)^{1/p} \sum_{j=1}^{\infty} A_{\bullet,j} Z_j + \sigma_{p,\beta}(I_k - AA^T)^{1/2} N_k \right)$$

for some matrix $A \in \mathcal{R}_2^{k \times \infty}$ with columns $A_{\bullet,1}, A_{\bullet,2}, \dots$ such that $\|AA^T\|_{\mathrm{op}} < 1$, where $Z_1, Z_2, \dots$ are i.i.d. p-generalized Gaussian random variables,

$$\sigma_{p,\beta}^2 := \left(\frac{p}{\beta} \right)^{2/p} \frac{\Gamma(3/p)}{\Gamma(1/p)},$$

and N_k is an independent k-dimensional standard Gaussian random vector. If ν does not admit a representation of this form, then $I(\nu) = \infty$.

Theorem 3.2 includes an additional integrability condition for W_n, due to the fact that W_n is no longer of the same order as the p-norm of $Z^{(n)}$. Why this is needed specifically can be seen in its proof. Note that a helpful sufficient condition for the uniform integrability of $(W_n/n^\kappa)^{-2/p}$ is given by

$$(3.1) \qquad \sup_{n\in\mathbb{N}} \mathbb{E}\left[\left(\frac{n^\kappa}{W_n}\right)^{4/p}\right] \leqslant C$$

for some absolute constant $C > 0$. In particular, it can be applied to verify the uniform integrability for certain gamma distributions.

LEMMA 3.1. *Fix $p \in (0,\infty)$. Suppose for each $n \in \mathbb{N}$ that W_n follows a gamma distribution with shape a_n and rate $b > 0$, where $(a_n)_{n\in\mathbb{N}}$ is a positive sequence and $b \in (0,\infty)$. Assume that $a_n \geqslant m$ and $a_n/n^\kappa \geqslant C$ for all $n \in \mathbb{N}$ for some $m > 4/p$, $\kappa \in (0,\infty)$ and $C \in (0,\infty)$. Then*

$$\sup_{n\in\mathbb{N}} \mathbb{E}\left[\left(\frac{n^\kappa}{W_n}\right)^{4/p}\right] \leqslant M_{p,m}\left(\frac{b}{C}\right)^{4/p} < \infty,$$

where $M_{p,m}^{-1} = \prod_{i=0}^{A_p}\left(1 - \frac{4}{p(m+i)}\right)$ and $A_p = \lceil 4/p \rceil$ is the smallest integer greater than or equal to $4/p$.

The proof of this lemma is postponed to the end of this paper. As a concrete and geometrically motivated example we consider the distribution on $\mathbb{B}_p^n$ arising as the projection to the first n coordinates of the cone probability measure $\mathbf{C}_{n+m_n,p}$ on $\mathbb{B}_p^{n+m_n}$, where m_n is a positive sequence satisfying $m_n \geqslant m$ and $m_n/n^\kappa \geqslant C$ for all $n \in \mathbb{N}$ for some $m > 4$, $\kappa \in (0,\infty)$ and $C \in (0,\infty)$. As discussed above, this case corresponds to $\mathbf{P}_{n,p,\mathbf{W}_n}$ with $\mathbf{W}_n = \gamma(m_n/p, 1/p)$ and fits the assumptions of Lemma 3.1. The same holds for the projection of the uniform distribution $\mathbf{U}_{n+m_n,p}$ corresponding to $\mathbf{P}_{n,p,\mathbf{W}_n}$ with $\mathbf{W}_n = \gamma(1 + m_n/p, 1/p)$. In particular, Theorem 3.2 applies to these situations.

4. PROOFS

In this section we prove Theorems 3.1 and 3.2. The proofs will follow in the footsteps of the proof of [10, Theorem C], adapting and generalizing the arguments where necessary. We start off by formulating some probabilistic representations of the target quantities and show some auxiliary results.

Assume the set-up of Theorem 3.1 and for a fixed Stiefel matrix $V \in \mathbb{V}_{n,k}$ denote by $V_{\bullet,j}$, $j = 1,\ldots,n$, its columns. Then, by (1.1) and the representation results from Proposition 2.1 it follows that for any Borel set $F \in \mathcal{B}(\mathbb{R}^k)$,

$$(4.1) \qquad \mu_{VX^{(n)}}(F) = \mathbb{P}(VX^{(n)} \in F)$$

$$= \mathbb{P}\left(\sum_{j=1}^n n^{1/p} \frac{Z_j}{(\|Z^{(n)}\|_p^p + W_n)^{1/p}} V_{\bullet,j} \in F\right),$$

where $Z^{(n)} = (Z_1,\ldots,Z_n)$ with $Z_j \sim \mathbf{N}_p$ i.i.d. and $W_n \sim \mathbf{W}_n$ independent

of $Z^{(n)}$. Moreover, let

$$(4.2) \qquad \tilde{\mu}_{VX^{(n)}}(F) := \mathbb{P}\left(\left(\frac{1}{1+\alpha}\right)^{1/p} \sum_{j=1}^{n} Z_j V_{\bullet,j} \in F\right),$$

again with i.i.d. $Z_j \sim \mathbf{N}_p$. We shall see that we can confine our analysis to $\tilde{\mu}_{VX^{(n)}}$ instead of $\mu_{VX^{(n)}}$, since they are arbitrarily close to each other in $n \in \mathbb{N}$ with respect to the Lévy–Prokhorov metric. On the space $\mathcal{M}_1(\mathbb{R}^k)$ of probability measures on $\mathbb{R}^k$, the Lévy–Prokhorov metric ρ_{LP} is defined by

$$\rho_{\mathrm{LP}}(\mu,\nu) := \inf\{\varepsilon > 0 \colon \mu(F) \leqslant \nu(F_\varepsilon) + \varepsilon \text{ and } \nu(F) \leqslant \mu(F_\varepsilon) + \varepsilon$$
$$\text{for all } F \in \mathcal{B}(\mathbb{R}^k)\},$$

where F_ε denotes the ε-neighbourhood of $F \in \mathcal{B}(\mathbb{R}^k)$, defined as

$$F_\varepsilon := \{x \in \mathbb{R}^k \colon \|a - x\|_2 < \varepsilon \text{ for some } a \in F\}, \quad \varepsilon > 0.$$

We shall now prove that for the Lévy–Prokhorov metric ρ_{LP} on $\mathcal{M}_1(\mathbb{R}^k)$, the distance $\rho_{\mathrm{LP}}(\mu_V, \tilde{\mu}_V)$ converges to 0 uniformly over all $V \in \mathbb{V}_{n,k}$ as $n \to \infty$.

LEMMA 4.1. *Fix* $p \in (0, \infty)$ *and* $k \in \mathbb{N}$. *For any* $n \in \mathbb{N}$ *let* $X^{(n)}$ *be as in Theorem* 3.1. *Then*

$$\lim_{n\to\infty} \sup_{V \in \mathbb{V}_{n,k}} \rho_{\mathrm{LP}}(\mu_{VX^{(n)}}, \tilde{\mu}_{VX^{(n)}}) = 0.$$

Proof. Let $F \in \mathcal{B}(\mathbb{R}^k)$ and $\varepsilon > 0$. Then

$$(4.3) \quad \tilde{\mu}_{VX^{(n)}}(F) = \mathbb{P}\left(\left(\frac{1}{1+\alpha}\right)^{1/p} \sum_{j=1}^{n} Z_j V_{\bullet,j} \in F\right)$$

$$\leqslant \mathbb{P}\left(\sum_{j=1}^{n} n^{1/p} \frac{Z_j}{(\|Z^{(n)}\|_p^p + W_n)^{1/p}} V_{\bullet,j} \in F_\varepsilon\right)$$

$$+ \mathbb{P}\left(\left\|\left(\frac{1}{1+\alpha}\right)^{1/p} \sum_{j=1}^{n} Z_j V_{\bullet,j} - \sum_{j=1}^{n} n^{1/p} \frac{Z_j}{(\|Z^{(n)}\|_p^p + W_n)^{1/p}} V_{\bullet,j}\right\|_2 \geqslant \varepsilon\right)$$

$$= \mu_{VX^{(n)}}(F_\varepsilon)$$

$$+ \mathbb{P}\left(\left\|\left(\frac{1}{1+\alpha}\right)^{1/p} \sum_{j=1}^{n} Z_j V_{\bullet,j} - \sum_{j=1}^{n} n^{1/p} \frac{Z_j}{(\|Z^{(n)}\|_p^p + W_n)^{1/p}} V_{\bullet,j}\right\|_2 \geqslant \varepsilon\right).$$

Let us prove that the second summand on the right-hand side converges to 0 as $n \to \infty$. By Markov's inequality,

$$\mathbb{P}\left(\left\|\left(\frac{1}{1+\alpha}\right)^{1/p}\sum_{j=1}^{n}Z_jV_{\bullet,j}-\sum_{j=1}^{n}n^{1/p}\frac{Z_j}{(\|Z^{(n)}\|_p^p+W_n)^{1/p}}V_{\bullet,j}\right\|_2\geqslant\varepsilon\right)$$

$$\leqslant\varepsilon^{-1}\mathbb{E}\left\|\left(\frac{1}{1+\alpha}\right)^{1/p}\sum_{j=1}^{n}Z_jV_{\bullet,j}-\sum_{j=1}^{n}n^{1/p}\frac{Z_j}{(\|Z^{(n)}\|_p^p+W_n)^{1/p}}V_{\bullet,j}\right\|_2,$$

and by the Cauchy–Schwarz inequality,

$$(4.4)\quad\mathbb{E}\left\|\left(\frac{1}{1+\alpha}\right)^{1/p}\sum_{j=1}^{n}Z_jV_{\bullet,j}-\sum_{j=1}^{n}n^{1/p}\frac{Z_j}{(\|Z^{(n)}\|_p^p+W_n)^{1/p}}V_{\bullet,j}\right\|_2$$

$$=\mathbb{E}\left(\left\|\sum_{j=1}^{n}Z_jV_{\bullet,j}\right\|_2\cdot\left|\left(\frac{1}{1+\alpha}\right)^{1/p}-\frac{n^{1/p}}{(\|Z^{(n)}\|_p^p+W_n)^{1/p}}\right|\right)$$

$$\leqslant\sqrt{\mathbb{E}\left\|\sum_{j=1}^{n}Z_jV_{\bullet,j}\right\|_2^2}\cdot\sqrt{\mathbb{E}\left|\left(\frac{1}{1+\alpha}\right)^{1/p}-\frac{n^{1/p}}{(\|Z^{(n)}\|_p^p+W_n)^{1/p}}\right|^2}.$$

As $Z_1,\ldots,Z_n$ are i.i.d. with mean zero and the rows $V_{1,\bullet},\ldots,V_{k,\bullet}$ of the Stiefel matrix V are orthonormal vectors, the first factor on the right-hand side of (4.4) satisfies

$$(4.5)\quad\mathbb{E}\left\|\sum_{j=1}^{n}Z_jV_{\bullet,j}\right\|_2^2=\mathbb{E}\left[\sum_{i,j=1}^{n}Z_iZ_j\langle V_{\bullet,i},V_{\bullet,j}\rangle_2\right]$$

$$=\mathbb{E}[Z_1^2]\sum_{j=1}^{n}\|V_{\bullet,j}\|_2^2=\mathbb{E}[Z_1^2]\sum_{i=1}^{k}\|V_{i,\bullet}\|_2^2=k\mathbb{E}[Z_1^2].$$

To address the second factor in (4.4), let us first argue that

$$(4.6)\quad\xi_n:=\left(\left(\frac{1}{1+\alpha}\right)^{1/p}-\frac{n^{1/p}}{(\|Z^{(n)}\|_p^p+W_n)^{1/p}}\right)^2\to 0$$

in probability as $n\to\infty$. Indeed, by the continuous mapping theorem, it suffices to show that

$$\frac{\|Z^{(n)}\|_p^p}{n}+\frac{W_n}{n}\to 1+\alpha$$

in probability. This follows from the fact that as $Z_1,\ldots,Z_n$ are i.i.d. p-generalized Gaussian random variables, we have $\mathbb{E}|Z_i|^p=1$, and moreover that by assumption, $W_n/n\to\alpha$ in probability. In fact, we even have $\xi_n\to 0$ in L^1. To see this, it suffices to show that $(\xi_n)_n$ is uniformly integrable, which in combination with (4.6) yields convergence in L^1. Clearly, $(\xi_n)_n$ is uniformly integrable if the sequence

$$\left(\frac{n}{\|Z^{(n)}\|_p^p+W_n}\right)^{2/p}\leqslant\left(\frac{n}{\|Z\|_p^p}\right)^{2/p}$$

is uniformly integrable, where we have used the fact that $W_n \geqslant 0$. This in turn follows from the fact that $\|Z^{(n)}\|_p^p \sim \gamma(n/p, 1/p)$ together with Lemma 3.1 for $\mathbf{W}_n = \gamma(n/p, 1/p)$, $\kappa = 1$ and $C = 1/p$. Indeed, observing that $\frac{n/p}{n} = 1/p$ and assuming without loss of generality that $n > 4$, so that we may choose $m := 5/p$, we obtain

$$\mathbb{E}\left(\frac{n}{\|Z^{(n)}\|_p^p}\right)^{4/p} \leqslant M_{p,m}$$

for all $n > 4$. Hence, $\xi_n \to 0$ in L^1, and as a consequence, the second factor in (4.4) converges to 0. This implies that the second summand in (4.3) converges to 0 uniformly in $n \in \mathbb{N}$. Altogether, we have proven that for any $\varepsilon > 0$,

$$\tilde{\mu}_{VX^{(n)}}(F) \leqslant \mu_{VX^{(n)}}(F_\varepsilon) + \varepsilon$$

for n sufficiently large. In the same way, we may also prove that

$$\mu_{VX^{(n)}}(F) \leqslant \tilde{\mu}_{VX^{(n)}}(F_\varepsilon) + \varepsilon$$

for n sufficiently large, which finishes the proof. ∎

Finally, let us replace $V \in \mathbb{V}_{n,k}$ by random variables $V_{n,k}$, i.e. the Stiefel matrix is chosen at random according to the uniform distribution $\mu_{n,k}$ on $\mathbb{V}_{n,k}$. Based on Lemma 4.1, we may prove that a weak LDP for the modified sequence $\tilde{\mu}_{V_{n,k}X^{(n)}}$ implies a weak LDP (in the sense of [4, Definition, p. 7]) for $\mu_{V_{n,k}X^{(n)}}$, both respectively defined as in (4.1) and (4.2) with respect to $V_{n,k}$.

LEMMA 4.2. *Assume the set-up of Theorem* 3.1 *and recall the notation* (4.1) *and* (4.2). *If the sequence* $\tilde{\mu}_{V_{n,k}X^{(n)}}$ *satisfies a weak LDP on* $\mathcal{M}_1(\mathbb{R}^k)$ *at speed n and with rate function I, then the sequence* $\mu_{V_{n,k}X^{(n)}}$ *satisfies the same weak LDP.*

Proof. It suffices to check the weak LDP on a basis of the topology of $\mathcal{M}_1(\mathbb{R}^k)$, e. g., the balls

$$B_r(\nu) := \{\mu \in \mathcal{M}_1(\mathbb{R}^k) : \rho_{\mathrm{LP}}(\mu, \nu) < r\}$$

for any $r \in (0, \infty)$. By Lemma 4.1, for n sufficiently large $\rho_{\mathrm{LP}}(\tilde{\mu}_{V_{n,k}X^{(n)}}, \mu_{V_{n,k}X^{(n)}})$ $< r/2$ uniformly over all realizations of $V_{n,k} \in \mathbb{V}_{n,k}$. Therefore, by the triangle inequality for ρ_{LP},

$$\frac{1}{n}\log \mathbb{P}(\tilde{\mu}_{V_{n,k}X^{(n)}} \in B_{r/2}(\nu)) \leqslant \frac{1}{n}\log \mathbb{P}(\mu_{V_{n,k}X^{(n)}} \in B_r(\nu))$$
$$\leqslant \frac{1}{n}\log \mathbb{P}(\tilde{\mu}_{V_{n,k}X^{(n)}} \in B_{3r/2}(\nu)),$$

and hence

$$\limsup_{n\to\infty} \frac{1}{n}\log\mathbb{P}(\tilde{\mu}_{V_{n,k}X^{(n)}} \in B_{r/2}(\nu)) \leqslant \liminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}(\mu_{V_{n,k}X^{(n)}} \in B_r(\nu))$$

$$\leqslant \limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}(\mu_{V_{n,k}X^{(n)}} \in B_r(\nu)) \leqslant \liminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}(\tilde{\mu}_{V_{n,k}X^{(n)}} \in B_{3r/2}(\nu)).$$

Thus, by monotonicity in r, taking the infimum over $r \in (0,\infty)$, the LDP for $\tilde{\mu}_{V_{n,k}X^{(n)}}$ yields

$$-I(\nu) \leqslant \inf_{r\in(0,\infty)} \liminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}(\mu_{V_{n,k}X^{(n)}} \in B_r(\nu))$$

$$\leqslant \inf_{r\in(0,\infty)} \limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}(\mu_{V_{n,k}X^{(n)}} \in B_r(\nu)) \leqslant -I(\nu).$$

The claim now follows from [4, Theorem 4.1.11]. ∎

On a compact space, weak and full LDPs coincide. Here, compactness is provided by the following lemma.

LEMMA 4.3. *There is a constant $C \in (0,\infty)$ such that for all $n \geqslant k$ and all $V \in \mathbb{V}_{n,k}$,*

$$\mu_{VX^{(n)}} \in M_C := \Big\{ \mu \in \mathcal{M}_1(\mathbb{R}^k) \colon \int_{\mathbb{R}^k} \|x\|_2\,\mu(\mathrm{d}x) \leqslant C \Big\},$$

where the set M_C is compact for any choice of $C \in (0,\infty)$.

Proof. The compactness of the set M_C in the weak topology on $\mathcal{M}_1(\mathbb{R}^k)$ has been shown in [10, proof of Lemma 5.3], so it remains to prove the first assertion. To this end, recalling the representation of the distribution $\mu_{VX^{(n)}}$ given in (4.1), it suffices to prove that

$$\limsup_{n\to\infty} \sup_{V\in\mathbb{V}_{n,k}} \mathbb{E}\left\|\sum_{j=1}^n n^{1/p}\frac{Z_j}{(\|Z^{(n)}\|_p^p + W_n)^{1/p}}V_{\bullet,j}\right\|_2 < \infty$$

for i.i.d. $Z_j \sim \mathbf{N}_p$ and $W_n \sim \mathbf{W}_n$ as in Theorem 3.1. By the triangle inequality it then follows that

$$\mathbb{E}\left\|\sum_{j=1}^n n^{1/p}\frac{Z_j}{(\|Z^{(n)}\|_p^p + W_n)^{1/p}}V_{\bullet,j}\right\|_2$$

$$\leqslant \mathbb{E}\left\|\Big(\frac{1}{1+\alpha}\Big)^{1/p}\sum_{j=1}^n Z_jV_{\bullet,j} - \sum_{j=1}^n n^{1/p}\frac{Z_j}{(\|Z^{(n)}\|_p^p + W_n)^{1/p}}V_{\bullet,j}\right\|_2$$

$$+ \mathbb{E}\left\|\Big(\frac{1}{1+\alpha}\Big)^{1/p}\sum_{j=1}^n Z_jV_{\bullet,j}\right\|_2.$$

The first summand on the right hand side converges to 0 uniformly in $n \in \mathbb{N}$, as was shown after (4.4). Moreover, by Hölder's inequality and (4.5), the second summand is uniformly bounded by $\sqrt{k\mathbb{E}[Z_1^2]}/(1+\alpha)^{1/p}$, and thus the claim follows. ∎

    Combining the accumulated auxiliary results, we now have the sufficient tools to prove Theorem 3.1.

Proof of Theorem 3.1. We apply Proposition 2.2 to the symmetric non-Gaussian random variables $Z_j/(1+\alpha)^{1/p}$, $Z_j \sim \mathbf{N}_p$, which, by (2.2), have finite moments of all orders and, in particular, variance

$$\sigma_{p,\alpha}^2 = \left(\frac{p}{1+\alpha}\right)^{2/p} \frac{\Gamma(3/p)}{\Gamma(1/p)}.$$

Hence, the sequence $\tilde{\mu}_{V_{n,k}X^{(n)}}$ satisfies an LDP on $\mathcal{M}_1(\mathbb{R}^k)$ with speed n and rate function I as stated in Theorem 3.1. Therefore, by Lemma 4.2, $\mu_{V_{n,k}X^{(n)}}$ satisfies the same weak LDP, which extends to a full LDP by the compactness arguments given in Lemma 4.3, thus finishing the proof. ∎

    The proof of Theorem 3.2 works in a very similar way to that of Theorem 3.1, hence we will only point out the steps where it differs from the previous proof. Given the different scaling of $X^{(n)}$, it follows that for a Stiefel matrix $V \in \mathbb{V}_{n,k}$, we have

$$(4.7) \qquad \mu_{VX^{(n)}}(F) := \mathbb{P}(VX^{(n)} \in F)$$
$$= \mathbb{P}\left(\sum_{j=1}^{n} n^{\kappa/p} \frac{Z_j}{(\|Z^{(n)}\|_p^p + W_n)^{1/p}} V_{\bullet,j} \in F\right)$$

for any $F \in \mathcal{B}(\mathbb{R}^k)$, and we set

$$(4.8) \qquad \tilde{\mu}_{VX^{(n)}}(F) := \mathbb{P}\left(\left(\frac{1}{\beta}\right)^{1/p} \sum_{j=1}^{n} Z_j V_{\bullet,j} \in F\right),$$

using the same notation as in Theorem 3.1 and its proof. The only argument that needs to be adapted is the proof of Lemma 4.1, which will be replaced by the following lemma.

    LEMMA 4.4. *For $p \in (0,\infty)$ and $k \in \mathbb{N}$. For any $n \in \mathbb{N}$ let $X^{(n)}$ be as in Theorem* 3.2. *Then*

$$\lim_{n\to\infty} \sup_{V\in\mathbb{V}_{n,k}} \rho_{\mathrm{LP}}(\mu_{VX^{(n)}}, \tilde{\mu}_{VX^{(n)}}) = 0.$$

    Proof. Let $F \in \mathcal{B}(\mathbb{R}^k)$, $V \in \mathbb{V}_{n,k}$ fixed, and $\varepsilon > 0$. Then, by an analogous expansion to (4.3), we have

$$(4.9) \quad \tilde{\mu}_{VX^{(n)}}(F) \leqslant \mu_{VX^{(n)}}(F_\varepsilon)$$
$$+ \mathbb{P}\left(\left\|\left(\frac{1}{\beta}\right)^{1/p} \sum_{j=1}^{n} Z_j V_{\bullet,j} - \sum_{j=1}^{n} n^{\kappa/p} \frac{Z_j}{(\|Z^{(n)}\|_p^p + W_n)^{1/p}} V_{\bullet,j}\right\|_2 \geqslant \varepsilon\right).$$

Again, we need to show that the second summand on the right hand side converges to 0 as $n \to \infty$. By Markov's inequality,

$$
\mathbb{P}\left(\left\|\left(\frac{1}{\beta}\right)^{1/p} \sum_{j=1}^{n} Z_j V_{\bullet,j} - \sum_{j=1}^{n} n^{\kappa/p} \frac{Z_j}{(\|Z^{(n)}\|_p^p + W_n)^{1/p}} V_{\bullet,j}\right\|_2 \geqslant \varepsilon\right)
$$

$$
\leqslant \varepsilon^{-1} \mathbb{E}\left\|\left(\frac{1}{\beta}\right)^{1/p} \sum_{j=1}^{n} Z_j V_{\bullet,j} - \sum_{j=1}^{n} n^{\kappa/p} \frac{Z_j}{(\|Z^{(n)}\|_p^p + W_n)^{1/p}} V_{\bullet,j}\right\|_2,
$$

and a further application of the Cauchy–Schwarz inequality as in (4.4) yields

$$
(4.10) \quad \mathbb{E}\left\|\left(\frac{1}{\beta}\right)^{1/p} \sum_{j=1}^{n} Z_j V_{\bullet,j} - \sum_{j=1}^{n} n^{\kappa/p} \frac{Z_j}{(\|Z^{(n)}\|_p^p + W_n)^{1/p}} V_{\bullet,j}\right\|_2
$$

$$
\leqslant \sqrt{\mathbb{E}\left\|\sum_{j=1}^{n} Z_j V_{\bullet,j}\right\|_2^2} \cdot \sqrt{\mathbb{E}\left|\left(\frac{1}{\beta}\right)^{1/p} - \frac{n^{\kappa/p}}{(\|Z^{(n)}\|_p^p + W_n)^{1/p}}\right|^2},
$$

with the first factor simplifying to $k\mathbb{E}[Z_1^2]$ as in (4.5). To address the second factor, it remains to show that

$$
(4.11) \quad \xi_n := \left(\left(\frac{1}{\beta}\right)^{1/p} - \frac{n^{\kappa/p}}{(\|Z^{(n)}\|_p^p + W_n)^{1/p}}\right)^2 \to 0
$$

in probability as $n \to \infty$. We again do so by showing that

$$
\frac{\|Z^{(n)}\|_p^p}{n^\kappa} + \frac{W_n}{n^\kappa} \to \beta
$$

in probability due to the continuous mapping theorem. Since $\kappa > 1$, by the same arguments as in the proof of Theorem 3.1 it follows that $\|Z^{(n)}\|_p^p/n^\kappa \to 0$ and the behaviour of W_n dominates. By assumption, $W_n/n^\kappa \to \beta$ in probability. In fact, we even have $\xi_n \to 0$ in L^1. Indeed, since the sequence of random variables $(W_n/n^\kappa)^{-2/p}$ is uniformly integrable by assumption, it follows that the sequence of random variables ξ_n is uniformly integrable as well, which in combination with (4.11) yields convergence in L^1. As a consequence, the second factor on the right hand side of (4.10) converges to 0. This implies that the second summand on the right hand side of (4.9) converges to 0 uniformly in $n \in \mathbb{N}$. Altogether, we have proven that for any $\varepsilon > 0$,

$$
\tilde{\mu}_{VX^{(n)}}(F) \leqslant \mu_{VX^{(n)}}(F_\varepsilon) + \varepsilon
$$

for n sufficiently large, and we can prove analogously that

$$
\mu_{VX^{(n)}}(F) \leqslant \tilde{\mu}_{VX^{(n)}}(F_\varepsilon) + \varepsilon
$$

for n sufficiently large, thus finishing the proof.   ∎

Since the rest of the proof of Theorem 3.1 does not depend on the specific choice of α or the scaling of the $X^{(n)}$, the remainder of the proof of Theorem 3.2 can proceed in the very same way. It remains only to present the proof of Lemma 3.1.

Proof of Lemma 3.1. We start by observing that

$$\mathbb{E}\left[\left(\frac{n^{\kappa}}{W_n}\right)^{4/p}\right] = n^{4\kappa/p}\frac{b^{a_n}}{\Gamma(a_n)}\int_0^{\infty} x^{a_n-1-4/p}e^{-bx}\,\mathrm{d}x = n^{4\kappa/p}\frac{b^{4/p}}{\Gamma(a_n)}\Gamma\left(a_n - \frac{4}{p}\right).$$

Applying first A_p times the defining property of the gamma function and using then the inequality of [9, (12)] for quotients of gamma functions (applied with $x = a_n + A_p + 1 - 4/p$ and $y = 4/p$, where we recall that $A_p = \lceil 4/p \rceil$) one can see that

$$\begin{aligned}
\frac{\Gamma\left(a_n - \frac{4}{p}\right)}{\Gamma(a_n)} &= \left(\prod_{i=0}^{A_p} \frac{a_n + i}{a_n - \frac{4}{p} + i}\right)\frac{\Gamma\left(a_n - \frac{4}{p} + A_p + 1\right)}{\Gamma(a_n + A_p + 1)} \\
&= \left(\frac{1}{\prod_{i=0}^{A_p}\left(1 - \frac{4}{p(a_n+i)}\right)}\right)\frac{1}{\frac{\Gamma(a_n+A_p+1)}{\Gamma(a_n-\frac{4}{p}+A_p+1)}} \\
&\leqslant \left(\frac{1}{\prod_{i=0}^{A_p}\left(1 - \frac{4}{p(m+i)}\right)}\right)\frac{1}{(a_n + A_p - \frac{4}{p})^{4/p}} \\
&\leqslant \left(\frac{1}{\prod_{i=0}^{A_p}\left(1 - \frac{4}{p(m+i)}\right)}\right)\frac{1}{a_n^{4/p}} = M_{p,m}\frac{1}{a_n^{4/p}},
\end{aligned}$$

where we have also used $a_n \geqslant m > 4/p$ for all $n \in \mathbb{N}$. By our assumption that $a_n/n^{\kappa} \geqslant C$ it follows that

$$\mathbb{E}\left[\left(\frac{n^{\kappa}}{W_n}\right)^{4/p}\right] \leqslant M_{p,m}\frac{n^{4\kappa/p}b^{4/p}}{a_n^{4/p}} = M_{p,m}b^{4/p}\left(\frac{n^{\kappa}}{a_n}\right)^{4/p} \leqslant M_{p,m}\left(\frac{b}{C}\right)^{4/p}$$

for all $n \in \mathbb{N}$. This completes the proof. $\blacksquare$

REFERENCES

[1] D. Alonso-Gutiérrez, J. Prochno, and C. Thäle, *Large deviations for high-dimensional random projections of ℓ_p^n-balls*, Adv. Appl. Math. 99 (2018), 1–35.

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

[3] D. Chafaï, O. Guédon, G. Lecué, and A. Pajor, *Interactions between Compressed Sensing Random Matrices and High Dimensional Geometry*, Soc. Math. France, 2012.

[4] A. Dembo and O. Zeitouni, *Large Deviations: Techniques and Applications*, Stochastic Modelling Appl. Probab. 38, Springer, Berlin, 2010 (corrected reprint of the second (1998) edition).

[5] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*, Appl. Numer. Harmonic Anal., Birkhäuser/Springer, New York, 2013.

[6] N. Gantert, S. S. Kim, and K. Ramanan, *Large deviations for random projections of ℓ^p balls*, Ann. Probab. 45 (2017), 4419–4476.

[7] A. Hinrichs, D. Krieg, E. Novak, J. Prochno, and M. Ullrich, *Random sections of ellipsoids and the power of random information*, Trans. Amer. Math. Soc. 374 (2021), 8691–8713.

[8] A. Hinrichs, J. Prochno, and M. Sonnleitner, *Random sections of ℓ_p-ellipsoids, optimal recovery and Gelfand numbers of diagonal operators*, arXiv:2109.14504 (2021).

[9] G. J. O. Jameson, *Inequalities for gamma function ratios*, Amer. Math. Monthly 120 (2013), 936–940.

[10] Z. Kabluchko and J. Prochno, *Large deviations for random matrices in the orthogonal group and Stiefel manifold with applications to random projections of product distributions*, arXiv:2110.12977 (2021).

[11] Z. Kabluchko, J. Prochno, and C. Thäle, *High-dimensional limit theorems for random vectors in ℓ_p^n-balls*, Commun. Contemp. Math. 21 (2019), art. 1750092, 30 pp.

[12] Z. Kabluchko, J. Prochno, and C. Thäle, *Sanov-type large deviations in Schatten classes*, Ann. Inst. H. Poincaré Probab. Statist. 56 (2019), 928–953.

[13] Z. Kabluchko, J. Prochno, and C. Thäle, *High-dimensional limit theorems for random vectors in ℓ_p^n-balls. II*, Commun. Contemp. Math. 23 (2021), art. 1950073, 35 pp.

[14] T. Kaufmann and C. Thäle, *Weighted p-radial distributions on Euclidean and matrix p-balls with applications to large deviations*, J. Math. Anal. Appl. 515 (2022), no. 1, art. 126377, 41 pp.

[15] S. S. Kim and K. Ramanan, *Large deviation principles induced by the Stiefel manifold, and random multi-dimensional projections*, Adv. in Appl. Math. 134 (2022), art. 102306, 64 pp.

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

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

[18] D. Krieg and M. Sonnleitner, *Random points are optimal for the approximation of Sobolev functions*, arXiv:2009.11275 (2020).

[19] J. Prochno, C. Thäle, and N. Turchi, *Geometry of ℓ_p^n-balls: Classical results and recent developments*, in: High Dimensional Probability VIII, Progr. Probab. 74, Birkhäuser, 2019, 121–150.

Tom Kaufmann, Holger Sambale, Christoph Thäle
Ruhr University Bochum
Faculty of Mathematics
Universitätsstr. 150
44780 Bochum, Germany
E-mail: tom.kaufmann@rub.de
     holger.sambale@rub.de
     christoph.thaele@rub.de