Berry-Esseen Bounds for Random Projections of $\ell_p^n$-Balls

SAMUEL JOHNSTON AND JOSCHA PROCHNO

Abstract. In this work we study the rate of convergence in the central limit theorem for the Euclidean norm of random orthogonal projections of vectors chosen at random from an $\ell_p^n$-ball which has been obtained in [Alonso-Gutiérrez, Prochno, Thäle: Gaussian fluctuations for high-dimensional random projections of $\ell_p^n$-balls, Bernoulli 25(4A), 2019, 3139–3174]. More precisely, for any $n \in \mathbb{N}$ let $E_n$ be a random subspace of dimension $k_n \in \{1, \ldots, n\}$, $P_{E_n}$ the orthogonal projection onto $E_n$, and $X_n$ be a random point in the unit ball of $\ell_p^n$. We prove a Berry-Esseen theorem for $\|P_{E_n}X_n\|_2$ under the condition that $k_n \to \infty$. This answers in the affirmative a conjecture of Alonso-Gutiérrez, Prochno, and Thäle who obtained a rate of convergence under the additional condition that $k_n/n^{2/3} \to \infty$ as $n \to \infty$. In addition, we study the Gaussian fluctuations and Berry-Esseen bounds in a 3-fold randomized setting where the dimension of the Grassmannian is also chosen randomly. Comparing deterministic and randomized subspace dimensions leads to a quite interesting observation regarding the central limit behavior. In this work we also discuss the rate of convergence in the central limit theorem of [Kabluchko, Prochno, Thäle: High-dimensional limit theorems for random vectors in $\ell_p^n$-balls, Commun. Contemp. Math. (2019)] for general $\ell_p$-norms of non-projected vectors chosen at random in an $\ell_p^n$-ball.

1. Introduction and Results

The last decades have shown the power of probability theory in the study of high-dimensional convex bodies. Although high dimensions typically increase the complexity of problems, it has become apparent that convexity enforces a certain regularity in the geometry of the underlying space and in the behavior of random objects. The area dealing with random phenomena as the space dimension tends to infinity is also known as high-dimensional probability theory and it combines powerful methods from probability theory, functional analysis, and convex geometry.

One of the first results in this direction is a high-dimensional central limit theorem commonly known as the Poincaré-Maxwell-Borel Lemma (see [6]), showing that the first $k$ coordinates of a point uniformly distributed over the $n$-dimensional Euclidean ball or sphere are independent standard normal variables in the limit as $n \to \infty$ with $k$ fixed. Later Stam studied the central limit behavior when $k$ is allowed to vary with the dimension $n$ [15]. Probably the best example of the last decade is a celebrated result of Klartag. He proved a central limit theorem for high-dimensional convex bodies, showing that most $k$-dimensional marginals of a random vector uniformly distributed in an isotropic convex body are approximately Gaussian, provided that $k = k_n$ is smaller than $n^\kappa$ for some absolute constant $\kappa \in (0,1)$ known to satisfy $\kappa < 1/14$ [11] [12]. Other than Klartag’s central limit theorem there are only a few random geometric quantities that have been shown to satisfy a central limit theorem. For instance, central limit phenomena for the log-volume of random simplices in high dimensions have recently been studied by Grote, Kabluchko, and Thäle in [7]. Then there is the central limit theorem for the volume of $k$-dimensional random projections of the $n$-dimensional cube obtained by Paouris, Pivovarov, and Zinn in [13], which in the case $k=1$ was obtained independently by Kabluchko, Litvak, and Zaporozhets [8]. In fact, when $k=1$ we can reformulate the result into a central limit theorem for the $\ell_p^1$-norm $\|\theta\|_1$, where $\theta$ is a random vector uniformly distributed on the Euclidean sphere $S^{n-1}$. This has recently been extended by Kabluchko, Prochno, and Thäle to a central limit theorem for arbitrary $\ell_p^n$-balls, $1 \leq p \leq \infty$ [9] [10]. Gaussian fluctuation in the Grassmannian setting were then studied by Alonso-Gutiérrez, Prochno, and Thäle in [2]. They proved a central
limit theorem for the Euclidean norm of random orthogonal projections and also obtained Berry-Esseen bounds on the rate of convergence in their central limit theorem whenever the subspace dimensions grow fast enough as the dimension \( n \) of the ambient space tends to infinity.

In this paper we investigate the rate of convergence in the Grassmannian setting studied in [2]. For \( n \in \mathbb{N} \), denote by \( E_n \) a random \( k_n \)-dimensional subspace of \( \mathbb{R}^n \), by \( P_{E_n} \) the orthogonal projection onto \( E_n \), and by \( X_n \) a random point in the unit ball of \( \ell_p^n \). Alonso-Gutiérrez, Prochno, and Thäle proved a Berry-Esseen theorem for the sequence \( \| P_{E_n} X_n \|_2, n \in \mathbb{N} \) of random variables under the condition that the subspace dimensions \( k_n \) tend to infinity and satisfy \( k_n / n^{2/3} \to \infty \) as \( n \to \infty \). Here we weaken the assumption considerably by showing that the condition \( k_n / n^{2/3} \to \infty \) as \( n \to \infty \) may be lifted. Contrary to the more analytic approach pursued in [2], we use a probabilistic one based on the geometry of Kolmogorov-Smirnov distances and the relationship between different Gaussian densities under this metric. One of the ingredients in our approach is giving a bound on the Kolmogorov-Smirnov distance between different Gaussians in terms of the variances. This can be seen as a version of a result appearing in a paper of Brehm and Voigt [5]. This tool teams up with basic properties of Kolmogorov-Smirnov distances and the classical Berry-Esseen theorem in allowing us to establish convergence in the Grassmannian setting studied in [2].

We also discuss a related problem to the one studied in [2], where the dimension \( K_n \) of the Grassmannian is chosen randomly according to a binomial distribution. Here an underlying probabilistic structure makes certain aspects of this 3-fold randomized problem simpler than its “deterministic” counterpart: the sum of \( n \) random variables may be represented as a sum of \( n \) random variables multiplied by indicator functions, and as a result the key random variables involved lend themselves more easily to use in the Berry-Esseen theorem.

In addition to the Grassmannian setting, we provide a Berry-Esseen type central limit theorem for the \( \ell_p^n \)-norm of a random vector in an \( \ell_p^n \) ball, where \( 0 < p \leq \infty \) and \( q \neq p \) with \( 0 < q < \infty \). This setting has been investigated by Kabluchko, Prochno, and Thäle in [9] [10]. For \( q = 2 \) the Berry-Esseen bound already follows from [2] Theorem 1.2], when the subspace dimensions \( k_n \) are chosen to be \( n \) (and so \( k_n / n^{2/3} \) tends to infinity as \( n \) does).

1.1. The distributions on the \( \ell_p^n \)-balls. Before stating our main results, we need to introduce our class of probability measures on unit balls of \( \ell_p^n \), following the description in [2] [3]. Let \( ||x||_p := (\sum_{i=1}^n |x_i|^p)^{1/p} \) be the \( p \)-norm, and let \( \mathbb{B}_p^n \) be the unit ball of \( \ell_p^n \). Let \( \mathbb{U}_{n,p} \) be the uniform distribution on \( \mathbb{B}_p^n \) and let \( \mathcal{C}_{n,p} \) be the cone probability measure on its boundary. (We give a full definition for the cone measure on Section 2.2.) Let \( \mathbb{W} \) be any Borel probability measure on \( [0, \infty) \). In this paper we consider distributions on \( \mathbb{B}_p^n \) of the form

\[
\mathbb{P}_{n,p,W} := \mathbb{W}(\{0\}) \mathcal{C}_{n,p} + \mathbb{U}_{n,p},
\]

where \( H : \mathbb{B}_p^n \to \mathbb{R} \) is given by \( H(x) = h(||x||_p) \) with

\[
h(r) := \frac{1}{\Gamma(1 + \frac{n}{p})} \frac{1}{(1 - r^p)^{1+n/p}} \int_0^\infty s^{n/p} e^{-s r^p/(1-r^p)} \mathbb{W}(ds), \quad r \in [0, 1]. \tag{1}
\]

In other words this means that

\[
\int_{\mathbb{B}_p^n} f(x) \mathbb{P}_{n,p,W}(dx) = \mathbb{W}(\{0\}) \int_{\mathbb{B}_p^{n-1}} f(x) \mathcal{C}_{n,p}(dx) + \int_{\mathbb{B}_p^n} f(x) H(x) \mathbb{U}_{n,p}(dx)
\]

for all non-negative measurable functions \( f : \mathbb{B}_p^n \to \mathbb{R} \).

It is a plain that when \( \mathbb{W} \) is the Dirac mass at 0 the measure \( \mathbb{P}_{n,p,W} \) coincides with the cone probability measure \( \mathcal{C}_{n,p} \). On the other hand, the choice

\[
\mathbb{W}(ds) := p^{-1} e^{-s/p} ds
\]

yields that \( H \equiv 1 \) and \( \mathbb{W}(\{0\}) = 0 \), and hence in this case \( \mathbb{P}_{n,p,W} \) is equal to the uniform measure on \( \mathbb{B}_p^n \). Let us also mention here that when \( \mathbb{W} \) is the gamma distribution with a certain parameter, a random vector with law \( \mathbb{P}_{n,p,W} \) may be understood as an \( n \) dimensional projection
of a random element of $\mathbb{B}^{n+m}_{p}$ \cite{2, 3}. With these special cases in mind, for the remainder of this article we will work in full generality, with the probability measure $\mathcal{W}$ at hand.

Define

$$M_p(r) := \frac{p^{r/p}}{r+1} \Gamma(1 + \frac{r+1}{p}).$$

As we will see below (Lemma 2.1), the constants $M_p(r)$ correspond to the moments of the so-called $p$-Gaussian distribution, and will feature in all of our results. Furthermore, in all three of our main results, Theorem A, Theorem B and Theorem C we will encounter the quantity

$$\sigma^2(p, q) := \frac{1}{q^2} \left( \frac{M_p(2q)}{M_p(q)^2} - 1 \right) - \frac{2}{pq} \left( \frac{M_p(p+q)}{M_p(q)} - 1 \right) + \frac{1}{p^2} (M_p(2p) - 1),$$

which is the variance of the centered random variable $\frac{|Z^n-M_p(q)|}{qM_p(q)} - \frac{|Z^n_p-1}{p}$ where $Z$ is $p$-Gaussian.

We emphasize that $\sigma^2(p, q)$ is not symmetric in $p$ and $q$.

Finally, the Kolmogorov-Smirnov distance between a pair of random variables is given by

$$d_{KS}(X, Y) := \sup_{t \in \mathbb{R}} |\mathbb{P}(X \leq t) - \mathbb{P}(Y \leq t)|.$$
1. Random projections onto subspaces of random dimension. We will also consider subspace-valued random variables of random dimension. For $n \in \mathbb{N}$, $k \in \{0, \ldots, n\}$, and $\lambda \in [0, 1]$ consider the probability measure $\mu_{n,\lambda}$ on the set $G_n := \bigcup_{k=0}^{n-1} G_{n,k}$ of all subspaces of $\mathbb{R}^n$ defined as

$$
\mu_{n,\lambda}(E_{n,k}) = {n \choose k} \lambda^k (1 - \lambda)^{n-k} \nu_{n,k}(E_{n,k}),
$$

whenever $E_{n,k}$ is a measurable subset of $G_{n,k}$. In particular, if $E_n$ is distributed according to $\mu_{n,\lambda}$, then the dimension of $E_n$ is binomially distributed with parameters $n$ and $\lambda$.

The measure $\mu_{n,\lambda}$ may be constructed as follows. Let $\{e_i\}_{1 \leq i \leq n}$ be the standard unit vector basis of $\mathbb{R}^n$, and let $I_\lambda$ be the random subset of $\{1, \ldots, n\}$ generated by independently including each element $1 \leq i \leq n$ in $I_\lambda$ with probability $\lambda$, and not including $i$ with probability $1 - \lambda$. Now consider the random subspace

$$
E_\lambda := \text{span} \{e_i : i \in I_\lambda\}
$$

of $\mathbb{R}^n$ with random dimension $\#I_\lambda$, where $\#$ denotes the cardinality of a finite set. Finally, let $T$ be a Haar distributed element of the orthogonal group of dimension $n$. Then the image $TE_\lambda$ of $E_\lambda$ under $T$ is distributed according to $\mu_{n,\lambda}$.

The following theorem concerns the asymptotic Euclidean size of a random vector with law $\mathbb{P}_{\mathbb{R}^n,\mathbb{W}}$ projected onto a random subspace $E_n$ of distribution $\mu_{n,\lambda_n}$. The sequence $(\lambda_n n)_{n \in \mathbb{N}}$ appearing in the theorem may be interpreted as the average dimension of a subspace with distribution $\mu_{n,\lambda}$ we project onto.

**Theorem B.** Let $1 \leq p < \infty$ and $p \neq 2$. Let $\lambda_n$ be a sequence of elements of $(0, 1]$ such that $\lambda_n$ converges to some $\lambda \in [0, 1]$. Let $X_n$ be a random variable with law $\mathbb{P}_{\mathbb{R}^n,\mathbb{W}}$, and let $E_n$ be a random subspace of $\mathbb{R}^n$ with law $\mu_{n,\lambda_n}$ which is independent of $X_n$. Define the random variable

$$
Y_n := \frac{n^{1/p}}{\sqrt{M_p(2)}} \|P_{E_n}X_n\|_2 - \sqrt{\lambda_n n}.
$$

Then, as $n \to \infty$, the random variables $Y_n$ converge in distribution to a centered Gaussian random variable $G$ with variance $\lambda_n^2(p, 2) + \frac{3}{4}(1 - \lambda)$. Moreover,

$$
d_{KS}(Y_n, G) \leq C_p \max\left\{\frac{\log(n)}{\sqrt{\lambda_n n}}, |\lambda_n - \lambda|, \mathbb{P}\left[W > \frac{c_p}{\lambda_n} \log(n)\right]\right\},
$$

where $W$ is a non-negative, independent random variable with distribution $\mathbb{W}$ and $c_p, C_p \in (0, \infty)$ are constants depending only on $p$.

We conclude this subsection with two interesting observations that arise from a comparison of Theorem \textbf{A} and Theorem \textbf{B}.

**Remark 1.1.** Comparing the variances that occur in Theorem \textbf{A} and Theorem \textbf{B}, we notice that the one in the latter is larger. This however seems natural given the increased randomness in that setting.

Moreover, we find it particularly interesting and surprising that the additional randomness gives rise to a term $-\frac{3}{4} \lambda$ in the variance, because exactly such an additional term also arises in the comparison of a central limit theorem for random projections and deterministic projections in [10] (see discussion after Theorem E in that paper).

**Remark 1.2.** Theorem \textbf{A} together with Theorem \textbf{B} leads us to the observation that we have the same central limit theorem regardless of whether we project onto uniform random subspaces of dimensions $k_n$ or onto random subspaces of random dimension, provided their dimensions are sufficiently large, i.e., $\lambda_n \to 1$ and $k_n/n \to 1$ as $n \to \infty$. On the other hand, if $\lambda_n$ coincides with the sequence $k_n/n$ in Theorem \textbf{A} and $\lambda_n \to \lambda \in [0, 1)$, then we still have central limit theorems with the same centering, but this time with different limiting variances.
1.4. The $q$-norm of random elements of $\mathbb{R}^n_p$. Finally, we study the asymptotics of the $\ell_q$-norm of a random element of $\mathbb{R}^n_p$ as the dimension $n$ tends to infinity. To this end, let us first recall an abbreviated version of the main result in [9] Theorem 1.1 (a), $d = 1$. Assume that $1 \leq p \neq q < \infty$ and let $X_n$ be a random vector uniformly distributed on $\mathbb{B}_p^n$, and define the random variable

$$Y_n := \sqrt{n} \left( n^{1/q-1/p} \frac{M_p(q)^{1/q}}{\|X_n\|_q} - 1 \right).$$

(6)

Then, as $n \to \infty$, $Y_n$ converges in distribution to $G$, a centered Gaussian random variable with variance $\sigma^2(p, q)$ as given in [14]. (The statement we give here is abbreviated in the sense that the authors of [9] work in a multivariate setting, considering the simultaneous convergence for different values of $q$.) Our final result, Theorem C, establishes an upper bound on the rate at which $Y_n$ converges to $G$ in terms of the Kolmogorov-Smirnov distance. In fact, we opt to work in a slightly more general setting, letting the random vector $X_n$ have distribution $\mathbb{P}_{n,p,W}$ for a general choice of $W$ as was considered recently in [10].

**Theorem C.** Let $1 \leq p \neq q < \infty$. Let $X_n$ be a random vector with distribution $\mathbb{P}_{n,p,W}$, $Y_n$ be defined as in (14), and let $G$ be a centered Gaussian random variable with variance $\sigma^2(p, q)$. Then,

$$d_{KS}(Y_n, G) \leq \frac{C_{p,q} \log n}{\sqrt{n}} + \mathbb{P} \left[ W > c_{p,q} \sqrt{n \log n} \right],$$

where $W$ is a non-negative, independent random variable with distribution $\mathbb{W}$ and $c_{p,q}, C_{p,q} \in (0, \infty)$ depend only on $p$ and $q$. In particular, the sequence $Y_n$ converges in distribution to $G$.

Since the choice $\mathbb{W}(ds) = p^{-1} e^{-s/p} ds$ results in $\mathbb{P}_{n,p,W}$ being the uniform distribution on the unit ball $\mathbb{B}_p^n$, Theorem C implies the central limit theorem of Kabluchko, Prochno, and Thäle in [9] and in [10] for $W_n \equiv W$ there. Here we refrain from working in this slightly more general setting where $W$ may depend on $n$.

1.5. **Overview.** The remainder of the paper is structured as follows. In Section 2 we introduce our notation, as well as some preliminary facts about the geometry of $\ell_p^n$-balls – including several probabilistic representations for random vectors that will be integral to our approach. We also initiate here our study of Kolmogorov-Smirnov distances, developing some basic facts like a triangle inequality as well as a Lipschitz continuity result for the variances of different Gaussians. These general facts will be used in the remainder of the article in Section 3 and Section 4 which are dedicated to proving Theorem A, Theorem B, and Theorem C respectively.

2. Preliminaries and notation

In this section we shall briefly present the necessary background material together with the notation we use throughout the text.

2.1. **Notation.** In this article we work with $\mathbb{R}^n$ equipped with the standard Euclidean structure. Given a subset $A$ of $\mathbb{R}^n$ and a subset $I$ of the real line, we define

$$IA := \{ra \in \mathbb{R}^n : r \in I, a \in A\}.$$

(7)

For $k \in \{1, \ldots, n\}$, we will denote by $G_{n,k}$ the Grassmannian of $k$-dimensional subspaces of $\mathbb{R}^n$. For $k = 0$, $G_{n,0}$ is simply the set containing the trivial vector space $\{0\}$, where $0$ is the origin. It is possible to endow $G_{n,k}$ with a unique rotationally invariant probability measure $\nu_{n,k}$, the Haar probability measure on $G_{n,k}$. For a subspace $E$ of $\mathbb{R}^n$, we denote by $P_E$ the orthogonal projection of $\mathbb{R}^n$ onto $E$ (with respect to the standard Euclidean structure).

For $1 \leq p < \infty$, we define the $\ell_p$-norm of an element $x = (x_1, \ldots, x_n)$ of $\mathbb{R}^n$ by

$$\|x\|_p := \left( \sum_{i=1}^n |x_i|^p \right)^{1/p}.$$
For $p \geq 1$, whenever $\mathbb{R}^n$ is endowed with the norm $|| \cdot ||_p$ it becomes a Banach space which we shall denote $\ell^n_p$. The unit ball of $\ell^n_p$ is given by $\mathbb{B}_p^n := \{ x \in \mathbb{R}^n : ||x||_p \leq 1 \}$. We shall write $S_p^{n-1} := \{ x \in \mathbb{R}^n : ||x||_p = 1 \}$ for the unit sphere in $\ell^n_p$.

Given a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and a probability measure $\mu$ on a measurable space $(E, \mathcal{E})$, we shall indicate by $X \sim \mu$ that the random object $X : \Omega \to E$ has distribution $\mu$. Given any pair of random variables $X$ and $Y$, we denote by

$$X \overset{d}{=} Y$$

their equality in distribution.

Finally, we adopt the following convention for constants. We will write $c_p, C_p$ and $c_{p,q}, C_{p,q}$ to denote positive constants depending only on $p$ and on $p, q$ respectively. Within subsections we will distinguish between constants by writing $C_p, C'_p, C''_p$ etc, but two instances of $C_p$ or $C'_p$ occurring in different sections will denote different constants. Constants $c, C \in (0, \infty)$ etc. without any subindex are absolute constants.

2.2. The $p$-Gaussian distribution. We say a real-valued random variable has the $p$-Gaussian distribution if its density with respect to the Lebesgue measure on $\mathbb{R}$ is given by

$$\phi_p(s) := \frac{e^{-|s|^p/p}}{2p^{1/p} \Gamma(1 + 1/p)} ds$$

In the case $p = 2$, we say that the random variable has the standard Gaussian distribution.

The following Lemma, lifted from [2] (see also [3], Lemma 4.1), gives the absolute moments of $p$-Gaussian random variables in terms of the gamma function.

Lemma 2.1 (Lemma 3.1 [2]). Let $1 \leq p \leq \infty$ and $r \geq 0$. Let $Z$ be $p$-Gaussian distributed. Then

$$E[|Z|^r] = M_p(r) := \frac{p^{r/p}}{r + 1} \frac{\Gamma(1 + \frac{r + 1}{p})}{\Gamma(1 + \frac{r}{p})}$$

and for $q \geq 0$,

$$\text{Cov}[|Z|^q, |Z|^r] = M_p(q + r) - M_p(q)M_p(r).$$

In light of Lemma 2.1 it is straightforward to prove that with $\sigma^2(p, q)$ as in (1.1), if $Z$ is $p$-Gaussian distributed, then

$$\text{Var} \left( \frac{|Z|^q - M_p(q)}{qM_p(q)} - \frac{|Z|^p - 1}{p} \right) = E \left( \left( \frac{|Z|^q - M_p(q)}{qM_p(q)} - \frac{|Z|^p - 1}{p} \right)^2 \right) = \sigma^2(p, q).$$

2.3. Probabilistic representations for measures on $\mathbb{B}_p^n$. We now consider probability measures on the unit ball $\mathbb{B}_p^n$ and its boundary $S_p^{n-1}$. Let $U_{n,p}$ be the uniform measure on $\mathbb{B}_p^n$. We define the cone measure $C_{n,p}$ on the boundary $S_p^{n-1}$ to be the unique probability measure with the property that

$$C_{n,p}(A) := U_{n,p}([0,1]A),$$

for measurable subsets $A$ of $S_p^{n-1}$, where $[0,1]A$ is defined as in (2.1). Recall that we defined the class of probability measures $\mathbb{P}_{n,p,W}$ on the unit ball given by

$$\mathbb{P}_{n,p,W} := \mathbb{W} \{ 0 \} C_{n,p} + HU_{n,p},$$

where $H$ is a $p$-radial function given by (1.1).

The first key result of this section is a lemma from [3], stating that a random vector distributed according to $\mathbb{P}_{n,p,W}$ may be represented in terms of independent $p$-Gaussian random variables.

Lemma 2.2 ([3], Theorem 3). Let $Z = (Z_1, \ldots, Z_n)$ be a random vector whose entries are independent and identically distributed $p$-Gaussians, and let $W$ be a non-negative real-valued random variable with law $\mathbb{W}$. Then the $n$-dimensional random vector

$$\frac{Z}{(||Z||_p^p + W)^{1/p}}$$
is distributed according to $\mathbb{P}_{n,p,W}$

We remark that in particular, Lemma 2.2 clarifies that the measure $\mathbb{P}_{n,p,W}$ is in fact a probability measure—a property that is not immediately obvious from the definition.

Alonso-Gutiérrez, Prochno, and Thäle 1 used Lemma 2.2 in conjunction with the rotational invariance of standard Gaussian random variables to give the following probabilistic representations for the Euclidean norm of orthogonal projections of random vectors onto random subspaces.

Proposition 2.3 first appeared as [1, Theorem 3.1], though the precise formulation we use is [2, Proposition 2.7].

**Proposition 2.3.** Let $X_n$ be a random vector in $\mathbb{R}^n$ with law $\mathbb{P}_{n,p,W}$, and let $E_n$ in $\mathbb{G}_{n,k}$ be a random subspace distributed according to $\nu_{n,k}$, independent of $X_n$. Then we have the distributional equality

$$||P_{E_n}X_n||_2 \overset{d}{=} \frac{\left(\sum_{i=1}^n Z_i^2\right)^{1/2}}{\left(\sum_{i=1}^n |Z_i|^p + W\right)^{1/p}} \left(\sum_{i=1}^n g_i^2\right)^{1/2},$$

where $Z_i$ are a collection of independent and identically distributed $p$-Gaussian random variables, and $g_i$ are a collection of standard Gaussian random variables.

Recall that in Section 1.3 we introduced the probability measure $\mu_{n,\lambda}$ defined on the set of all subspaces of $\mathbb{R}^n$. We require an analogue of Proposition 2.3 where the random subspace is a random subspace measure $\lambda$ rather than the Grassmannian measure $\nu_{n,k}$. Lemma 2.4 gives us such a representation.

**Lemma 2.4.** Let $X_n$ be a random vector in $\mathbb{R}^n$ with law $\mathbb{P}_{n,p,W}$, and let $E_n$ be a random subspace of $\mathbb{R}^n$ distributed according to $\mu_{n,\lambda}$, independent of $X_n$, where $\lambda \in [0,1]$. Then we have the distributional identity

$$||P_{E_n}X_n||_2 \overset{d}{=} \frac{\left(\sum_{i=1}^n Z_i^2\right)^{1/2}}{\left(\sum_{i=1}^n |Z_i|^p + W\right)^{1/p}} \left(\sum_{i=1}^n I_i^{(\lambda)} g_i^2\right)^{1/2},$$

where $Z_i$ are independent $p$-Gaussians, $g_i$ are independent standard Gaussians, and $I_i^{(\lambda)}$ are independent Bernoulli random variables with probability $\lambda$ of taking the value 1 and probability $1 - \lambda$ of taking the value 0.

**Proof.** By definition, the random subspace measure $\mu_{n,\lambda}$ is equal to the random measure $\nu_{n,K_n}$, where $K_n$ is binomially distributed with parameters $n$ and $\lambda$. The result follows by noting that the if $K_n$ is independent of the $g_i$ and $Z_i$, and is binomially distributed with parameters $n$ and $\lambda$, then we have the equality in law

$$\sum_{i=1}^n I_i^{(\lambda)} g_i^2 \overset{d}{=} \sum_{i=1}^{K_n} g_i^2.$$ Comparing (2.3) with (2.4), this implies the result. \qed

In the next section we introduce the Kolmogorov-Smirnov distance between random variables, and develop some first properties of this distance.

### 2.4. Kolmogorov-Smirnov distances

Recall that the Kolmogorov-Smirnov distance between two real-valued random variables $X$ and $Y$ is given by

$$d_{KS}(X,Y) := \sup_{t \in \mathbb{R}} |\mathbb{P}(X \leq t) - \mathbb{P}(Y \leq t)|.$$ (10)

In the cases that $F$ and $G$ are distribution functions associated with $X$ and $Y$, or $f$ and $g$ are the probability densities of $X$ and $Y$, we will abuse notation and we will write $d_{KS}(F,G)$ and $d_{KS}(f,g)$ for $d_{KS}(X,Y)$, the Kolmogorov-Smirnov distance between $X$ and $Y$. (The meaning will always be clear from context.)
We now elaborate on some properties of Kolmogorov-Smirnov distance. First we note that if \( X \) and \( Y \) have continuous densities \( f \) and \( g \), then the supremum is attained at a point \( t \in \mathbb{R} \) such that \( f(t) = g(t) \). (There may be more than one such point.)

The total variation distance between real-valued random variables \( X \) and \( Y \) is defined as

\[
d_{TV}(X,Y) := \sup_{A} |\mathbb{P}(X \in A) - \mathbb{P}(Y \in A)|,
\]

where the supremum is taken over all Borel subsets of the real numbers. Again, we will write \( d_{TV}(F,G) \) and \( d_{TV}(f,g) \) for the total variation distance between distribution functions and densities. We will also use the fact below that the total variation distance between two densities \( f \) and \( g \) may also be written

\[
d_{TV}(f,g) = \frac{1}{2} \int_{-\infty}^{\infty} |f(s) - g(s)| \, ds.
\]

In any case, by considering Borel subsets of the form \( A_t := \{ s \in \mathbb{R} : s \leq t \} \) in (2.4) and comparing with (2.4), it is plain that

\[
d_{K}(f,g) \leq d_{TV}(f,g).
\]

Equality holds in (2.4) whenever \( f \) and \( g \) are continuous and the equation \( f(t) = g(t) \) has exactly one real solution.

We also remark that the Kolmogorov-Smirnov distance is invariant under rescaling of the random variables. Namely, if \( X \) and \( Y \) are random variables and \( \lambda \) is a non-zero real number, then

\[
d_{K}(\lambda X, \lambda Y) = d_{K}(X,Y).
\]

Moreover, it is immediate that Kolmogorov-Smirnov distances satisfy the triangle inequality

\[
d_{K}(X,Z) \leq d_{K}(X,Y) + d_{K}(Y,Z).
\]

2.4.1. Kolmogorov-Smirnov distances between different Gaussian densities. Now we take a moment to consider the Kolmogorov-Smirnov distances between centered Gaussian random variables with different variances \( \sigma \) and \( \tau \). We remark that by (2.4), when \( N_{\sigma} \) and \( N_{\tau} \) are Gaussian random variables of variance \( \sigma^2 \) and \( \tau^2 \) respectively, the quantity \( d_{K}(N_{\sigma}, N_{\tau}) \) only depends on the ratio \( \tau / \sigma \).

**Lemma 2.5.** For \( \nu \in (0, \infty) \), let \( f_{\nu}(s) := \frac{1}{\sqrt{2\pi s^2}} e^{-s^2/2\nu^2} \), \( s \in \mathbb{R} \). Then, whenever \( \sigma < \tau \),

\[
d_{K}(f_{\sigma}, f_{\tau}) \leq \frac{1}{4} \left(1 - \frac{\sigma}{\tau} \right) + \frac{1}{8} \left( \frac{\tau^2}{\sigma^2} - 1 \right).
\]

**Proof.** Since \( \tau > \sigma \), we note that there is a \( s_0 > 0 \) such that \( f_{\tau}(s) > f_{\sigma}(s) \) whenever \( s < -s_0 \) or \( s > s_0 \), and \( f_{\tau}(s) < f_{\sigma}(s) \) whenever \( -s_0 < s < s_0 \). In particular,

\[
d_{K}(f_{\sigma}, f_{\tau}) = \int_{-\infty}^{-s_0} f_{\tau}(s) - f_{\sigma}(s) \, ds = \frac{1}{4} \int_{-\infty}^{\infty} |f_{\sigma}(s) - f_{\tau}(s)| \, ds.
\]

Recall that \( \sigma < \tau \). Using the triangle inequality to obtain the first inequality below, we have

\[
\frac{1}{\sqrt{2\pi \sigma^2}} e^{-s^2/2\sigma^2} - \frac{1}{\sqrt{2\pi \tau^2}} e^{-s^2/2\tau^2} \leq \frac{1}{\sqrt{2\pi \tau^2}} e^{-s^2/2\tau^2} - \frac{1}{\sqrt{2\pi \tau^2}} e^{-s^2/2\tau^2} + \frac{1}{\sqrt{2\pi \tau^2}} e^{-s^2/2\tau^2} - \frac{1}{\sqrt{2\pi \tau^2}} e^{-s^2/2\tau^2} = \frac{1}{\sqrt{2\pi \sigma^2}} e^{-s^2/2\sigma^2} \left( 1 - \frac{\sigma}{\tau} \right) + \frac{1}{\sqrt{2\pi \tau^2}} \left( \frac{s^2}{\sigma^2} - 1 \right)
\]

\[
\leq \frac{1}{\sqrt{2\pi \sigma^2}} e^{-s^2/2\sigma^2} \left( 1 - \frac{\sigma}{\tau} \right) + \frac{1}{\sqrt{2\pi \tau^2}} e^{-s^2/2\tau^2} \left( \frac{1}{\sigma^2} - \frac{1}{\tau^2} \right),
\]
where the final inequality above follows from the fact that \(|1 - e^{-\alpha}| < \alpha\) for positive \(\alpha\). Integrating over \(\mathbb{R}\), we obtain

\[
\int_{-\infty}^{\infty} \left| \frac{1}{\sqrt{2\pi \sigma^2}} e^{-s^2/2\sigma^2} - \frac{1}{\sqrt{2\pi \tau^2}} e^{-s^2/2\tau^2} \right| \, ds \leq \left( 1 - \frac{\sigma}{\tau} \right) + \frac{\tau^2}{2} \left( \frac{1}{\sigma^2} - \frac{1}{\tau^2} \right) .
\]  

(16)

By combining (2.4.1) with (2.4.1), we obtain the result. \(\square\)

The following proposition states that Kolmogorov-Smirnov distance between different Gaussians is locally Lipschitz in the variance. A similar result stated in terms of total variation distance appears in Brehm and Voigt \[5, \text{Theorem 3.1}\], where a proof is given using the heat semigroup.

**Proposition 2.6.** For any \(\alpha, \beta \in (0, \infty)\) such that \(\frac{\beta}{\alpha} > \frac{1}{2}\),

\[
d_{KS}(f_{\alpha}, f_{\beta}) \leq \frac{3}{8} \frac{|\alpha^2 - \beta^2|}{\alpha^2}.
\]

**Proof.** If \(\beta \geq \alpha\), then we may write \(\beta = \alpha(1 + \epsilon)\) for some suitable \(\epsilon \geq 0\) so that \(\frac{|\alpha^2 - \beta^2|}{\alpha^2} = 2\epsilon + \epsilon^2\). Then using Lemma 2.3 to obtain the first inequality below, and the fact that \(\epsilon \geq 0\) to obtain the second, we have

\[
d_{KS}(f_{\alpha}, f_{\beta}) \leq \frac{1}{4} (1 - 1/(1 + \epsilon)) + \frac{1}{8} ((1 + \epsilon)^2 - 1)
\]

\[
\leq \frac{1}{4} (2\epsilon + \epsilon^2) = \frac{1}{4} \frac{|\alpha^2 - \beta^2|}{\alpha^2} \leq \frac{3}{8} \frac{|\alpha^2 - \beta^2|}{\alpha^2} .
\]

On the other hand, if \(\beta < \alpha\), then we may write \(\beta = \alpha(1 - \epsilon)\) for some suitable \(0 < \epsilon < \frac{1}{2}\) so that in this case \(\frac{|\alpha^2 - \beta^2|}{\alpha^2} = 2\epsilon - \epsilon^2\). Again using Lemma 2.3 to obtain the first equality below, as well as the fact that \(0 < \epsilon < \frac{1}{2}\) to obtain the second inequality below, we have

\[
d_{KS}(f_{\alpha}, f_{\beta}) \leq \frac{1}{4} (1 - (1 - \epsilon)) + \frac{1}{8} \left( \frac{1}{(1 - \epsilon)^2} - 1 \right)
\]

\[
< \frac{3}{8} (2\epsilon - \epsilon^2) = \frac{3}{8} \frac{|\alpha^2 - \beta^2|}{\alpha^2} .
\]

This completes the proof. \(\square\)

### 2.4.2. The Berry-Esseen theorem.

Under certain conditions, the central limit theorem states that the sum of many independent random variables is approximately Gaussian. We are now ready to state the most prominent result using Kolmogorov-Smirnov distances, the Berry-Esseen theorem, which gives an upper bound on the Kolmogorov-Smirnov distance between a sum \(S\) of independent centered random variables and a Gaussian random variable with the suitable variance. The version of the theorem we use, allowing the random variables occurring in the sum to be non-identically distributed, is due to Berry \[1\].

**Theorem 2.7** (Berry-Esseen). There exists a universal constant \(K_{BE} \in (0, \infty)\) with the following property: for \(m \in \mathbb{N}\) let \(S_m\) be a sum of independent and centered random variables \(X_1, \ldots, X_m\) with variances \(\sigma_1^2, \ldots, \sigma_m^2\) and third order absolute moments \(\rho_1, \ldots, \rho_m\). Let \(G_m\) be a centered Gaussian random variable with the same variance as \(S_m\). Then

\[
d_{KS}(S_m, G_m) \leq K_{BE} \max_{1 \leq i \leq m} \frac{\rho_i}{\sigma_i^3} \sqrt{\frac{\sigma_i^2}{m}} .
\]

In particular, if \(X_1, \ldots, X_m\) are identically distributed with variance \(\sigma^2\) and third absolute moment \(\rho\), then

\[
d_{KS}(S_m, G_m) \leq K_{BE} \frac{\rho}{\sigma^3 \sqrt{m}} .
\]
We remark that in Berry’s original paper, $K_{BE}$ may be taken to be 1.88. This estimate has since been sharpened significantly, most recently by Shetsova [14] who showed the conclusions of Theorem [2.7] remain valid with $K_{BE} = 0.5583$.

2.5. Some further tools. In Sections 2 and 3 we require the following lemma, lifted from [2] Lemma 4.1.

**Lemma 2.8.** Let $X^{(1)}, X^{(2)}, X^{(3)}$ be (possibly dependent) random variables, and let $G$ be a centered Gaussian random variable with variance $\sigma^2$. Then, for any $\epsilon \in (0, \infty)$,\[
sup_{t \in \mathbb{R}} \left| \mathbb{P}[X^{(1)} + X^{(2)} + X^{(3)} \geq t] - \mathbb{P}[G \geq t] \right| \\
\leq \sup_{t \in \mathbb{R}} \left( \mathbb{P}[X^{(1)} \geq t] - \mathbb{P}[G \geq t] \right) + \frac{\mathbb{P}(|X^{(2)}| > \epsilon/2) + \mathbb{P}(|X^{(3)}| > \epsilon/2)}{\sqrt{2\pi\sigma^2}}.
\]

We will use the Berry-Esseen theorem both directly and indirectly. In the latter case we have the following lemma, Lemma 2.9 which is a generalization of an idea appearing at the bottom of page 21 in [2]. Since this lemma is rather technical, before giving a statement we would like to take a moment to provide some motivation.

In our proofs of Theorems A and B we will encounter situations where for each $m \in \mathbb{N}$ we have a rescaled sum $S_m$ of $m$ i.i.d. random variables whose distributions may depend on $m$. Given a sequence $(\alpha_m)$ of reals tending to infinity, we would like the find the smallest sequence $(\alpha_m)$ ensuring that $\mathbb{P}[|S_m| > \alpha_m] \leq C/\sqrt{\beta_m}$ for some constant $C \in (0, \infty)$ not depending on $m$. Lemma 2.9 states the best possible choice is to take $\alpha_m$ of the order $\sqrt{\log(\beta_m)}$.

In the proof of Lemma 2.8 we will use the notation\[
\sqrt{a_+} := \begin{cases} \sqrt{a}, & \text{if } a \geq 0, \\ 0, & \text{if } a < 0, \end{cases}
\]
as well as the following well known inequality for Gaussian integrals, which, for any $t > 0$, provides the estimate\[
\int_t^\infty \frac{1}{\sqrt{2\pi}} e^{-u^2/2} \, du \leq \frac{1}{t} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}.
\]

**Lemma 2.9.** Let $X^{(m)}$, $m \in \mathbb{N}$ be a sequence of centered random variables with $\mathbb{E}[|X^{(m)}|^3] = \rho_m$ and $\mathbb{E}[|X^{(m)}|^2] = \sigma_m^2$. Suppose $\sigma_{\max} \geq \sigma_m$ for every $m \in \mathbb{N}$, and that $\gamma_m, m \in \mathbb{N}$ is any sequence satisfying\[
\frac{\rho_m}{\sigma_m^2} \leq \gamma_m.
\]
Suppose for each $m \in \mathbb{N}$ we have a sum\[
S_m = \frac{1}{\sqrt{m}} \left( X_1^{(m)} + \ldots + X_m^{(m)} \right),
\]
where $X_1^{(m)}, \ldots, X_m^{(m)}$ are independent copies of $X^{(m)}$. Suppose $\beta_m, m \in \mathbb{N}$ is a sequence of positive reals such that $\beta_m \leq m/\gamma_m^2$ for every $m \in \mathbb{N}$. Then, for any sequence $(\alpha_m)_{m\geq1}$ of positive reals satisfying $\alpha_m \geq \sigma_{\max}\sqrt{\log(\beta_m)}$ for every $m \in \mathbb{N}$, we have\[
\mathbb{P}[|S_m| > \alpha_m] \leq \frac{C^*}{\sqrt{\beta_m}} \quad \text{for every } m \in \mathbb{N},
\]
for $C^* = 2(K_{BE} + 1)$.

**Proof.** Since $C^* \geq 2$, the bound (2.9) holds trivially when $\beta_m \leq 4$, and we may assume without loss of generality that $\beta_m > 4$ for the remainder of the proof. Let $G_m$ be a centered Gaussian random variable with variance $\sigma_m^2$. Then using the definition of Kolmogorov-Smirnov distance and the symmetry of the Gaussian density, for any $\alpha_m > 0$ we have the inequality\[
\mathbb{P}[|S_m| > \alpha_m] \leq 2d_{KS}(S_m, G_m) + 2\mathbb{P}[G_m > \alpha_m].
\]
Indeed, it holds that
\[
\Pr(|S_m| > \alpha_m) = \Pr(|S_m| > \alpha_m) - \Pr(|G_m| > \alpha_m) + \Pr(|G_m| > \alpha_m)
\]
and since
\[
\Pr(|S_m| > \alpha_m) = 1 - \left(\Pr(S_m \leq \alpha_m) - \Pr(S_m \leq -\alpha_m)\right)
\]
as well as
\[
\Pr(|G_m| > \alpha_m) = 1 - \left(\Pr(G_m \leq \alpha_m) - \Pr(G_m \leq -\alpha_m)\right),
\]
we obtain
\[
\Pr(|S_m| > \alpha_m) - \Pr(|G_m| > \alpha_m) = \Pr(G_m \leq \alpha_m) - \Pr(S_m \leq \alpha_m) + \Pr(S_m \leq -\alpha_m) - \Pr(G_m \leq -\alpha_m)
\]
\[
\leq 2d_K(S_m, G_m).
\]
Now, on the one hand, using the Berry-Esseen theorem to obtain the first inequality below, the estimate (2.9) to obtain the second, and the fact that \(\alpha_m \geq \beta_m / \gamma_m\) to obtain the third, we have
\[
d_K(S_m, G_m) \leq \frac{K_{BE} \rho_m / \sigma_m^3}{\sqrt{m}} \leq \frac{K_{BE} \gamma_m}{\sqrt{m}} \leq \frac{K_{BE}}{\beta_m}. \tag{21}
\]
On the other hand, \(G_m / \sigma_m\) is a standard Gaussian random variable, and using the inequality (2.10) we have
\[
\Pr(G_m > \alpha_m) = \Pr\left[\frac{G_m}{\sigma_m} > \frac{\alpha_m}{\sigma_m} \leq \frac{1}{\alpha_m / \sigma_m} \right] \leq \frac{1}{\alpha_m / \sigma_m} e^{-\frac{1}{2} \left(\frac{\alpha_m}{\sigma_m}\right)^2}, \tag{22}
\]
and plugging \(\alpha_m \geq \sigma_{\max} \sqrt{\log(\beta_m)_+} \geq \sigma_m \sqrt{\log(\beta_m)_+}\) into (2.11) we have
\[
\Pr(G_m > \alpha_m) \leq \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{\beta_m \sqrt{\log(\beta_m)_+}}} \leq \frac{1}{\beta_m}, \tag{23}
\]
where in the final inequality above we used the fact that \(\beta_m > 4 > \exp(1)\) and the crude bound \(\frac{1}{\sqrt{2\pi}} \leq 1\). Combining (2.5) and (2.6) in (2.5), and using the definition of \(C^\ast\), the result follows.

\[\square\]

2.5.1. Bounds for second order functions. Our final tool provides us with a bound on the probability that a function with local second order growth takes on values bounded away from 0.

Lemma 2.10. Suppose that \(v\) is a random vector in \(\mathbb{R}^d\) and \(\Psi: \mathbb{R}^d \to \mathbb{R}\) is a function with the property that there exist \(\delta, M \in (0, \infty)\) such that
\[
|\Psi(x)| \leq M||x||_2^2 \quad \text{whenever} \quad ||x||_2 < \delta. \tag{24}
\]
Then, for any \(c \in (0, \infty)\),
\[
\Pr[|\Psi(v)| > c] \leq \Pr[||v||_2 > \sqrt{\frac{c}{M}}] + \Pr[||v||_2 \geq \delta].
\]

Proof. Using the property (2.11) to obtain the second inequality below, we have
\[
\Pr[|\Psi(v)| > c] = \Pr[|\Psi(v)| > c, ||v||_2 < \delta] + \Pr[|\Psi(v)| > c, ||v||_2 \geq \delta]
\]
\[
\leq \Pr[M||v||_2^2 > c, ||v||_2 < \delta] + \Pr[|\Psi(v)| > c, ||v||_2 \geq \delta]
\]
\[
\leq \Pr[M||v||_2^2 > c] + \Pr[||v||_2 \geq \delta].
\]
This establishes the result.

\[\square\]

3. Proof of Theorem

In this section we prove Theorem which establishes Berry-Esseen bounds for the \(\ell_2\)-norm of the rescaled projection of a random vector with distribution \(P_{n,p,W}\) onto a random \(k\)-dimensional subspace of \(\mathbb{R}^n\) distributed according to \(\nu_{n,k}\).
3.1. The set up. Let $X_n$ be a random vector with law $\mathbb{P}_{n,p,W}$, and let $E_n$ be a random subspace of $\mathbb{R}^n$ with law $\nu_{n,k_n}$, $k_n \in \{1, \ldots, n\}$. Then by Proposition 2.3 we have the identity in law

$$||P_{E_n} X_n||_2 = \frac{\left(\sum_{i=1}^n Z_i^2\right)^{1/2} \left(\sum_{i=1}^{k_n} g_i^2\right)^{1/2}}{\left(\sum_{i=1}^n |Z_i|^p + W\right)^{1/p} \left(\sum_{i=1}^{k_n} g_i^2\right)^{1/2}},$$

where $Z_i$ be a collection of independent $p$-Gaussian random variables, $W$ has law $\mathcal{W}$, and $g_i$ are independent standard Gaussian random variables.

Shifting the sums to have expectation zero, we may write

$$||P_{E_n} X_n||_2 = \frac{(nM_p(2) + \sum_{i=1}^n (Z_i^2 - M_p(2)))^{1/2} \left(k_n + \sum_{i=1}^{k_n} (g_i^2 - 1)\right)^{1/2}}{(n + \sum_{i=1}^n (|Z_i|^p - 1) + W)^{1/p} \left(n + \sum_{i=1}^{k_n} (g_i^2 - 1)\right)^{1/2}},$$

where $F(a, b, c, w, d) = (1/a)^{1/2}(1+b)^{1/2}/(1+c+w)^{1/p} + (1+d)^{1/2}$, and

$$a_n = \sum_{i=1}^n (Z_i^2 - M_p(2)), \quad b_n = \sum_{i=1}^{k_n} (g_i^2 - 1),$$

$$c_n = \sum_{i=1}^n (|Z_i|^p - 1), \quad d_n = \sum_{i=1}^{k_n} (g_i^2 - 1).$$

The function $F : \mathbb{R}^5 \to \mathbb{R}$ is twice differentiable in a neighborhood around the origin in $\mathbb{R}^5$, and hence, by Taylor’s theorem we may write

$$F(a, b, c, w, d) = 1 + \frac{1}{2}(a + b - d) - \frac{1}{p}(c + w) + \Psi(a, b, c, w, d),$$

where $\Psi : \mathbb{R}^5 \to \mathbb{R}$ is a function such that there exist $M, \delta \in (0, \infty)$ satisfying

$$|\Psi(x)| \leq M||x||_2^2 \quad \text{whenever} \quad ||x||_2 \leq \delta.$$

It follows from plugging (3.1) into (3.1) that the focal random variable $Y_n$ may be decomposed as follows:

$$Y_n := \frac{n^{1/p}}{\sqrt{M_p(2)}} ||P_{E_n} X_n||_2 - \sqrt{k_n} \xi_n - \sqrt{k_n} W \frac{a_n}{pn} + \sqrt{k_n} \Psi \left(\frac{a_n}{nM_p(2)} \frac{b_n}{k_n} \frac{c_n}{n} \frac{W}{n} \frac{d_n}{n}\right),$$

where $\xi_n := \frac{1}{\sqrt{2nM_p(2)}} a_n + \frac{1}{2n} b_n - \frac{1}{p} c_n - \frac{1}{2n} d_n$.

Assuming that $\lambda_n := k_n/n$ converges to some $\lambda \in [0, 1]$, we recall that the goal of this proof is to establish a bound for the Kolmogorov-Smirnov distance between $Y_n$ and a Gaussian random variable $G$ with variance $\nu(\lambda)$, where $\nu : [0, 1] \to \mathbb{R}$ is given by

$$\nu(s) := s \sigma^2(p, 2) + \frac{1}{2}(1 - s).$$

The main idea of the proof is that we may apply the Berry-Esseen theorem to show that $\xi_n$, which is a sum of many random variables, is close to a Gaussian random variable, and the remaining terms in the expansion (3.1) for $Y_n$ are negligible.

To this end, using Lemma 2.8 to obtain the first inequality, and the triangle inequality (2.4) for Kolmogorov-Smirnov distances to obtain the second, for any $\epsilon > 0$ we have

$$d_{KS}(Y_n, G) \leq d_{KS}(\xi_n, G) + \mathbb{P} \left[\frac{W}{pn} > \frac{\epsilon}{2\sqrt{k_n}}\right].$$
where we recall $v$ where

$$3.2. \text{ A bound for } d_{KS}(\xi_n, G_n). \text{ We note that the constituent parts } a_n, b_n, c_n, d_n \text{ of the random variable } \xi_n \text{ are clearly not independent of one another. However, } \xi_n \text{ may be written as a sum of } n + k_n + (n - k_n) \text{ independent random variables by writing}

\[
\xi_n = \sum_{i=1}^{n} \frac{\sqrt{\lambda_n} \varphi_i}{\sqrt{n}} + \sum_{i=1}^{k_n} \frac{1 - \lambda_n}{2\sqrt{k_n}} \phi_i + \sum_{i=k_n+1}^{n} \left( -\frac{\sqrt{\lambda_n(1-\lambda_n)}}{2\sqrt{n-k_n}} \right) \phi_i,
\]

where we recall $\lambda_n := k_n/n$, and define for $i \in \{1, \ldots, n\}$ the random variables

$$\varphi_i := \frac{|Z_i|^2 - M_p(2)}{2M_p(2)} - \frac{|Z_i|^p - 1}{p} \quad \text{and} \quad \phi_i := g_i^2 - 1.$$

Let $\varphi$ and $\phi$ be random variables with the same law as $\varphi_1$ and $g_1$ respectively. Then noting that the variance of $\varphi$ is given by $\sigma^2(p, 2)$ (see the display right after Lemma 2.1) and the variance of $\phi$ is simply 2, we obtain that

$$\text{Var}[\xi_n] = \lambda_n \text{Var}[\varphi] + \left( \frac{1-\lambda_n}{2} \right)^2 \text{Var}[\phi] + \frac{\lambda_n(1-\lambda_n)}{4} \text{Var}[\phi] = \lambda_n \sigma^2(p, 2) + \frac{1}{2}(1-\lambda_n) = v(\lambda_n),$$

where $v(s)$ is given by (3.1).

Now let $G_n$ be a centered Gaussian random variable with variance $v(\lambda_n)$. Then in the context of the Berry-Esseen Theorem (Theorem 2.7), with $m = n + k_n + (n - k_n)$, we have

$$\max_{1 \leq i \leq m} \frac{\rho_i}{\sigma_i^2} = \max \left\{ \left( \frac{E[|\varphi|^3]}{E[|\varphi|^2]^2} \right)^{1/2}, \left( \frac{E[|\phi|^3]}{E[|\phi|^2]} \right)^{1/2} \right\} \leq C_p A_n,$$

where

$$C_p := \max \left\{ \frac{E[|\varphi|^3]}{E[|\varphi|^2]}, \frac{E[|\phi|^3]}{E[|\phi|^2]} \right\} \quad \text{and} \quad A_n := \max \left\{ \frac{\sqrt{\lambda_n}}{\sqrt{n}}, \frac{1-\lambda_n}{2\sqrt{k_n}}, \frac{\sqrt{\lambda_n(1-\lambda_n)}}{2\sqrt{n-k_n}} \right\}.$$

Using the fact that $\lambda_n \leq 1$ and $\lambda_n := k_n/n$, it follows that $A_n \leq \frac{1}{\sqrt{k_n}}$. Consequently,

$$\max_{1 \leq i \leq m} \frac{\rho_i}{\sigma_i^2} \leq \frac{C_p}{\sqrt{k_n}},$$

and by applying the Berry-Esseen theorem to the sequence $\xi_n, n \in \mathbb{N}$ we have

$$d_{KS}(\xi_n, G_n) \leq K_{BE} \frac{C_p}{\sqrt{v(\lambda_n)\sqrt{k_n}}} \leq K_{BE} \frac{C_p}{\sqrt{J_p\sqrt{k_n}}} =: C'_p \frac{1}{\sqrt{k_n}}.$$

(31)
where the second inequality above follows from using the definition (3.1) of $J_p$.

3.3. A bound for $d_{KS}(G_n, G)$. The random variables $G_n$ and $G$ are centered Gaussians with variances given by $v(\lambda_n)$ and $v(\lambda)$ respectively. It follows from Proposition 2.6 and the definition of $J_p$ in (3.1) that

$$d_{KS}(G_n, G) \leq \frac{3}{8} |v(\lambda_n) - v(\lambda)| \leq \frac{3}{8} \frac{\sigma^2(p, 2) - \frac{1}{2}||\lambda_n - \lambda||}{J_p} = C_p |\lambda_n - \lambda|,$$

where

$$C_p := \frac{3}{8} \min(\sigma^2(p, 2), \frac{1}{2}).$$

3.4. A bound for the $\Psi$ term. In this section we take a sequence $(\epsilon_n)_{n \in \mathbb{N}}$ decreasing down to 0 in such a way that the probability

$$P \left[ \Psi \left( \frac{a_n}{nM_p(2)} \frac{b_n}{k_n}, \frac{c_n}{n}, \frac{d_n}{n} \right) > \frac{\epsilon_n}{2\sqrt{k_n}} \right]$$

may be bounded by a function of $n$ decreasing down to 0 as $n \to \infty$.

To this end, using Lemma 2.10 to obtain the first inequality below, and the triangle inequality to obtain the second, with $v_n := \left( \frac{a_n}{nM_p(2)}, \frac{b_n}{k_n}, \frac{c_n}{n}, \frac{d_n}{n} \right) \in \mathbb{R}^5$ we have

$$P \left[ \Psi (v_n) > \frac{\epsilon_n}{2\sqrt{k_n}} \right] \leq P \left[ ||v_n||_2 \geq \sqrt{\epsilon_n} \right] + P \left[ ||v_n||_2 \geq \delta \right] \leq Q_n \left( \frac{a_n}{nM_p(2)} \right) + Q_n \left( \frac{b_n}{k_n} \right) + Q_n \left( \frac{c_n}{n} \right) + Q_n \left( \frac{d_n}{n} \right),$$

where for a random variable $X$,

$$Q_n(X) := P \left[ |X| > \sqrt{\frac{\epsilon_n}{10M\sqrt{k_n}}} \right] + P \left[ |X| > \frac{\delta}{\sqrt{3}} \right].$$

(We emphasize that $Q_n$ depends on $\epsilon_n$.) We now show that with a suitable choice of $\epsilon_n$, there is a constant $C \in (0, \infty)$ such that the terms involving $a_n, b_n, c_n, d_n$ in the final expression in (3.4) may be bounded above by $C/\sqrt{k_n}$.

First consider obtaining an upper bound for

$$Q_n \left( \frac{a_n}{nM_p(2)} \right) = P \left[ \frac{a_n}{\sqrt{n}} > c_p \sqrt{n} \right] + P \left[ \frac{a_n}{\sqrt{n}} > c_p' \sqrt{\frac{n\epsilon_n}{k_n}} \right],$$

where $c_p = \frac{\delta M_p(2)}{\sqrt{3}}$ and $c_p' = \frac{M_p(2)}{\sqrt{10M}}$.

Letting $S_n := \frac{a_n}{\sqrt{n}}$, $n \in \mathbb{N}$ we are in the set up of Lemma 2.9 with $m = n$, and with two different tail probabilities we would like to bound, namely

$$\alpha_n := c_p \sqrt{n} \quad \text{and} \quad \alpha_n' := c_p' \sqrt{\frac{n\epsilon_n}{k_n}}.$$
for every \( n \in \mathbb{N} \). Now consider setting \( \epsilon_n = \Theta_p \log(k_n)/\sqrt{k_n} \) and \( \beta_n = \theta_p k_n \). Then if \( \theta_p \) is sufficiently small and \( \Theta_p \) is sufficiently large such that both the inequalities in (3.4) hold, then (3.4) holds, and there is a constant \( C'_p \in (0, \infty) \) such that

\[
Q_n \left( \frac{a_n}{nM_p(2)} \right) \leq C'_p \frac{\epsilon_n}{\sqrt{k_n}}
\]

for every \( n \in \mathbb{N} \). We may work analogously with the terms \( Q_n \left( \frac{\epsilon_n}{n} \right) \) and \( Q_n \left( \frac{d_n}{n} \right) \), which also involve tail bounds for sums of \( n \) independent random variables. It follows that for sufficiently large constants \( C''_p \) and \( \Theta'_p \), by setting \( \epsilon_n = \Theta'_p \log(k_n)/\sqrt{k_n} \) we ensure that

\[
Q_n \left( \frac{a_n}{nM_p(2)} \right) + Q_n \left( \frac{\epsilon_n}{n} \right) + Q_n \left( \frac{d_n}{n} \right) \leq C''_p \frac{\epsilon_n}{\sqrt{k_n}}
\]

for every \( n \in \mathbb{N} \).

We now consider obtaining an upper bound for \( Q_n \left( \frac{b_n}{k_n} \right) \), which this time involves finding tail probabilities for a sum of \( k_n \) random variables. We may write

\[
Q_n \left( \frac{b_n}{k_n} \right) = \mathbb{P} \left( \left| \frac{b_n}{\sqrt{k_n}} \right| > c''_p \sqrt{k_n} \right) + \mathbb{P} \left( \left| \frac{b_n}{\sqrt{k_n}} \right| > c''_p \sqrt{\epsilon_n k_n} \right),
\]

where \( c''_p = \frac{\delta}{\sqrt{n}} \) and \( c''_p = \frac{\delta}{\sqrt{10M}} \). Again we are in the context of Lemma 2.9 but this time with \( m = k_n \) and

\[
\alpha''_m = c''_p \sqrt{k_n} \quad \text{and} \quad \alpha''_m = c''_p \sqrt{\epsilon_n k_n}.
\]

(We will continue to index all variables here with a subscript \( n \) as opposed to \( m \).) Now \( b_n \) is a sum of \( k_n \) random variables distributed like \( \phi = \sqrt{q^2 - 1} \), where \( q \) is a standard Gaussian. This time take \( \gamma'_n := \mathbb{E}[|\phi|^3]/\mathbb{E}[|\phi|^2] \) for every \( n \in \mathbb{N} \), as well as \( \sigma'_{\max,p} := \mathbb{E}[|\phi|^2] \). Then by Lemma 2.9 whenever \( \beta_n \leq n/\gamma'_n \) and \( \epsilon_n \) are chosen so that the inequalities

\[
c''_p \sqrt{k_n} \geq \sigma'_{\max,p} \sqrt{\log(\beta_n)} \quad \text{and} \quad c''_p \sqrt{\epsilon_n k_n} \geq \sigma'_{\max,p} \sqrt{\log(\beta_n)}
\]

hold for every \( n \in \mathbb{N} \), we have

\[
Q_n \left( \frac{b_n}{k_n} \right) \leq \frac{4(K_{BE} + 1)}{\sqrt{\beta_n}}.
\]

In particular, setting \( \beta_n = \theta'_p k_n \) and \( \epsilon_n := \Theta'_{p} \log(k_n)/\sqrt{k_n} \), for sufficiently small \( \theta'_p \) and sufficiently large \( \Theta'_p \), the inequalities (3.4) hold for every \( n \in \mathbb{N} \). It follows that by (3.4) there is a constant \( C'''_p \in (0, \infty) \) such that

\[
Q_n \left( \frac{b_n}{k_n} \right) \leq C'''_p \frac{\epsilon_n}{\sqrt{k_n}}.
\]

In particular, by setting \( \epsilon_n := \Theta^*_p \log(k_n)/\sqrt{k_n} \) with \( \Theta^*_p := \Theta'_p \log(k_n)/\sqrt{k_n} \), both (3.4) and (3.4) hold for every \( n \in \mathbb{N} \), and hence there is a constants \( C^*_p \in (0, \infty) \) such that

\[
Q_n \left( \frac{a_n}{nM_p(2)} \right) + Q_n \left( \frac{b_n}{k_n} \right) + Q_n \left( \frac{\epsilon_n}{n} \right) + Q_n \left( \frac{d_n}{n} \right) \leq C^*_p \frac{\epsilon_n}{\sqrt{k_n}}.
\]

In particular, by plugging (3.4) into (3.4), there are constants \( \Theta^*_p, C^*_p \in (0, \infty) \) depending only on \( p \) such that with the choice \( \epsilon_n := \Theta^*_p \log(k_n)/\sqrt{k_n} \) we have

\[
\mathbb{P} \left[ \left| \frac{\Psi \left( \frac{a_n}{nM_p(2)}, \frac{b_n}{k_n}, \frac{c_n}{n}, W, \frac{d_n}{n} \right)}{\sqrt{2k_n}} \right| > \frac{\epsilon_n}{\sqrt{2k_n}} \right] \leq \frac{C^*_p}{\sqrt{k_n}} + \mathbb{P} \left[ W > \frac{\epsilon_n}{\sqrt{10Mk_n}} \right] + \mathbb{P} \left[ W > \frac{\delta}{\sqrt{n}} \right].
\]
3.5. **Wrapping things together.** Recall our original bound (3.1). We obtained the bounds (3.4) and (3.4) for the \( \epsilon_n \)-independent terms. Then in Section 3.4 we showed there are constants \( C^*_p, \Theta^*_p \in (0, \infty) \) such that if \( \epsilon_n := \Theta^*_p \log(k_n)/\sqrt{k_n} \), then (3.4) holds. Consider the collection of terms involving \( W \) that occur in (3.1) and (3.4). Noting that by definition \( W \) is non-negative, with \( \epsilon_n := \Theta^*_p \log(k_n)/\sqrt{k_n} \) it follows that there is a constant \( c_p \in (0, \infty) \) such that the \( W \) terms may be bounded by

\[
P \left[ W > n \frac{\epsilon_n}{10 M \sqrt{k_n}} \right] + P \left[ W > \frac{\delta}{\sqrt{n}} \right] + P \left[ W > \frac{\epsilon_n}{2 \sqrt{k_n}} \right] \leq 3 P \left[ W > c_p n \frac{\log(k_n)}{k_n} \right].
\]

(42)

Plugging (3.4), (3.4), (3.4) into (3.1), and using \( \epsilon_n := \Theta^*_p \log(k_n)/\sqrt{k_n} \), and (3.3), it follows that there exist constants \( c_p, C_p \in (0, \infty) \) depending on \( p \) but independent of \( \lambda_n, \lambda, \) and \( n \) such that

\[
d_{K_S}(Y_n, G) \leq C_p \max \left\{ \frac{\log(k_n)}{\sqrt{k_n}}, |\lambda_n - \lambda| \right\} + 3 P \left( W > c_p n \frac{\log(k_n)}{k_n} \right),
\]

which completes the proof of Theorem A.

### 4. Proof of Theorem B

In this section we prove Theorem B which gives Berry-Esseen bounds for the rescaled \( \ell_2 \)-norm of the orthogonal projection of a random vector with distribution \( P_{n,p,W} \) onto a \( \mu_{n,\lambda_n} \)-distributed random subspace of \( \mathbb{R}^n \). In fact, in this setting \( n \lambda_n \) has the significance of being the average dimension of a subspace with distribution \( \mu_{n,\lambda_n} \).

#### 4.1. Setting up the proof.

Let \( X_n \) be a random vector with law \( P_{n,p,W} \), and let \( E_n \) be a random subspace of \( \mathbb{R}^n \) with law \( \mu_{n,\lambda_n} \). Then by Lemma 2.4 we have the identity in law

\[
|P_{E_n}X_n|_2 = \frac{d}{\left( \sum_{i=1}^n Z_i^2 \right)^{1/2} \left( \sum_{i=1}^n I_i^{(\lambda_n)} g^2_i \right)^{1/2} \left( \sum_{i=1}^n |Z_i|^p + W \right)^{1/p} \left( \sum_{i=1}^n g_i^2 \right)^{1/2}},
\]

where \( Z_i \) are \( p \)-Gaussian random variables, \( W \) has distribution \( W \), \( g_i \) are standard Gaussian random variables, and \( I_i^{(\lambda_n)} \) are Bernoulli random variables with parameter \( \lambda_n \).

Recentering the random variables so the sums have zero expectation, we may write

\[
|P_{E_n}X_n|_2 = \frac{d}{\left( n M_p(2) + \sum_{i=1}^n (Z_i^2 - M_p(2)) \right)^{1/2} \left( n \lambda_n + \sum_{i=1}^n (I_i^{(\lambda_n)} g^2_i - \lambda_n) \right)^{1/2} \left( n + \sum_{i=1}^n (|Z_i|^p - 1) + W \right)^{1/p} \left( n + \sum_{i=1}^n g_i^2 - 1 \right)^{1/2}}
\]

\[
= n^{1/2 - 1/p} \sqrt{\lambda_n M_p(2)} F \left( \frac{a_n}{n M_p(2)}, \frac{b'_n}{\lambda_n}, \frac{c_n w}{n}, \frac{d_n}{n} \right),
\]

where \( F \) and \( a_n, c_n, d_n \) are as introduced in Section 3.1 and

\[
b'_n := \sum_{i=1}^n (I_i^{(\lambda_n)} g^2_i - \lambda_n).
\]

With \( \Psi \) as in Section 3.1 it follows that the random variable \( Y_n := \frac{n^{1/p}}{\sqrt{M_p(2)}} |P_{E_n}X_n|_2 - \sqrt{\lambda_n n} \) may be written

\[
Y_n := \xi_n - \frac{\sqrt{\lambda_n n} W}{p\sqrt{n}} + \sqrt{\lambda_n n} \Psi \left( \frac{a_n}{n M_p(2)}, \frac{b'_n}{\lambda_n}, \frac{c_n}{n}, \frac{w}{n}, \frac{d_n}{n} \right),
\]

where \( \xi_n = \sqrt{\lambda_n n} \left( \frac{a_n}{2 n M_p(2)} + \frac{b'_n}{2 n \lambda_n} - \frac{c_n}{pm} - \frac{d_n}{2 n} \right) \). The goal of this section is to obtain an upper bound on the Kolmogorov-Smirnov distance between \( Y_n \) and a Gaussian random variable \( G \) with
4.2. A bound for \( d_{KS}(\xi_n, G) \). In the final part we collect our bounds together to prove the result. Again using Lemma [2,3] to obtain the first inequality below, and the triangle inequality [2,4] for Kolmogorov-Smirnov distances to obtain the second, for any \( \epsilon > 0 \), we have

\[
d_{KS}(Y_n, G) \leq d_{KS}(\xi_n, G) + \mathbb{P}\left[ \frac{\sqrt{\lambda_n}W}{p\sqrt{n}} > \frac{\epsilon}{2} \right] + \mathbb{P}\left[ |\Psi\left( \frac{a_n}{nM_p(2)}, \frac{b_n'}{n\lambda_n}, \frac{c_n}{n}, \frac{W}{n}, \frac{d_n}{n} \right)| > \frac{\epsilon}{2\sqrt{n\lambda_n}} \right]
\]

\[
+ \frac{\epsilon}{\sqrt{2\pi w(\lambda)}}
\]

\[
\leq d_{KS}(\xi_n, G_n) + d_{KS}(G_n, G) + \mathbb{P}\left[ \frac{\sqrt{\lambda_n}W}{p\sqrt{n}} > \frac{\epsilon}{2} \right]
\]

\[
+ \mathbb{P}\left[ |\Psi\left( \frac{a_n}{nM_p(2)}, \frac{b_n'}{n\lambda_n}, \frac{c_n}{n}, \frac{W}{n}, \frac{d_n}{n} \right)| > \frac{\epsilon}{2\sqrt{n\lambda_n}} \right] + \frac{\epsilon}{\sqrt{2\pi J_p}},
\]

where \( G_n \) is a centered Gaussian random variable with the same variance as \( \xi_n \), and

\[
J_p := \inf_{s \in [0,1]} w(s) = \min\{\sigma^2(p,2), 3/4\}
\]

is positive.

The remainder of the proof is thus split into four parts. First we use the Berry-Esseen theorem to obtain an upper bound for \( d_{KS}(\xi_n, G_n) \). Then we use Proposition [2,6] to give an upper bound on \( d_{KS}(G_n, G) \). Then finally we take a choice of \( \epsilon = \epsilon_n \) to give an upper bound for the term involving \( \Psi \). In the final part we collect our bounds together to prove the result.

#### 4.2. A bound for \( d_{KS}(\xi_n, G_n) \)

A calculation establishes that \( \xi_n \) may be written as a sum \( \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \varphi_i^{(\lambda)} \) of independent and identically distributed centered random variables distributed like \( \varphi^{(\lambda)} \), where for each \( s \in (0,1] \), \( \varphi^{(s)} \) is the centered random variable

\[
\varphi^{(s)} := \sqrt{s} \left( \frac{Z^2 - M_p(2)}{2M_p(2)} + \frac{I^{(s)}g^2 - s}{2s} - \frac{|Z|^p - 1}{p} - \frac{g^2 - 1}{2} \right)
\]

where \( g \) is a standard Gaussian, \( Z \) is \( p \)-Gaussian, and \( I^{(s)} \) is Bernoulli distributed with \( \mathbb{P}[I^{(s)} = 1] = s = 1 - \mathbb{P}[I^{(s)} = 0] \).

We now take a moment to investigate the behavior of the moments of \( \varphi^{(s)} \) as \( s \) varies in \((0,1]\), with particular care for the case \( s \to 0 \). Considering first the second moment, a calculation shows that \( \mathbb{E}[|\varphi^{(s)}|^2] = w(s) \), where \( w(s) \) is given in (4.1). As for the third moment, we may write

\[
\varphi^{(s)} = \sqrt{s}W_p + \frac{1}{2\sqrt{s}}I^{(s)}g^2,
\]

where \( W_p \) is a random variable independent of \( I^{(s)} \) and whose distribution is independent of \( s \).

Using the definition of \( I^{(s)} \) and the independence to obtain the second equality below, we have

\[
\mathbb{E}[|\varphi^{(s)}|^3] = \frac{1}{s^{3/2}} \mathbb{E}\left[ \frac{|I^{(s)}g^2}{2} + sW_p \right]^3
\]

\[
= \frac{1}{s^{3/2}} \mathbb{E}\left[ \frac{g^2}{2} + sW_p \right]^3 + \frac{1}{s^{3/2}} (1 - s) \mathbb{E}[|sW_p|^3]
\]

\[
= \frac{1}{s} f_p(s)
\]

for a bounded function \( f_p : [0,1] \to [0,\infty) \). Since \( w(s) := \mathbb{E}[|\varphi^{(s)}|^2] \) is bounded below on \([0,1]\) by \( J_p > 0 \), and \( \sqrt{s} \mathbb{E}[|\varphi^{(s)}|^3] = f_p(s) \) is bounded above on \([0,1]\), there is a constant \( C_p \in (0,\infty) \).
such that
\[ C_p \geq \sup_{s \in [0,1]} \sqrt{s} \frac{\mathbb{E}[\| \varphi^{(s)} \|^2]}{\mathbb{E}[\| \varphi^{(s)} \|^2]}. \]

We are now ready to use the Berry-Esseen theorem. Recall that \( \xi_n := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \varphi_i^{(\lambda_n)} \) is a rescaled sum of independent and identically distributed random variables with law \( \varphi^{(\lambda_n)} \). If \( G_n \) is a Gaussian random variable with variance \( w(\lambda_n) \), then using the Berry-Esseen theorem to obtain the first inequality below, and the definition of \( C_p \) to obtain the second, we obtain
\[ d_{KS}(\xi_n, G_n) \leq K_{BE} \frac{\mathbb{E}[\| \varphi^{(\lambda_n)} \|^2]}{\mathbb{E}[\| \varphi^{(\lambda_n)} \|^2]} \frac{1}{\sqrt{n}} \leq K_{BE} \frac{C_p}{\sqrt{\lambda_n n}} \leq \frac{C'_p}{\sqrt{\lambda_n n}}, \]
where \( C'_p := K_{BE} C_p \in (0, \infty) \) depends only on \( p \).

4.3. A bound for \( d_{KS}(G_n, G) \). The random variables \( G_n \) and \( G \) are centered Gaussians with variances \( w(\lambda_n) \) and \( w(\lambda) \) respectively. Since \( w \) is Lipschitz on \([0, 1]\), and bounded below on \([0, 1]\) by a positive constant, it follows from Proposition 2.10 that there is a constant \( C_p \in (0, \infty) \) depending on \( p \) but not on \( n, \lambda_n, \lambda \) such that
\[ d_{KS}(G_n, G) \leq \frac{3}{8} \frac{|w(\lambda_n) - w(\lambda)|}{w(\lambda)} \leq C_p |\lambda_n - \lambda|. \]

4.4. Bounding the \( \epsilon \)-dependent terms. Let \( (\epsilon_n)_{n \geq 1} \) be a sequence of positive reals decreasing down to 0. By Lemma 2.10 and the triangle inequality, with \( v_n := \left( \frac{a_n}{n M_p(2)}, \frac{b'_n}{n \lambda_n}, \frac{c_n}{n}, \frac{W_n}{n}, \frac{d_n}{n} \right) \) we have
\[ \mathbb{P} \left[ \left| \Psi(v_n) \right| > \frac{\epsilon_n}{2 \sqrt{n \lambda_n}} \right] \leq \mathbb{P} \left[ \left| v_n \right|_2 \geq \sqrt{\frac{\epsilon_n}{2 M \sqrt{n \lambda_n}}} \right] + \mathbb{P} \left[ \left| v_n \right|_2 \geq \delta \right] \leq Q_n \left( \frac{a_n}{n M_p(2)} \right) + Q_n \left( \frac{b'_n}{n \lambda_n} \right) + Q_n \left( \frac{c_n}{n} \right) + Q_n \left( \frac{W_n}{n} \right) + Q_n \left( \frac{d_n}{n} \right), \]
where for a random variable \( X \),
\[ Q_n(X) := \mathbb{P} \left[ \left| X \right| > \frac{\delta}{\sqrt{n \lambda_n}} \right] + \mathbb{P} \left[ \left| X \right| > \frac{\epsilon_n}{10 M \sqrt{n \lambda_n}} \right]. \]

Considering first the terms involving \( a_n, b_n, \) and \( d_n \), we may use the bounds developed in Section 3.4. In particular there are sufficiently large constants \( C_p \in (0, \infty) \) and \( \Theta_p \in (0, \infty) \), such that by setting \( \epsilon_n = \Theta_p \frac{\log(k_n)}{\sqrt{k_n}} \), we ensure that
\[ Q_n \left( \frac{a_n}{n M_p(2)} \right) + Q_n \left( \frac{c_n}{n} \right) + Q_n \left( \frac{d_n}{n} \right) \leq \frac{C_p}{\sqrt{\lambda_n n}}, \]
for every \( n \in \mathbb{N} \).

Now we look to obtain a bound for \( Q_n \left( \frac{b'_n}{n \lambda_n} \right) \), which may be written
\[ Q_n \left( \frac{b'_n}{n \lambda_n} \right) = \mathbb{P} \left[ \left| \frac{b'_n}{\sqrt{n \lambda_n}} \right| > c_p \sqrt{n \lambda_n} \right] + \mathbb{P} \left[ \left| \frac{b'_n}{\sqrt{n \lambda_n}} \right| > c'_p \sqrt{\epsilon_n \sqrt{n \lambda_n}} \right], \]
where \( c_p := \delta/\sqrt{5} \) and \( c'_p := \frac{1}{\sqrt{10 M}} \). Recalling that \( b'_n := \sum_{i=1}^{n} \left( \varphi_i^{(\lambda_n)} g_i^2 - \lambda_n \right) \) is a sum of independent and identically distributed random variables, we see that we are in the setting of Lemma 2.9 with tails
\[ \alpha_n := c_p \sqrt{n \lambda_n} \quad \text{and} \quad \alpha'_n := c'_p \sqrt{\epsilon_n \sqrt{n \lambda_n}}. \]

In order to use Lemma 2.9 first we need to take a moment to study the moments of the random variables in \( b'_n \). Let \( I^{(s)} \) have the Bernoulli distribution with parameter \( s, g \) be a standard
Gaussian random variable, and define the random variable $\phi^{(s)} := \frac{1}{\sqrt{s}} (f^{(s)}g^2 - s)$. Then

$$S_n := \frac{b_n}{\sqrt{n}} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \phi_i^{(\lambda_n)},$$

(49)

where $\phi_i^{(\lambda_n)}$ are independent and identically distributed like $\phi^{(s)}$. Using the fact that $E[g^2] = 1, E[g^4] = 3, E[g^6] = 15$, it is straightforward to check that $\phi^{(s)}$ is centered and has absolute second and third moments given by

$$E[|\phi_s|^2] = (3 - s) \quad \text{and} \quad E[|\phi_s|^3] = s^{-3/2} \left(15s - 9s^2 + 4s^3 - 2s^4\right),$$

the latter of which is bounded above by $19/\sqrt{s}$ for $s \in [0, 1]$. Using the fact that $E[|\phi_s|^2] \geq 2$ on $[0, 1]$, we have the bound

$$\sup_{s \in [0,1]} \sqrt{s} \frac{E[|\phi_s|^3]}{E[|\phi_s|^2]^{3/2}} \leq \frac{19}{2^{3/2}}.$$

It follows that with the sum $S_n$ in (4.3), we are in the set up of Lemma 2.9 with

$$\gamma_n = \frac{19}{2^{3/2}} \frac{1}{\sqrt{\lambda_n}},$$

and $\sigma_{\max} = 3 \geq E[|\phi_{\lambda_n}|^2]$ for all $\lambda_n$ and $n \in \mathbb{N}$.

Now let $\beta_n$ and $\epsilon_n$ be sequences such that $n/\gamma_n^2 \geq \beta_n$ for every $n \in \mathbb{N}$, and additionally

$$c_p \sqrt{n \lambda_n} \geq \sigma_{\max} \sqrt{\log(\beta_n)} \quad \text{and} \quad c_p' \sqrt{\epsilon_n} \sqrt{n \lambda_n} \geq \sigma_{\max} \sqrt{\log(\beta_n)}$$

(50)

for every $n \in \mathbb{N}$. Then by Lemma 2.9 we have

$$Q_n \left( \frac{b_n}{n \lambda_n} \right) \leq \frac{4(K_{BE} + 1)}{\sqrt{\beta_n}}.$$

In particular, consider setting $\beta_n = \theta_p^\prime \lambda_n$ and $\epsilon_n = \Theta_p \log(\lambda_n)$. Then if $\theta_p^\prime$ is sufficiently small and $\Theta_p^\prime$ is sufficiently large such that $n/\gamma_n^2 \geq \beta_n$ and the inequalities (5.3) hold for every $n \in \mathbb{N}$, then for a sufficiently large constant $C_p^\prime \in (0, \infty)$ we have

$$Q_n \left( \frac{b_n}{n \lambda_n} \right) \leq \frac{C_p^\prime}{\sqrt{n \lambda_n} \beta_n^\prime}.$$

(51)

In particular, setting $\Theta_p^\prime := \max\{\Theta_p, \Theta_p^\prime\}$, both (4.3) and (4.4) hold, and there is a constant $C_p^\prime \in (0, \infty)$ such that

$$Q_n \left( \frac{a_n}{n M_p(2)} \right) + Q_n \left( \frac{c_n}{n} \right) + Q_n \left( \frac{d_n}{n} \right) + Q_n \left( \frac{b_n'}{n \lambda_n} \right) \leq \frac{C_p^\prime}{\sqrt{n \lambda_n}}.$$

(52)

With $\epsilon_n := \Theta_p \log(\lambda_n)$, by plugging (4.4) into (4.3) we obtain

$$\mathbb{P} \left[ \left| \frac{a_n}{n M_p(2)} \frac{b_n'}{n \lambda_n} \frac{c_n}{n} \frac{d_n}{n} \right| \geq \frac{\epsilon_n}{2 \sqrt{\lambda_n}} \right] \leq \frac{C_p^\prime}{\sqrt{n \lambda_n}} + \mathbb{P} \left[ W > \frac{n \delta}{\sqrt{5}} \right] + \mathbb{P} \left[ W > n \frac{\epsilon_n}{10 M \sqrt{\lambda_n}} \right].$$

(53)

4.5. Wrapping things together. Considering the original bound (4.1), and the bounds for the respective parts (4.2), (4.3), and (4.4). With the choice $\epsilon_n = \Theta_p \log(n)$, there is some constant $c_p \in (0, \infty)$ such that the terms involving tail probabilities for $W$ in (4.1) and (4.4) can be bounded by writing

$$\mathbb{P} \left[ \frac{\sqrt{\lambda_n} W}{p \sqrt{n}} > \frac{\epsilon_n}{2} \right] + \mathbb{P} \left[ W > \frac{n \delta}{\sqrt{5}} \right] + \mathbb{P} \left[ W > n \sqrt{\frac{\epsilon_n}{10 M \sqrt{\lambda_n}}} \right] \leq 3 \mathbb{P} \left[ W > c_p \frac{1}{\lambda_n} \log(n) \right],$$

where we used the fact that $W$ is non-negative.
In particular, combining (4.2), (4.3), and (4.4) in (4.1), it follows that there exist constants $c_p, C_p \in (0, \infty)$ depending only on $p$ such that

$$d_{KS}(Y_n, G) \leq C_p \max \left\{ \frac{\log(n)}{\sqrt{\lambda_n n}}, |\lambda_n - \lambda| \right\} + 3 \mathbb{P} \left[ W > c_p \frac{1}{\lambda_n} \log(n) \right],$$

completing the proof of Theorem C.

5. Proof of Theorem C

In this section we prove Theorem C which establishes Berry-Esseen type bounds for the $q$-norm of a vector-valued random variable with law $\mathbb{P}_{n,p,W}$.

5.1. Setting up the proof. Let $X_n$ have law $\mathbb{P}_{n,p,W}$. Then $X_n$ has a representation

$$X_n \overset{d}{=} \frac{Z}{(\sum_{i=1}^{n} |Z_i|^p + W)^{1/p}},$$

where $Z_1, \ldots, Z_n$ are independent and identically distributed $p$-Gaussian random variables, and $W$ is independent of the $Z_i$ with distribution $\mathcal{W}$. In particular, the $q$-norm of $X_n$ may be written

$$||X_n||_q \overset{d}{=} \frac{(\sum_{i=1}^{n} |Z_i|^q)^{1/q}}{(\sum_{i=1}^{n} |Z_i|^p + W)^{1/p}} = \frac{(nM_p(q) + \sum_{i=1}^{n} (|Z_i|^q - M_p(q)))^{1/q}}{(n + \sum_{i=1}^{n} (|Z_i|^p - 1) + W)^{1/p}} = \frac{(nM_p(q))^{1/q}}{n^{1/p}} F \left( \frac{\xi_1}{\sqrt{n}}, \frac{\xi_2}{\sqrt{n}}, W/n \right),$$

(54)

where $F(a, b, c) = \frac{(1+a)^{1/q}}{(1+b+c)^{1/p}}$, and $\xi_1^1$ and $\xi_2^1$ are given by

$$\xi_1^1 = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{|Z_i|^q - M_p(q)}{M_p(q)} \quad \text{and} \quad \xi_2^1 = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{|Z_i|^p - 1}{1}.$$

Since $F$ is twice differentiable in a neighborhood about the origin in $\mathbb{R}^3$, we may write

$$F(a, b, c) = 1 + \frac{1}{p} a - \frac{1}{q} (b + c) + \Psi(a, b, c),$$

(55)

where $\Psi : \mathbb{R}^3 \to \mathbb{R}$ is such that there exists $M, \delta \in (0, \infty)$ with

$$||\Psi(x)|| \leq M||x||_2^2 \quad \text{whenever} \quad ||x||_2 < \delta.$$

(56)

Combining (5.1) and (5.1), we have

$$Y_n := \sqrt{n} \left( n^{1/q-1/p} M_p(q)^{1/q} ||X_n||_q - 1 \right) = \xi_n - \frac{W}{\sqrt{n} \Psi} + \sqrt{n} \Psi \left( \frac{\xi_1^1}{\sqrt{n}}, \frac{\xi_2^1}{\sqrt{n}}, W/n \right),$$

where $\xi_n = \frac{1}{p} \xi_1^1 - \frac{1}{q} \xi_2^1$. Let $G$ be a Gaussian random variable with variance $\sigma^2(p, q)$. In light of Lemma 2.8 for any $\epsilon > 0$ we have the following bound

$$d_{KS}(Y_n, G) \leq d_{KS}(\xi_n, G) + \mathbb{P} \left[ \frac{W}{n} > \epsilon/2 \right] + \mathbb{P} \left[ \left| \sqrt{n} \Psi \left( \frac{\xi_1^1}{\sqrt{n}}, \frac{\xi_2^1}{\sqrt{n}}, \frac{W}{n} \right) \right| > \epsilon/2 \right] + \frac{\epsilon}{\sqrt{2\pi} \sigma^2(p, q)}.$$ 

(57)

To give a brief overview of the remainder of the proof, first we use the Berry-Esseen theorem to show that the random variable $\xi_n$ is close to a Gaussian random variable with variance $\sigma^2(p, q)$ in terms of Kolmogorov-Smirnov distances, obtaining an upper bound for $d_{KS}(\xi_n, G)$. Then we obtain an upper bound for remaining terms.
5.2. A bound for \(d_{KS}(\xi_n, G)\). First of all, we note the random variable \(\xi_n\) may be written \(\xi_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_i\), where

\[
Z_i := \frac{|Z_i|^q - M_p(q)}{qM_p(q)} - \frac{|Z_i|^p - 1}{p}, \quad i \in \{1, \ldots, n\}.
\]

Each \(Z_i\) has variance \(\sigma^2(p,q)\) as in (1.1). In particular, \(\xi_n\) also has variance \(\sigma^2(p,q)\). Let \(\rho(p,q)\) denote the third moment of \(|Z_i|\). Then by the Berry-Esseen theorem (see Theorem 2.7), if \(G\) is a centered Gaussian random variable with variance \(\sigma^2(p,q)\), then we have

\[
d_{KS}(\xi_n, G) \leq K_{BE} \frac{\rho(p,q)}{\sigma^3(p,q)\sqrt{n}} =: C_{p,q} \sqrt{n}.
\]

(58)

5.3. A bound for the \(\Psi\) term. Now consider obtaining an upper bound for the term involving \(\Psi\) in (5.1). With \(M\) and \(\delta\) as in (5.1), using Lemma 2.9 and the triangle inequality, we have

\[
P \left( \sqrt{n} \Psi \left( \frac{\xi_n}{\sqrt{n}}, \frac{\xi_n^2}{\sqrt{n}}, \frac{W}{n} \right) > \frac{\epsilon_n}{2} \right) \leq Q_n \left( \frac{\epsilon_n^1}{\sqrt{n}} \right) + Q_n \left( \frac{\epsilon_n^2}{\sqrt{n}} \right) + Q_n \left( \frac{W}{n} \right),
\]

(59)

where for real-valued random variables \(X\),

\[
Q_n(X) := P \left[ |X| > \sqrt{\frac{\epsilon_n}{6M\sqrt{n}}} \right] + P \left[ |X| > \frac{\delta}{\sqrt{3}} \right] .
\]

Let us start with an upper bound for

\[
Q_n \left( \frac{\epsilon_n^1}{\sqrt{n}} \right) = P \left[ |\xi_n^1| > \sqrt{\frac{\epsilon_n^1\sqrt{n}}{6M}} \right] + P \left[ |\xi_n^1| > \frac{\epsilon_n^1\sqrt{n}}{\sqrt{3}} \right].
\]

We may use Lemma 2.9 to bound the tail probabilities of \(\xi_n^1\). To set this up, let \(m = n, \beta_n = \theta n\) for some \(\theta \in (0,1]\),

\[
\gamma_n := \frac{\mathbb{E} \left[ |Z_i|^q - M_p(q) | M_p(q) \right]^3}{\mathbb{E} \left[ |Z_i|^q - M_p(q) | M_p(q) \right]^{3/2}}
\]

for every \(n \in \mathbb{N}\), and let \(\sigma_{\max}\) be the variance of \(|Z_i|^q - M_p(q) | M_p(q)\). In particular, whenever \(\epsilon_n\) and \(\beta_n \leq n/\gamma_n^2\) are sequences such that

\[
\sqrt{\frac{\epsilon_n\sqrt{n}}{6M}} \geq \sigma_{\max} \sqrt{\log(\beta_n)} \quad \text{and} \quad \frac{\delta \sqrt{n}}{\sqrt{3}} \geq \sigma_{\max} \sqrt{\log(\beta_n)}
\]

(60)

for every \(n \in \mathbb{N}\), then

\[
Q_n \left( \frac{\epsilon_n^1}{\sqrt{n}} \right) \leq \frac{4(K_{BE} + 1)}{\sqrt{\beta_n}}.
\]

In particular, consider setting \(\beta_n = \theta_{p,q} n\), and \(\epsilon_n := \Theta_{p,q} \log(n) / \sqrt{n}\), where \(\theta_{p,q}\) is small enough and \(\Theta_{p,q}\) is large enough so that both the inequalities in (5.3) hold for every \(n \in \mathbb{N}\). Then there is a constant \(C_{p,q} \in (0, \infty)\) such that with \(\epsilon_n := \Theta_{p,q} \log(n) / \sqrt{n}\), we have

\[
Q_n \left( \frac{\epsilon_n^1}{\sqrt{n}} \right) \leq \frac{C_{p,q}}{\sqrt{n}}
\]

for every \(n \in \mathbb{N}\). Making an identical argument with the term \(Q_n \left( \frac{\epsilon_n^2}{\sqrt{n}} \right)\), it follows that there are sufficiently large constants \(\Theta'_{p,q}\) and \(C'_{p,q}\) such that taking \(\epsilon_n := \Theta'_{p,q} \log(n) / \sqrt{n}\), we have

\[
Q_n \left( \frac{\epsilon_n^1}{\sqrt{n}} \right) + Q_n \left( \frac{\epsilon_n^2}{\sqrt{n}} \right) \leq \frac{C'_{p,q}}{\sqrt{n}}
\]
for every $n \in \mathbb{N}$. Hence, by (5.3) and the definition of $Q_n(W)$, with $\epsilon_n := \Theta_{p,q}' \frac{\log(n)}{\sqrt{n}}$, we have

$$
\begin{align*}
\mathbb{P} \left[ \sqrt{n} \Psi \left( \frac{e_n}{\sqrt{n}}, \frac{W}{n} \right) > \frac{\epsilon_n}{2} \right] &\leq \frac{C_{p,q}'}{\sqrt{n}} + \mathbb{P} \left[ W > \frac{\delta n}{\sqrt{3}} \right] + \mathbb{P} \left[ W > n \sqrt{\frac{e_n}{6M\sqrt{n}}} \right] + \mathbb{P} \left[ \frac{W}{n} > \frac{\epsilon_n}{2} \right] + \frac{\epsilon_n}{2 \sqrt{2\pi \sigma^2(p,q)}}.
\end{align*}
$$

(61)

5.4. Wrapping things together. Recall our original bound (5.1) for $d_{KS}(Y_n, G)$. Plugging the intermediate bounds (5.2) and (5.3) into (5.1), we recall that there are constants $\Theta_{p,q}', C_{p,q}', C_{p,q}'' \in (0, \infty)$ such that with $\epsilon_n := \Theta_{p,q}' \frac{\log(n)}{\sqrt{n}}$ we have

$$
d_{KS}(Y_n, G) \leq \frac{C_{p,q}'}{\sqrt{n}} + \frac{C_{p,q}''}{\sqrt{n}} + \mathbb{P} \left[ W > \frac{\delta n}{\sqrt{3}} \right] + \mathbb{P} \left[ W > n \sqrt{\frac{e_n}{6M\sqrt{n}}} \right] + \mathbb{P} \left[ \frac{W}{n} > \frac{\epsilon_n}{2} \right] + \frac{\epsilon_n}{2 \sqrt{2\pi \sigma^2(p,q)}}.
$$

For a sufficiently small constant $c_{p,q} \in (0, \infty)$, the terms in (61) involving tail probabilities for $W$ may be collected by the simple bound

$$
\mathbb{P} \left[ W > \frac{\delta n}{\sqrt{3}} \right] + \mathbb{P} \left[ W > n \sqrt{\frac{e_n}{6M\sqrt{n}}} \right] + \mathbb{P} \left[ \frac{W}{n} > \frac{\epsilon_n}{2} \right] \leq \mathbb{P} \left[ W > c_{p,q} \sqrt{n \log(n)} \right],
$$

(62)

and it follows from plugging (5.4) and the definition $\epsilon_n := \Theta_{p,q}' \frac{\log(n)}{\sqrt{n}}$ into (4.3) that with $c_{p,q}$ as above, for a sufficiently large $C_{p,q}'' \in (0, \infty)$ we have

$$
d_{KS}(Y_n, G) \leq c_{p,q} \frac{\log(n)}{\sqrt{n}} + \mathbb{P} \left[ W > c_{p,q} \sqrt{n \log(n)} \right],
$$

which completes the proof of Theorem C.

Acknowledgment. SJ and JP have been supported by the Austrian Science Fund (FWF) Project P32405 “Asymptotic Geometric Analysis and Applications” of which JP is principal investigator. JP has also been supported by a Visiting International Professor Fellowship from the Ruhr University Bochum and its Research School PLUS.

References

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

[2] D. Alonso-Gutiérrez, J. Prochno, and C. Thäle. Gaussian fluctuations for high-dimensional random projections of $\ell_p^n$-balls. *Bernoulli*, 25(4A):3139–3174, 2019.

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

[4] A. C. Berry. The accuracy of the Gaussian approximation to the sum of independent variates. *Trans. Amer. Math. Soc.*, 49:122–136, 1941.

[5] U. Brehm and J. Voigt. Asymptotics of cross sections for convex bodies. *Beiträge Algebra Geom.*, 41(2):437–445, 2000.

[6] P. Diaconis and D. Freedman. A dozen de Finetti-style results in search of a theory. *Ann. Inst. H. Poincaré Probab. Statist.*, 23(2, suppl.):397–423, 1987.

[7] J. Grote, Z. Kabluchko, and C. Thäle. Limit theorems for random simplices in high dimensions. *ALEA Lat. Am. J. Probab. Math. Stat.*, 16(1):141–177, 2019.

[8] Z. Kabluchko, A.E. Litvak, and D. Zaporozhets. Mean width of regular polytopes and expected maxima of correlated Gaussian variables. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 442(Veroyatnost’ i Statistika, 23):75–96, 2015. Reprinted in *J. Math. Sci. (N.Y.)* 225 (2017), no. 5, 770–787.

[9] Z. Kabluchko, J. Prochno, and C. Thäle. High-dimensional limit theorems for random vectors in $\ell_p^n$-balls. *Commun. Contemp. Math.*, 21(1):1750092, 30, 2019.

[10] Z. Kabluchko, J. Prochno, and C. Thäle. High-dimensional limit theorems for random vectors in $\ell_p^n$-balls. ii. *Commun. Contemp. Math.* (to appear), 2019+.

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

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

[13] G. Paouris, P. Pivovarov, and J. Zinn. A central limit theorem for projections of the cube. *Probab. Theory Related Fields*, 159(3-4):701–719, 2014.
[14] I.G. Shevtsova. On the absolute constants in the berry-esseen inequality and its structural and nonuniform improvements. Inform. Primen., 7(1):124–125, 2013.
[15] A. J. Stam. Limit theorems for uniform distributions on spheres in high-dimensional Euclidean spaces. J. Appl. Probab., 19(1):221–228, 1982.

Institut für Mathematik & Wissenschaftliches Rechnen, University of Graz, Heinrichstrasse 36, 8010 Graz, Austria
E-mail address: samuel.johnston@uni-graz.at

Institut für Mathematik & Wissenschaftliches Rechnen, University of Graz, Heinrichstrasse 36, 8010 Graz, Austria
E-mail address: joscha.prochno@uni-graz.at