ASYMPTOTIC NORMALITY FOR RANDOM SIMPLICES AND CONVEX BODIES IN HIGH DIMENSIONS

D. ALONSO-GUTIÉRREZ, F. BESAU, J. GROTE, Z. KABLUCHKO, M. REITZNER, C. THÄLE, B.-H. VRITSIOU, AND E. WERNER

Abstract. Central limit theorems for the log-volume of a class of random convex bodies in $\mathbb{R}^n$ are obtained in the high-dimensional regime, that is, as $n \to \infty$. In particular, the case of random simplices pinned at the origin and simplices where all vertices are generated at random is investigated. The coordinates of the generating vectors are assumed to be independent and identically distributed with subexponential tails. In addition, asymptotic normality is established also for random convex bodies (including random simplices pinned at the origin) when the spanning vectors are distributed according to a radially symmetric probability measure on the n-dimensional ℓ_p-ball. In particular, this includes the cone and the uniform probability measure.

1. Introduction and main results

1.1. **Motivation.** Central limit theorems for random polytopes in $\mathbb{R}^n$ are widely known if the space dimension n is kept fixed, while the number of generating points tends to infinity. We refer, for example, to the survey articles of Bárány [3], Hug [8] and Reitzner [17] for results in this direction and for further references. In the present paper we investigate the case where the number of generating points is essentially equal to the space dimension and both tend to infinity *simultaneously*. To be more precise, we consider the case of random n-dimensional simplices in $\mathbb{R}^n$, where we distinguish between the case of $(n+1)$ generating points chosen at random, or the case where we only have n random points and the $(n+1)$st vertex is fixed at the origin. The latter construction is called a *pinned simplex* in the following. Asymptotic normality for the log-volume of random simplices in high dimensions has previously been considered by Ruben [18], Maehara [11] and Mathai [10]. Note however, that in their results the dimension of the random (pinned) simplices is kept fixed, while the space dimension n tends to infinity. For Gaussian and so-called beta simplices Eichelsbacher and Knichel [6] and Grote, Kabluchko and Thäle [7] recently also studied a number of probabilistic limit theorems where the simplex dimension tends to infinity as well.

Our main result is a central limit theorem for the log-volume of a random n-dimensional simplex in $\mathbb{R}^n$, see Theorem 1.1 below. For an n-dimensional random simplex that is pinned, it is known that its volume is determined by the absolute value of the determinant of the matrix whose columns are given by the generating vectors from the origin. As a consequence, if these columns are filled by independent

2010 *Mathematics Subject Classification.* 52A22, 52A23, 60D05, 60F05.
Key words and phrases. central limit theorem, high dimensions, ℓ_p-ball, random convex body, random determinant, random parallelotope, random polytope, random simplex, stochastic geometry.

and identically distributed (i.i.d.) random variables, a central limit theorem for the log-volume of random pinned simplex follows from the central limit theorem for random determinants with i.i.d. entries established by Nguyen and Vu [13].

The same arguments cannot directly be applied if the coordinates of the generating points are not independent, that is, for example if the points are chosen with respect to a probability measure in the ℓ_p-ball $B_p^n \subset \mathbb{R}^n$ with $p \neq \infty$. However, we still succeed in establishing a central limit theorem for random pinned simplices in the ℓ_p-ball for certain radially symmetric probability measures, which include in particular the uniform probability measure and the cone-volume measure, see Theorem 1.3 below. In our proof we employ different tools, most notably a Schechtman-Zinn-type probabilistic representation of Barthe, Guédon, Mendelson and Naor [4], which allow us to relate the log-volume of the random pinned simplex to the determinant of a matrix with independent entries. Hence, although the coordinates of the generating vectors are now no longer independent, at the core of our argument we can still rely on the central limit theorem for the determinant of random matrices with i.i.d. entries.

1.2. **Main results: the case of independent coordinates.** Let μ be a probability measure on $\mathbb{R}^n$ $(n \geq 1)$ and let $X_0, \ldots, X_n$ be independent random vectors distributed according to μ. We define the random simplex

$$\Sigma_n := \operatorname{conv}\big(\{X_0, X_1, \ldots, X_n\}\big),$$

as well as the random pinned simplex

$$\Sigma_n^0 := \operatorname{conv}\big(\{0, X_1, \ldots, X_n\}\big).$$

In what follows, we shall focus on the case where the coordinates ξ_j^i, $j \in \{1, \ldots, n\}$, of the random vectors $X_i = (\xi_1^i, \ldots, \xi_n^i)$ are independent copies of a random variable ξ. Furthermore, we assume that the random variable ξ is symmetric, has variance one and subexponential tails with exponent $\alpha > 0$. By the latter we mean that there are constants $c_1, c_2 > 0$ such that

$$\mathbb{P}\big(|\xi| \geq t^\alpha\big) \leq c_1 e^{-c_2 t}, \qquad t > 0. \tag{1}$$

Examples are the uniform distribution on the cube $[-\sqrt{3}, \sqrt{3}]^n$, the uniform distribution on the discrete cube $\{-1, +1\}^n$, the two-sided exponential distribution, standard Gaussian distribution on $\mathbb{R}^n$ or, more generally, the p-generalized Gaussian distribution with density proportional to $e^{-|t|^p/a}$ (for an appropriate choice of $a > 0$) for any $p > 0$.

In our first result we establish a central limit theorem for the log-volume of the high-dimensional random simplices Σ_n and Σ_n^0, as $n \to \infty$. In the following $Z \sim \mathcal{N}(0, 1)$ always denotes a standard Gaussian random variable and $\xrightarrow{d}$ indicates convergence in distribution.

Theorem 1.1 (CLT for random simplices)**.** *Let ξ be a symmetric random variable with variance one and subexponential tails with exponent $\alpha > 0$.*

i) *Assume $\Sigma_n := \operatorname{conv}\big(\{X_0, X_1, \ldots, X_n\}\big)$ is a random simplex in $\mathbb{R}^n$ with X_i having i.i.d. coordinates $\xi_j^i \sim \xi$. Then*

$$\frac{\ln \operatorname{vol}_n(\Sigma_n) + \frac{n}{2}\ln n - \frac{n}{2} + \frac{1}{4}\ln n}{\sqrt{\frac{1}{2}\ln n}} \xrightarrow{d} Z, \qquad \text{as } n \to \infty.$$

ii) *Assume* $\Sigma_n^0 := \operatorname{conv}\big(\{0, X_1, \ldots, X_n\}\big)$ *is a random simplex in* $\mathbb{R}^n$ *with* X_i *having i.i.d. coordinates* $\xi_j^i \sim \xi$. *Then*

$$\frac{\ln \operatorname{vol}_n(\Sigma_n^0) + \frac{n}{2}\ln n - \frac{n}{2} + \frac{3}{4}\ln n}{\sqrt{\frac{1}{2}\ln n}} \xrightarrow{d} Z, \qquad \textit{as } n \to \infty.$$

Part ii) of Theorem 1.1 can be reformulated for the parallelotope spanned by the vertices of the pinned simplex from the origin. Even more generally, for a convex body $K \subset \mathbb{R}^n$, that is a compact convex subset with non-empty interior, and n random points $X_1, \ldots, X_n \in \mathbb{R}^n$ we define the random convex body

$$\Xi_n(K) := K[X_1, \ldots, X_n] = \left\{ \sum_{i=1}^n y_i X_i : (y_1, \ldots, y_n) \in K \right\}. \tag{2}$$

We note that for each $n \in \mathbb{N}$ and each convex body $K \subset \mathbb{R}^n$, $\Xi_n(K)$ is a random closed set in the usual sense of stochastic geometry, cf. [9, Chapter 16]. In particular, this implies that that the volume $\operatorname{vol}_n(\Xi_n(K))$ of $\Xi_n(K)$ is an ordinary random variable. As observed by Paouris and Pivovarov [15, 16], this concept generalizes a number of common constructions. Namely,

a) if K is the standard simplex

$$T^n = \left\{ (x_1, \ldots, x_n) \in \mathbb{R}^n : x_i \geq 0 \text{ and } \sum_{i=1}^n x_i \leq 1 \right\},$$

then $\Xi_n(T^n)$ coincides with the pinned simplex Σ_n^0.

b) if K is the unit cube $C^n = [0,1]^n$, then $\Xi_n(C^n)$ is the parallelotope spanned by $(X_i)_{i=1}^n$ from the origin.

c) if $K = B_\infty^n = [-1,1]^n$ is the symmetric cube, then $\Xi_n(B_\infty^n)$ is the zonotope generated by the segments $[-X_i, X_i]$, i.e.,

$$\Xi_n(B_\infty^n) = \left\{ \sum_{i=1}^n \lambda_i X_i : \lambda_i \in [-1,1] \right\}.$$

d) if $K = B_1^n$ is the cross-polytope, then $\Xi_n(B_1^n)$ is the symmetric convex hull of the $2n$ points $\{\pm X_i : i = 1, \ldots, n\}$.

e) if $K = B_2^n$ is the unit ball, then $\Xi_n(B_2^n)$ is an ellipsoid, that is, it is the image of the unit ball under the linear map whose matrix is generated by the random points $(X_i)_{i=1}^n$.

As a generalization of part ii) of Theorem 1.1 we obtain the following central limit theorem for the log-volume of the random convex bodies $\Xi_n(K)$. In the following we denote by $\operatorname{dist_K}(-,-)$ the Kolmogorov distance between random variables, that is, for two random variables X, Y we have

$$\operatorname{dist_K}(X,Y) := \sup_{t \in \mathbb{R}} \big| \mathbb{P}(X \leq t) - \mathbb{P}(Y \leq t) \big|.$$

Note that convergence in the Kolmogorov distance implies convergence in distribution. Also, by $o(1)$ we denote some sequence $(a_n)_{n \in \mathbb{N}}$ with $a_n \to 0$, as $n \to \infty$.

K_n	$\mathrm{vol}_n(K_n)$	central limit theorem
T^n	$\frac{1}{n!}$	$\frac{\ln \mathrm{vol}_n(\Xi_n(T^n))+\frac{n}{2}\ln n-\frac{n}{2}+\frac{3}{4}\ln n}{\sqrt{\frac{1}{2}\ln n}} \xrightarrow{d} Z$
C^n	1	$\frac{\ln \mathrm{vol}_n(\Xi_n(C^n))-\frac{n}{2}\ln n+\frac{n}{2}+\frac{1}{4}\ln n}{\sqrt{\frac{1}{2}\ln n}} \xrightarrow{d} Z$
B^n_∞	2^n	$\frac{\ln \mathrm{vol}_n(\Xi_n(B^n_\infty))-\frac{n}{2}\ln n-(\ln 2-\frac{1}{2})n+\frac{1}{4}\ln n}{\sqrt{\frac{1}{2}\ln n}} \xrightarrow{d} Z$
B^∞_1	$\frac{2^n}{n!}$	$\frac{\ln \mathrm{vol}_n(\Xi_n(B^n_1))+\frac{n}{2}\ln -(\ln 2+\frac{1}{2})n+\frac{3}{4}\ln n}{\sqrt{\frac{1}{2}\ln n}} \xrightarrow{d} Z$

TABLE 1. Special cases of the central limit theorem for the log-volume of random covex bodies (Theorem 1.2). Here, Z is a standard Gaussian random variable.

Theorem 1.2. *Let ξ be a symmetric random variable with variance one and subexponential tails with exponent $\alpha > 0$. Let $(K_n)_{n\in\mathbb{N}}$ be a sequence of convex bodies such that $K_n \subset \mathbb{R}^n$. If $X_1, \dots, X_n$ are random points in $\mathbb{R}^n$ with i.i.d. coordinates $\xi^i_j \sim \xi$, and $\Xi_n(K_n)$ is the random convex body as defined by* (2), *then*

$$S_n := \frac{\ln \mathrm{vol}_n(\Xi_n(K_n)) - \ln \mathrm{vol}_n(K_n) - \frac{n}{2}\ln n + \frac{n}{2} + \frac{1}{4}\ln n}{\sqrt{\frac{1}{2}\ln n}} \xrightarrow{d} Z, \qquad \text{as } n \to \infty.$$

More precisely, we have that

$$\mathrm{dist}_{\mathrm{K}}(S_n, Z) \le (\ln n)^{-\frac{1}{3}+o(1)}.$$

As an application of Theorem 1.2 we may revisit the special cases a) – d) of the random convex bodies $\Xi_n(K_n)$ mentioned above. The resulting central limit theorems are summarized in Table 1. In particular, taking $K_n = T^n$ for each $n \in \mathbb{N}$, we also obtain part ii) of Theorem 1.1.

1.3. **Main results: the case of ℓ_p-balls.** For $0 < p \le \infty$ the n-dimensional ℓ_p-ball $B^n_p \subset \mathbb{R}^n$ is defined as

$$B^n_p := \{x \in \mathbb{R}^n : \|x\|_p \le 1\},$$

where the p-norm (or quasi-norm if $0 < p < 1$) of $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ is

$$\|x\|_p := \begin{cases} \left(\sum\limits_{i=1}^n |x_i|^p\right)^{1/p} & \text{if } 0 < p < \infty, \\ \max\{|x_1|, \dots, |x_n|\} & \text{if } p = \infty. \end{cases}$$

In our next result we consider pinned simplices, denoted by $\Sigma^0_n(\nu)$, which are spanned by the origin and n points $X_1, \dots, X_n$ chosen at random with respect to a radially symmetric probability measure $\nu = \nu_n(m_n, p)$ on the n-dimensional ℓ_p-ball B^n_p. More specifically, ν belongs to a family of measures including the cone probability measure and the uniform probability measure on B^n_p which is driven by a parameter $m_n \ge 0$. This model contains a number of special cases that are of particular interest (see Theorem 1, Theorem 3, Corollary 3 and Corollary 4 in

[4] as well as the discussion before Theorem 1.1 in [2]). Namely, if $m_n = 0$, then the random points $X_1, \dots, X_n$ are distributed according to the cone probability measure on the boundary of B_p^n, i.e., the ℓ_p-sphere in $\mathbb{R}^n$. It is well known that this measure coincides with the normalized surface measure precisely if $p \in \{1, 2, \infty\}$. Next, if $m_n = 1$, then $X_1, \dots, X_n$ are selected according to the uniform distribution on B_p^n. Finally, if $m_n = m/p$ for some $m \in \mathbb{N}$, then the distribution corresponds to the image of the cone probability measure on B_p^{n+m} under the orthogonal projection onto the first n coordinates. Similarly, if $m_n = 1 + m/p$, then the points are sampled according to the image of the uniform distribution on B_p^{n+m} under the same projection. We refer to Section 3 for the precise construction of ν.

Theorem 1.3 (CLT for random convex bodies in the ℓ_p-ball)**.** *Let $X_1, \dots, X_n$ be n independent random points in the ℓ_p-ball B_p^n with respect to a probability measure $\nu = \nu_n(m_n, p)$ as defined in Section 3.1. Let $(K_n)_{n\in\mathbb{N}}$ be a sequence of convex bodies such that $K_n \subset \mathbb{R}^n$ and $\Xi_n(K_n)$ be the random convex body generated by ν-distributed random points $X_1, \dots, X_n$ as defined by (2). Then*

$$S_n := \frac{\ln \mathrm{vol}_n(\Xi_n(K_n)) - \ln \mathrm{vol}_n(K_n) - \frac{1}{2}\ln (n-1)! + \frac{n}{p}\ln(a(m_n + \frac{n}{p}))}{\sqrt{\frac{1}{2}\ln n}} \xrightarrow{d} Z,$$

as $n \to \infty$, where $a = (\Gamma(1/p)/\Gamma(3/p))^{p/2}$. More precisely, we have that

$$\mathrm{dist}_{\mathrm{K}}(S_n, Z) \le (\ln n)^{-\frac{1}{3}+o(1)}.$$

By choosing for each $n \in \mathbb{N}$, K_n as the standard simplex T^n we obtain the following central limit theorem as a direct corollary to Theorem 1.3.

Corollary 1.4 (CLT for random pinned simplices in the ℓ_p-ball)**.** *Let $\Sigma_n^0(\nu)$ be the random pinned simplex that is spanned by the origin and n independent random points in the ℓ_p-ball B_p^n which are distributed according to a probability measure $\nu = \nu_n(m_n, p)$ as defined in Section 3.1. Then*

$$\frac{\ln \mathrm{vol}_n(\Sigma_n^0(\nu)) + \frac{n}{2}\ln n - \frac{n}{2} + \frac{3}{4}\ln n + \frac{n}{p}\ln(a(m_n + \frac{n}{p}))}{\sqrt{\frac{1}{2}\ln n}} \xrightarrow{d} Z, \qquad \text{as } n \to \infty,$$

where $a = (\Gamma(1/p)/\Gamma(3/p))^{p/2}$.

1.4. **Plan of the paper.** In the next Section we outline the proof of Theorem 1.1 and of Theorem 1.2. We will collect the relevant tools along the way. As mentioned in the introduction, the proof of Theorem 1.1 essentially relies on the central limit theorem for determinants of random matrices of Nguyen and Vu [13] with additional arguments for the non-pinned case. In Section 3 we present the details of the proof of Theorem 1.3 and recall the definition of the special measure $\nu = \nu(m_n, p)$. For the proof of Theorem 1.3 we especially need the Schechtman-Zinn-type probabilistic representation from [4].

Acknowledgment. The authors started this project within a working group that formed during the Mini-Workshop *Perspectives in High-dimensional Probability and Convexity* at the Mathematisches Forschungsinstitut Oberwolfach (MFO). All support is gratefully acknowledged.

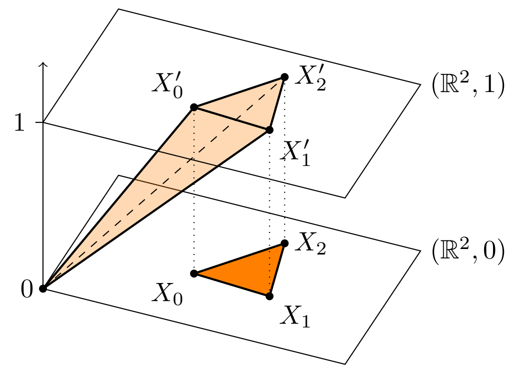

FIGURE 1. Illustration of the identification of the random simplex Σ_n with a pinned simplex Σ_{n+1}^0 in the two dimensional case.

D.A.-G. is partially supported by DGA grant `E26_17R` and MINECO grant `MTM2016-77710-P` and IUMA. F.B. was partially supported by the Deutsche Forschungsgemeinschaft (DFG) grant `BE 2484/5-2`. J.G. was supported by DFG via `RTG 2131` "High-dimensional Phenomena in Probability – Fluctuations and Discontinuity". E.W. was partially supported by NSF grant `DMS-1811146`.

2. Proof of Theorem 1.1 and Theorem 1.2

2.1. **Proof of Theorem 1.2 and part ii) of Theorem 1.1.** Let us first recall the central limit theorem of Nguyen and Vu for the log-determinant of random matrices with independent entries.

Theorem 2.1 ([13, Theorem 1.1]). *Let A_n be an $n \times n$ random matrix whose entries are independent random variables with zero mean, variance one and subexponential tails with exponent $\alpha > 0$. Further, let $Z \sim \mathcal{N}(0, 1)$ be a standard Gaussian random variable. Then,*

$$Z_n := \frac{\ln|\det A_n| - \frac{1}{2}\ln(n-1)!}{\sqrt{\frac{1}{2}\ln n}} \xrightarrow{d} Z, \quad \text{as } n \to \infty.$$

More precisely, the rate of convergence is

$$\operatorname{dist}_{\mathrm{K}}(Z_n, Z) \le (\ln n)^{-1/3+o(1)}, \tag{3}$$

for all n large enough.

Next, we recall the definition (2) of the random convex bodies $\Xi_n(K_n)$ and observe that

$$\operatorname{vol}_n(\Xi_n(K_n)) = |\det(X_1|\dots|X_n)| \operatorname{vol}_n(K_n), \tag{4}$$

see, for example, [16, Proposition 2.1]. This implies that

$$\frac{\ln \operatorname{vol}_n(\Xi_n(K_n)) - \ln \operatorname{vol}_n(K_n) - \frac{1}{2}\ln(n-1)!}{\sqrt{\frac{1}{2}\ln n}}$$
$$= \frac{\ln|\det(X_1|\dots|X_n)| - \frac{1}{2}\ln(n-1)!}{\sqrt{\frac{1}{2}\ln n}}$$

and hence Theorem 1.2 is a direct consequence of Theorem 2.1. Moreover, taking $K_n = T^n$ for each $n \in \mathbb{N}$ and recalling that $\Sigma_n^0 = \Xi_n(T^n)$, we may derive part ii) in Theorem 1.1 by Stirling's formula,

$$\ln n! = n \ln n - n + \frac{1}{2} \ln n + O(1).$$

2.2. **Proof of part i) in Theorem 1.1.** We now prove the central limit theorem for the log-volume of random simplices Σ_n. First, notice that the volume of Σ_n can be identified with the volume of a pinned simplex in $\mathbb{R}^{n+1}$ by the classical projective construction (see also Figure 1). Given $n+1$ points $X_0, X_1, \dots, X_n$ in $\mathbb{R}^n$ we consider the points $X_i' := (X_i, 1) \in \mathbb{R}^{n+1}$ for $i = 0, \dots, n$. We set $\Sigma_{n+1}^0 := \operatorname{conv}(\{0, X_0', \dots, X_n'\})$ and find that

$$\operatorname{vol}_n(\Sigma_n) = (n+1) \operatorname{vol}_{n+1}(\Sigma_{n+1}^0) = \frac{1}{n!} |\det(X_0'| \dots |X_n')|.$$

Now let $Y_i \in \mathbb{R}^{n+1}$ be the vector whose entries are given by the ith row of the $(n+1) \times (n+1)$ matrix $(X_0'| \dots |X_n')$ for $i = 1, \dots, n$. Then Y_i is a random vector in $\mathbb{R}^{n+1}$ with independent coordinates distributed like ξ. Notice that the last row in the matrix $(X_0'| \dots |X_n')$ is just the constant vector $U_{n+1} := (1, \dots, 1) \in \mathbb{R}^{n+1}$. Let Y_{n+1} be another random vector with independent entries distributed like ξ. We compare the random matrix $(Y_1| \dots |Y_n|U_{n+1})$ with the random matrix $(Y_1| \dots |Y_n|Y_{n+1})$. Notice that the latter now has independent and identically distributed entries. We have

$$\frac{1}{n!} |\det(Y_1| \dots |Y_n|U_{n+1})| = \operatorname{dist}(U_{n+1}, L_n) \operatorname{vol}_n(\operatorname{conv}\{0, Y_1, \dots, Y_n\})$$

and

$$\frac{1}{n!} |\det(Y_1| \dots |Y_n|Y_{n+1})| = \operatorname{dist}(Y_{n+1}, L_n) \operatorname{vol}_n(\operatorname{conv}\{0, Y_1, \dots, Y_n\}),$$

where L_n is the n-dimensional linear subspace spanned by $Y_1, \dots, Y_n$ in $\mathbb{R}^{n+1}$ and $\operatorname{dist}(v, L_n)$ denotes the distance of a vector $v \in \mathbb{R}^{n+1}$ to L_n. Collecting all of the above we conclude that

$$\begin{aligned} \ln \operatorname{vol}_n(\Sigma_n) &= \ln |\det(X_0'| \dots |X_n')| - \ln n! \\ &= \ln |\det(Y_1| \dots |Y_{n+1})| - \ln \operatorname{dist}(Y_{n+1}, L_n) + \ln \operatorname{dist}(U_{n+1}, L_n) - \ln n!. \end{aligned}$$

We will show that

$$(5) \qquad \frac{\ln \operatorname{dist}(U_{n+1}, L_n)}{\sqrt{\frac{1}{2} \ln n}} \xrightarrow{d} 0 \qquad \text{and} \qquad \frac{\ln \operatorname{dist}(Y_{n+1}, L_n)}{\sqrt{\frac{1}{2} \ln n}} \xrightarrow{d} 0,$$

as $n \to \infty$. Then we may conclude by Slutsky's theorem (see, for example, [5, Proposition A.42 (b)]) and the central limit theorem for the log-determinant, Theorem 2.1, that

$$\begin{aligned} \frac{\ln \operatorname{vol}_n(\Sigma_n) + \frac{1}{2} \ln n!}{\sqrt{\frac{1}{2} \ln n}} &= \frac{\ln |\det(Y_1| \dots |Y_{n+1})| - \frac{1}{2} \ln n!}{\sqrt{\frac{1}{2} \ln n}} \\ &\quad + \frac{\ln \operatorname{dist}(U_{n+1}, L_n)}{\sqrt{\frac{1}{2} \ln n}} - \frac{\ln \operatorname{dist}(Y_{n+1}, L_n)}{\sqrt{\frac{1}{2} \ln n}} \end{aligned}$$

converges in distribution to a standard Gaussian random variable $Z \sim \mathcal{N}(0,1)$, as $n \to \infty$.

To prove (5) we need the following two auxiliary results.

Lemma 2.2 (Berry-Esseen inequality [13, Lemma 8.1])**.** *Let $B_n = (b_1, \dots, b_n)$ be a random vector whose coordinates are independent copies of a random variable ξ with mean zero, variance one and subexponential tails with exponent $\alpha > 0$ and let $V_n = (v_1, \dots, v_n)$ be a fixed unit vector in $\mathbb{R}^n$. Then there exists a constant $c \in (0, \infty)$ such that*

$$\mathrm{dist}_{\mathrm{K}}(|\langle V_n, B_n \rangle|, |Z|) \le c \|V_n\|_\infty. \tag{6}$$

Lemma 2.3 ([14, Theorem 1.4] for subexponential tails, see Remark 2.4 below)**.** *Suppose L_n is a linear subspace spanned by n independent random vectors $Y_1, \dots, Y_n$ in $\mathbb{R}^{n+1}$ each of whose coordinates are independent copies of a random variable ξ with zero mean, variance one and subexponential tails with exponent $\alpha > 0$. Let N_{n+1} be a unit normal vector to L_n.*

- *Then there are constants $c_1, c_2, c_3 \in (0, \infty)$ such that*

$$\mathbb{P}\Big(\|N_{n+1}\|_\infty \ge \sqrt{\frac{m}{n}} \Big) \le c_2\, n^2 \exp\Big(-c_3 \Big(\frac{m}{\log n} \Big)^{\frac{1}{2\alpha+1}} \Big),$$

  for every $m \ge c_1 (\log n)^{2\alpha+2}$. As a consequence, with probability $1 - c_4 n^{-10}$, say, we have that

$$\|N_{n+1}\|_\infty \le c_5 \frac{(\ln n)^{\alpha+1}}{\sqrt{n}}, \tag{7}$$

  for some constant $c_4, c_5 > 0$.
- *If $(V_n)_{n\in\mathbb{N}}$ is a fixed sequence of unit random vectors $V_n \in \mathbb{S}^{n-1}$, then*

$$\sqrt{n} \langle V_n, N_n \rangle \xrightarrow{d} Z, \quad \text{as } n \to \infty. \tag{8}$$

Remark 2.4. *Note that* [14, Theorem 1.4] *is stated for subgaussian random variables. By* [14, Remark 2.3] *the theorem holds true also for random variables with subexponential tails with exponent $\alpha > 0$, but one has to be more generous with the estimates.*

With probability one the vectors $Y_1, \dots, Y_n$ span a random n-dimensional linear subspace L_n in $\mathbb{R}^{n+1}$ and we denote by $N_{n+1} \in \mathbb{S}^n$ a unit normal vector to L_n. Then

$$\mathrm{dist}(U_{n+1}, L_n) = |\langle U_{n+1}, N_{n+1} \rangle| \qquad \text{and} \qquad \mathrm{dist}(Y_{n+1}, L_n) = |\langle Y_{n+1}, N_{n+1} \rangle|,$$

where $\langle \cdot, \cdot \rangle$ is the standard scalar product in $\mathbb{R}^{n+1}$.

For the first estimate we set

$$V_{n+1} := \frac{1}{\sqrt{n+1}} U_{n+1},$$

use (8), and conclude by the continuous mapping theorem [9, Lemma 4.3], applied to the absolute-value function, that

$$\frac{\ln \mathrm{dist}(U_{n+1}, L_n)}{\sqrt{\frac{1}{2} \ln n}} \xrightarrow{d} 0 \tag{9}$$

as $n \to \infty$. This settles the first case of (5).

The second statement of (5) also follows by Slutsky's theorem once we show that

$$\mathrm{dist}(Y_{n+1}, L_n) \xrightarrow{d} |Z|, \quad \text{as } n \to \infty. \tag{10}$$

To prove this we use Lemma 2.2 and the first part of Lemma 2.3. We condition on L_n to fix N_{n+1} and combine (7) with (6). To be more precise, we have

$$\mathbb{P}(\mathrm{dist}(Y_{n+1}, L_n) \le t) = \mathbb{E}\,\mathbb{P}(|\langle Y_{n+1}, N_{n+1}\rangle| \le t | L_n)).$$

Here $\mathbb{P}(\,\cdot\,|L_n)$ denotes the conditional probability given L_n, where L_n is the linear space spanned by $Y_1, \dots, Y_n \in \mathbb{R}^{n+1}$ and $\mathbb{E}$ denotes expectation with respect to $Y_1, \dots, Y_n$. If we condition on L_n, then N_{n+1} is a fixed unit vector in $\mathbb{R}^{n+1}$ and we may apply the Berry-Esseen inequality (6) with $V_{n+1} = N_{n+1}$ there to deduce that

$$\sup_{t\in\mathbb{R}} \big|\mathbb{P}(|\langle Y_{n+1}, N_{n+1}\rangle| \le t \,|\, L_n) - \mathbb{P}(|Z| \le t)\big| \le c\|N_{n+1}\|_\infty.$$

Moreover, from (7) we conclude that there exist constants $c_1, c_2, c_3 \in (0, \infty)$ such that

$$0 \le \sup_{t\in\mathbb{R}} \big|\mathbb{P}(|\langle Y_{n+1}, N_{n+1}\rangle| \le t \,|\, L_n) - \mathbb{P}(|Z| \le t)\big| \le c_2 \frac{(\ln n)^{c_3}}{\sqrt{n}}$$

holds true with probability $1 - c_1 n^{-10}$ for sufficiently large n. Hence we have

$$\begin{aligned}\mathrm{dist}_{\mathrm{K}}(\mathrm{dist}(Y_{n+1}, L_n), |Z|) &\le \mathbb{E} \sup_{t\in\mathbb{R}} |\mathbb{P}(|\langle Y_{n+1}, N_{n+1}\rangle| \le t|L_n) - \mathbb{P}(|Z| \le t)| \\ &\le c_2 \frac{(\ln n)^{c_3}}{\sqrt{n}} (1 - c_1 n^{-10}) + 2c_1 n^{-10}.\end{aligned}$$

Finally, we notice that the last expression tends to zero, as $n \to \infty$. This yields (10) and completes the proof. □

3. Proof of Theorem 1.3

3.1. **The probability measures $\nu = \nu_n(m_n, p)$ on the ℓ_p-ball.** Denote for each $i \in \{1, \dots, n\}$ by $G_1^i, \dots, G_n^i$ random variables with density

$$t \mapsto \frac{e^{-|t|^p/a}}{2a^{1/p}\Gamma(1 + 1/p)}, \qquad \text{where} \qquad a = \left(\frac{\Gamma(1/p)}{\Gamma(3/p)}\right)^{p/2}, \qquad t \in \mathbb{R}. \tag{11}$$

Note that G_j^i has zero mean and variance one and subexponential tails with exponent $\alpha = 1/p$. In addition, let $m_n \in [0, \infty)$ and, for each $i \in \{1, \dots, n\}$, Q_i be random variables which are gamma distributed with shape m_n and rate $1/a$. More specifically this means that Q_i has density $t \mapsto a^{-m_n}\Gamma(m_n)^{-1}t^{m_n-1}e^{-t/a}$ for $t > 0$, provided that $m_n > 0$, and we use the convention that $Q_i = 0$ with probability one in case that $m_n = 0$. We shall assume that all the random variables we are considering are independent.

Next, we define the random vectors $X_1, \dots, X_n \in \mathbb{R}^n$ by putting

$$X_i := \frac{G_i}{(\|G_i\|_p^p + Q_i)^{1/p}}, \qquad \text{where} \qquad G_i := (G_1^i, \dots, G_n^i),$$

for $i \in \{1, \dots, n\}$. By $\nu_n(m_n, p)$ we denote the distribution of the random variables X_i on the ℓ_p-ball B_p^n. Finally, we let $\Xi_n(K, \nu)$ be the random convex body that is generated by a fixed convex body $K \subset \mathbb{R}^n$ and the independent random points $X_1, \dots, X_n$ for a distribution $\nu = \nu_n(m_n, p)$, which we consider to be fixed in this section.

3.2. **Proof of Theorem 1.3.** Recall that by (4) we have

$$\begin{aligned}\mathrm{vol}_n(\Xi_n(K_n,\nu)) &= \mathrm{vol}_n(K_n)|\det(X_1|\dots|X_n)| \\ &= \mathrm{vol}_n(K_n)|\det(G_1|\dots|G_n)|\prod_{i=1}^n \big(\|G_i\|_p^p+Q_i\big)^{-1/p},\end{aligned} \tag{12}$$

which yields

$$\ln \mathrm{vol}_n(\Xi_n(K_n,\nu)) = \ln|\det(G_1|\dots|G_n)| + \ln \mathrm{vol}_n(K_n) - \frac{1}{p}\sum_{i=1}^n \ln\big(\|G_i\|_p^p+Q_i\big).$$

Now we may apply the central limit theorem for the log-determinant Theorem 2.1 and obtain

$$O_n := \frac{\ln|\det(G_1|\dots|G_n)| - \frac{1}{2}\ln(n-1)!}{\sqrt{\frac{1}{2}\ln n}} \xrightarrow{d} Z, \quad \text{as } n\to\infty, \tag{13}$$

and moreover $\mathrm{dist}_{\mathrm{K}}(O_n,Z) \le (\ln n)^{-\frac{1}{3}+o(1)}$. To complete the proof of Theorem 1.3 we need to show that

$$P_n := \frac{\frac{1}{p}\sum_{i=1}^n \ln\big(\|G_i\|_p^p+Q_i\big) - \frac{n}{p}\ln(a(m_n+\frac{n}{p}))}{\sqrt{\frac{1}{2}\ln n}} \xrightarrow{d} 0, \tag{14}$$

and then apply Slutsky's theorem. Indeed, putting together (12) with (13) and (14) implies that

$$\begin{aligned}&\frac{\ln \mathrm{vol}_n(\Xi_n(K_n,\nu)) - \ln \mathrm{vol}_n(K_n) - \frac{1}{2}\ln(n-1)! + \frac{n}{p}\ln(a(m_n+\frac{n}{p}))}{\sqrt{\frac{1}{2}\ln n}} \\ &= O_n - P_n \xrightarrow{d} Z, \quad \text{as } n\to\infty,\end{aligned}$$

as desired. For all $\varepsilon>0$ we have that

$$\mathrm{dist}_{\mathrm{K}}(O_n-P_n,Z) \le \mathrm{dist}_{\mathrm{K}}(O_n,Z) + \mathbb{P}(|P_n|>\varepsilon)+\varepsilon,$$

see for example [2, Lemma 4.1]. Hence, once we show that

$$\mathbb{P}(|P_n| > (\ln n)^{-\frac{1}{3}}) \le (\ln n)^{-\frac{1}{3}+o(1)}, \tag{15}$$

we may conclude, by setting $\varepsilon = (\ln n)^{-\frac{1}{3}}$ and applying Theorem 2.1, that

$$\mathrm{dist}_{\mathrm{K}}(O_n-P_n,Z) \le (\ln n)^{-\frac{1}{3}+o(1)}.$$

This will finish the proof of Theorem 1.3.

3.3. **Proof of** (15)**.** We observe that the representation

$$\|G_i\|_p^p = \sum_{j=1}^n |G_j^i|^p$$

and the semigroup property of the gamma distributions imply that $\|G_i\|_p^p$ is gamma distributed with shape n/p and rate $1/a$, i.e., the Lebesgue density of $\|G_i\|_p^p$ on $\mathbb{R}$ is given by

$$\frac{1}{a^{\frac{n}{p}}\Gamma(\frac{n}{p})} x^{\frac{n}{p}-1} e^{-\frac{x}{a}}, \qquad x>0.$$

Recalling that by assumption Q_i is gamma distributed with shape m_n and rate $1/a$ we find that $\|G_i\|_p^p + Q_i$ is also gamma distributed with parameter shape $m_n + \frac{n}{p}$ and rate $1/a$. Hence, $\ln(\|G_i\|_p^p + Q_i)$ is log-gamma distributed with Lebesgue density on $\mathbb{R}$ given by

$$\frac{1}{a^{m_n+\frac{n}{p}}\Gamma(m_n+\frac{n}{p})} e^{x(m_n+\frac{n}{p})-\frac{e^x}{a}}, \qquad x > 0.$$

In particular, by direct computation, we find that

$$\mu_n := \mathbb{E}\ln(\|G_i\|_p^p + Q_i) = \psi\Big(m_n + \frac{n}{p}\Big) + \ln a, \tag{16}$$

where $\psi(x) := \frac{d}{dx}\ln\Gamma(x)$ is the digamma function (see e.g. [1, page 259]). Similarly, for the variance, one has that

$$\sigma_n^2 := \mathrm{Var}\ln(\|G_i\|_p^p + Q_i) = \psi_1\Big(m_n + \frac{n}{p}\Big),$$

with $\psi_1(x) := \frac{d}{dx}\psi(x)$ being the trigamma function (see e.g. [1, page 260]).

This implies that the auxiliary random variables

$$A_n := \Bigg[\frac{1}{p}\sum_{i=1}^n \ln\big(\|G_i\|_p^p + Q_i\big)\Bigg] - \frac{n}{p}\mu_n$$

satisfy

$$\mathbb{E}A_n = 0 \qquad \text{and} \qquad \mathrm{Var}\,A_n = \frac{n}{p^2}\psi_1\Big(m_n + \frac{n}{p}\Big).$$

The asymptotic expansions of the digamma and trigamma functions are

$$\psi(x) = \ln x - \frac{1}{2x} - \frac{1}{12x^2} + o(x^{-2}), \qquad \psi_1(x) = \frac{1}{x} + \frac{1}{2x^2} + o(x^{-2}),$$

for $x \to \infty$ (see [1, page 260]). Hence

$$\frac{\frac{n}{p}\mu_n - \frac{n}{p}\ln(a(m_n + \frac{n}{p}))}{\sqrt{\frac{1}{2}\ln n}} = \frac{1}{\sqrt{2}}\Big(1 + p\frac{m_n}{n}\Big)^{-1}(\ln n)^{-\frac{1}{2}} + o((\ln n)^{-\frac{1}{2}}) \longrightarrow 0,$$

for $n \to \infty$, and

$$\mathrm{Var}\,A_n = \frac{n}{p}\frac{1}{pm_n + n} + \frac{n}{2}\frac{1}{(pm_n+n)^2} + o(n^{-1}).$$

In particular, for all choices of m_n we find that $\mathrm{Var}\,A_n = O(1)$. By the triangle inequality we have

$$|P_n| \le \frac{|A_n|}{\sqrt{\frac{1}{2}\ln n}} + \frac{1}{\sqrt{2}}\Big(1 + p\frac{m_n}{n}\Big)^{-1}(\ln n)^{-\frac{1}{2}} + o((\ln n)^{-\frac{1}{2}})$$

and therefore there exists $c_1 > 0$ such that

$$|P_n| \le \frac{|A_n| + c_1}{\sqrt{\frac{1}{2}\ln n}},$$

for all n large enough. By the Chebyshev inequality this yields

$$\mathbb{P}(|P_n| > (\ln n)^{-\frac{1}{3}}) \leq \mathbb{P}\bigg(|A_n| > \frac{1}{\sqrt{2}}(\ln n)^{\frac{1}{6}} - c_1\bigg)$$
$$\leq \mathbb{P}\bigg(|A_n| > \frac{1}{2}(\ln n)^{\frac{1}{6}}\bigg) \leq \frac{\sqrt{2}\operatorname{Var} A_n}{(\ln n)^{\frac{1}{3}}} = (\ln n)^{-\frac{1}{3}+o(1)},$$

for all n large enough. Thus, (15) holds true and the proof is complete. □

Remark 3.1. *More general distributions for the random variables Q_i are possible. For example, our proof shows that, as long as Q_i is a non-negative random variable and $\operatorname{Var}\ln(\|G_i\|_p^p + Q_i) = O\big(\operatorname{Var}\ln(\|G_i\|_p^p)\big) = O(p/n)$, we have a CLT as above with the same scaling factor and a suitably modified final centering term.*

References

[1] M. Abramowitz and I.A. Stegun, *Handbook of Mathematical Functions*, Dover Publications (1965).

[2] D. Alonso-Gutiérrez, J. Prochno, and C. Thäle, *Gaussian fluctuations for high-dimensional random projections of ℓ_p^n-balls*, accepted for publication in Bernoulli (2018$^+$).

[3] I. Bárány, *Random polytopes, convex bodies, and approximation*, in Weil, W. (Ed.) *Stochastic Geometry*, Lecture Notes in Mathematics **1892**, Springer (2007).

[4] 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.

[5] R.F. Bass, *Stochastic Processes*, Cambridge University Press (2011).

[6] P. Eichelsbacher and L. Knichel, *Fine asymptotics for models with Gamma type moments*, arXiv: 1710.06484.

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

[8] D. Hug, *Random polytopes*, in: Spodarev, E. (Ed.), *Stochastic Geometry, Spatial Statistics and Random Fields. Asymptotic Methods*, Lecture Notes in Mathematics **2068**, Springer (2013).

[9] O. Kallenberg, *Foundations of Modern Probability*, Second Edition, Springer, 2002.

[10] A.M. Mathai, *On a conjecture in geometric probability regarding asymptotic normality of a random simplex*, Ann. Probab. **10** (1982), 247–251.

[11] H. Maehara, *On random simplices in product distributions*, J. Appl. Probab. **17** (1980), 553–558.

[12] R.E. Miles, *Isotropic random simplices*, Advances in Appl. Probability **3** (1971), 353–382.

[13] H.H. Nguyen and V. Vu, *Random matrices: law of the determinant*, Ann. Probab. **42** (2014), 146–167.

[14] H.H. Nguyen and V.H. Vu, *Normal vector of a random hyperplane*, Int. Math. Res. Not. IMRN (2018), 1754–1778.

[15] G. Paouris and P. Pivovarov, *A probabilistic take on isoperimetric-type inequalities*, Adv. Math. **230** (2012), 1402–1422.

[16] G. Paouris and P. Pivovarov, *Small-ball probabilities for the volume of random convex sets* Discrete Comput. Geom. **49** (2013), 601–646.

[17] M. Reitzner, *Random polytopes*, In: Molchanov, I., and Kendall, W. (eds.): *New Perspectives in Stochastic Geometry*, pp. 45–76, Oxford University Press (2010).

[18] H. Ruben, *The volume of a random simplex in an n-ball is asymptotically normal*, J. Appl. Probability **14** (1977), 647–653.

University of Zaragoza, Spain
E-mail address: alonsod@unizar.es

Vienna University of Technology, Austria
E-mail address: florian.besau@tuwien.ac.at

University of Ulm, Germany
E-mail address: julian.grote@uni-ulm.de

University of Münster, Germany
E-mail address: zakhar.kabluchko@wwu.de

University of Osnabrück, Germany
E-mail address: matthias.reitzner@uni-osnabrueck.de

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

University of Alberta in Edmonton, Canada
E-mail address: vritsiou@ualberta.ca

Case Western Reserve University, USA
E-mail address: elisabeth.werner@case.edu