Uniform behaviors of random polytopes under the Hausdorff metric

VICTOR-EMMANUEL BRUNEL

Department of Mathematics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139-4307, USA. E-mail: veb@mit.edu

We study the Hausdorff distance between a random polytope, defined as the convex hull of i.i.d. random points, and the convex hull of the support of their distribution. As particular examples, we consider uniform distributions on convex bodies, densities that decay at a certain rate when approaching the boundary of a convex body, projections of uniform distributions on higher dimensional convex bodies and uniform distributions on the boundary of convex bodies. We essentially distinguish two types of convex bodies: those with a smooth boundary and polytopes. In the case of uniform distributions, we prove that, in some sense, the random polytope achieves its best statistical accuracy under the Hausdorff metric when the support has a smooth boundary and its worst statistical accuracy when the support is a polytope. This is somewhat surprising, since the exact opposite is true under the Nikodym metric. We prove rate optimality of most our results in a minimax sense. In the case of uniform distributions, we extend our results to a rescaled version of the Hausdorff metric. We also tackle the estimation of functionals of the support of a distribution such as its mean width and its diameter. Finally, we show that high dimensional random polytopes can be approximated with simple polyhedral representations that significantly decrease their computational complexity without affecting their statistical accuracy.

Keywords: computational geometry; convex bodies; convex hull; deviation inequality; Hausdorff metric; high dimension; minimax estimation; random polytope

1. Introduction

A simple representation of random polytopes consists of taking the convex hull of random points in the Euclidean space. Other representations have been suggested and studied, such as random projections of high dimensional polytopes or intersection of random halfspaces (about random polytopes, see [25,39,43] and the references therein).

The study of random polytopes goes back to Rényi and Sulanke's seminal works [40,41], where the authors studied the area and the number of vertices of the convex hull of independent and identically distributed (i.i.d.) points uniformly distributed in a planar convex body. A long series of work followed so as to better understand the behavior of the volume of the random polytope in higher dimensions, depending on the structure of the supporting convex body [1,4,5, 7,13,17,18,37,45,46,49] also allowing for more general distributions (see [10] and the references therein).

Combinatorial properties of random polytopes are well understood in expectation [3] as well as very precise results about their functionals, such as intrinsic volumes [6,8,9,34].

Beyond these probabilistic results, random polytopes are useful to approximate convex bodies by polytopes with few vertices [16,22,38,42] and they are also in order to estimate density

supports under convexity restrictions [11,15,28–30,33] (see also [35] for weaker convexity type restrictions) or their volume [2,20].

In the aforementioned works, the random polytope is described as the convex hull of i.i.d. points whose distribution support is a convex body and an important quantity is its missing volume, that is, the Lebesgue measure of the set-difference between the support and the random polytope. When the support has volume one, it is well known that the expected missing volume is of order at least $(\ln n)^{d-1}/n$ and at most $n^{-2/(d+1)}$, where n is the number of points and d is the dimension of the Euclidean space [3,23]. Moreover, the former rate is achieved when the support has a smooth boundary, whereas the latter is achieved when the support is a polytope [3]. The Hausdorff distance between the random polytope and the convex support has attracted less interest and most significant results appear in works on approximation and estimation of convex bodies [16,30]. In [30], it is shown that the random polytope is a rate optimal estimator of the support, in a minimax sense, when the distribution of the random points is uniform. In [16], more general distributions are considered, including uniform on the boundary of a convex body, or nearly uniform distributions in smooth convex bodies or in polytopes: Almost sure rates of convergence are proven for the Hausdorff distance between the support and the random polytope, as the number n of points goes to infinity. One of the objectives of this article is to refine these results by proving general yet explicit deviation inequalities and show that the two extreme cases mentioned above (smooth boundary vs. polytope) are reversed under the Hausdorff metric.

2. Notation, outline and contributions

Throughout this work, d is a fixed positive integer that stands for the dimension of the ambient Euclidean space. A convex body in $\mathbb{R}^d$ is a compact and convex set with nonempty interior. We denote by $\mathcal{K}_d$ the class of all convex bodies in $\mathbb{R}^d$ and by $\mathcal{K}_d^{(1)}$ the class of all convex bodies included in the closed unit Euclidean ball in $\mathbb{R}^d$.

For all positive integers p, the p-dimensional closed Euclidean ball with center $a \in \mathbb{R}^p$ and radius $r > 0$ is denoted by $B_p(a, r)$ and the $(p-1)$-dimensional Euclidean unit sphere is denoted by $\mathbb{S}^{p-1}$. We write κ_p for the p-dimensional volume of $B_p(0, 1)$ and ω_{p-1} for the surface area of the $(p-1)$-dimensional unit sphere (note that $\omega_{p-1} = p\kappa_p$). The p-dimensional volume of a Borel set A in $\mathbb{R}^p$ is denoted by $\mathsf{Vol}_p(A)$.

The Euclidean norm in $\mathbb{R}^d$ is denoted by $\|\cdot\|$ and the Euclidean distance is denoted by $\rho(\cdot, \cdot)$.

The interior of a set $A \subseteq \mathbb{R}^d$ is denoted by $\mathsf{int}(A)$ and its boundary by ∂A. The complement of a set or event A is denoted by $A^{\complement}$.

The support function h_A of a compact set $A \subset \mathbb{R}^d$ is the mapping $h_A(u) = \max\{\langle u, x \rangle : x \in A\}$, $u \in \mathbb{R}^d$, where $\langle \cdot, \cdot \rangle$ stands for the canonical inner product in $\mathbb{R}^d$: It is the largest signed distance between the origin and a supporting hyperplane of A orthogonal to u.

Support functions have two important properties, which will be useful in the sequel. Let K be a compact and convex set. Then, h_K is positively homogeneous, that is,

$$h_K(\lambda u) = \lambda h_K(u) \tag{1}$$

for all $u \in \mathbb{R}^d$, $\lambda \geq 0$, and it is subadditive:

$$h_K(u + v) \leq h_K(u) + h_K(v) \tag{2}$$

for all $u, v \in \mathbb{R}^d$. In particular, (2) implies the reverse triangle inequality

$$h_K(u - v) \geq h_K(u) - h_K(v) \tag{3}$$

for all $u, v \in \mathbb{R}^d$.

Let $K \in \mathcal{K}_d$ with $0 \in \operatorname{int}(K)$. The pseudo-norm $\|\cdot\|_K$ (also called the Minkowski functional of K) is defined as $\|x\|_K = \min\{\lambda \geq 0 : x \in \lambda K\}, x \in \mathbb{R}^d$. It is a norm if and only if K is symmetric, that is, $K = -K$. If $K = B_d(0, 1)$, then $\|\cdot\|_K$ is the Euclidean norm in $\mathbb{R}^d$.

The polar set K° of a convex body K is the set of all $x \in \mathbb{R}^d$ such that $h_K(x) \leq 1$. It is clear that 0 is an interior point of K° as soon as K is bounded. Moreover, if $0 \in \operatorname{int}(K)$, then K° is compact and $h_K(x) = \|x\|_{K^\circ}$ for all $x \in \mathbb{R}^d$. In particular, if K is bounded and has 0 in its interior, then $K^\circ \in \mathcal{K}_d$.

The Hausdorff distance between two sets $K, K' \subseteq \mathbb{R}^d$ is denoted by $\mathsf{d}_\mathsf{H}(K, K')$ and it is defined as $\mathsf{d}_\mathsf{H}(K, K') = \max\{\max_{x \in K} \rho(x, K'), \max_{y \in K'} \rho(y, K)\}$. If K and K' are compact and convex, it can be expressed in terms of their support functions: $\mathsf{d}_\mathsf{H}(K, K') = \sup_{u \in \mathbb{S}^{d-1}} |h_K(u) - h_{K'}(u)|$.

If $K \in \mathcal{K}_d$, $u \in \mathbb{S}^{d-1}$ and $\varepsilon \geq 0$, we denote by $C_K(u, \varepsilon)$ the cap of K in the direction of u and of width ε:

$$C_K(u, \varepsilon) = \{x \in K : \langle u, x \rangle \geq h_K(u) - \varepsilon\}.$$

In the paper, all random variables are defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For notation convenience and with no loss of generality, if μ is a given distribution in $\mathbb{R}^d$ equipped with its Borel σ-algebra, we identify $\mathbb{P}$ and μ (by assuming that $\Omega = \mathbb{R}^d$). The support of μ is denoted by $\operatorname{supp}(\mu)$. The uniform distribution on a compact set $A \subseteq \mathbb{R}^d$ is denoted by $\mathbb{P}_A$ and the corresponding expectation operator is denoted by $\mathbb{E}_A$.

If (u_n) and (v_n) are two positive sequences, we write that $u_n = O(v_n)$ or, equivalently, $v_n = \omega(u_n)$, if the sequence (u_n/v_n) is bounded.

The paper is organised as follows. In Section 3, we state and prove our main theorems. The first one is a uniform deviation inequality that yields a stochastic upper bound for the Hausdorff distance between the random polytope and the convex hull of the support of the random points. It is uniform in the sense that the bounds do not depend on the probability measure of the points or its support. The second one gives the rate of convergence of functionals of the random polytope, such as the mean width or the maximal width, to those of the convex hull of the support of the random points. Both theorems hold under an assumption on the probability mass of caps of the convex hull of the support of the random points. That assumption is discussed in the last part of Section 3, where we list particular cases that are most relevant to the stochastic geometry and the statistics literature. In Section 4, we assume that the distribution of the random points in uniform on a convex body and we prove rate-optimality of our first theorem in a minimax sense. In Section 5, we use a different technique in order to prove a refined deviation inequality for the convex hull of uniform points in polytopes. In addition, we show that up to a logarithmic factor, the polytopal case is the least favorable for the statistical accuracy of random polytopes, in some sense. This result is surprising since the opposite is true under the Nikodym metric (i.e., volume of the symmetric difference): polytopal supports are the most favorable, whereas convex supports with smooth boundary are the least [3,23]. In Section 6, we focus on uniform distributions and we extend our results under a rescaled version of the Hausdorff metric, which allows

to consider a broader class of convex bodies. In Section 7, we tackle the problem of computational complexity of high dimensional random polytopes and propose a random approximation of random polytopes that achieves a significant gain in the computational cost without affecting the statistical accuracy.

3. Behavior of the random polytope and its functionals

3.1. Behavior of the random polytope

Let $K \in \mathcal{K}_d$ and μ be a probability measure in $\mathbb{R}^d$, whose support is contained in K. We make the following assumption on the pair (μ, K).

Assumption 1. Let α, L and ε_0 be positive numbers with $\varepsilon_0 \leq 1$. For all $u \in \mathbb{S}^{d-1}$ and $\varepsilon \in [0, \varepsilon_0]$, $\mu(C_K(u, \varepsilon)) \geq L\varepsilon^\alpha$.

Note that if (μ, K) satisfy Assumption 1, then K is necessarily the convex hull of the support of μ. Indeed, K is convex and it is easy to see that every closed halfspace that contains either $\mathsf{supp}(\mu)$ or K needs to contain the other as well.

Consider a collection $X_1, \ldots, X_n$ of i.i.d. random points with probability distribution μ, where n is a positive integer and let $\hat{K}_n$ be their convex hull. The following theorem shows that under Assumption 1, $\hat{K}_n$ concentrates around K under the Hausdorff metric, at an explicit rate that depends on α. For all $\alpha > 0$, we set

$$C_\alpha = \inf_{t>0} \frac{(1+t)^\alpha}{1+t^\alpha} = \begin{cases} 1 & \text{if } \alpha \geq 1, \\ 2^{\alpha-1} & \text{otherwise.} \end{cases}$$

Theorem 1. *Let α, L, ε_0 be positive numbers with $0 < \varepsilon_0 \leq 1$. Assume Assumption 1 and that $K \subseteq B_d(0, 1)$. Set $\tau_1 = \max(1, \frac{d}{C_\alpha \alpha L})$, $a_n = (\frac{\tau_1 \ln n}{n})^{\frac{1}{\alpha}}$ and $b_n = n^{\frac{-1}{\alpha}}$. Then, the random polytope $\hat{K}_n$ satisfies*

$$\mathbb{P}\big[\mathsf{d_H}(\hat{K}_n, K) \geq 2a_n + 2b_n x\big] \leq 12^d \exp\big(-C_\alpha L x^\alpha\big),$$

for all $x \geq 0$ with $a_n + b_n x \leq \varepsilon_0$.

Proof. Let $\varepsilon \in [0, \varepsilon_0]$ and $\delta = \varepsilon/4$. For all $u \in \mathbb{S}^{d-1}$, we have $h_{\hat{K}_n}(u) \leq h_K(u) - \varepsilon$ if and only if all of the points $X_1, \ldots, X_n$ lie outside of $C_K(u, \varepsilon)$. Hence,

$$\mathbb{P}\big[h_{\hat{K}_n}(u) \leq h_K(u) - \varepsilon\big] = \big(1 - \mu\big(C_K(u, \varepsilon)\big)\big)^n$$

$$\leq \exp\big(-Ln\varepsilon^\alpha\big).$$

(4)

The following lemma can be found in [19], Lemma 5.2.

Lemma 1. *Let $K \in \mathcal{K}_d$ with $0 \in \mathrm{int}(K)$ and $\delta \in (0, 1/2)$. There exists a subset $\mathcal{N}_\delta \subseteq \partial K$ of cardinality at most $\left(\frac{3}{\delta}\right)^d$ such that, for all $u \in \partial K$, there exist two sequences $(u_j)_{j \geq 0} \subseteq \mathcal{N}_\delta$ and $(\delta_j)_{j \geq 1} \subseteq \mathbb{R}$ such that $0 \leq \delta_j \leq \delta^j$ for all $j \geq 1$, and $u = u_0 + \sum_{j=1}^\infty \delta_j u_j$.*

Let $\mathcal{N}_\delta$ be a subset of $\mathbb{S}^{d-1}$ satisfying Lemma 1 applied to the unit Euclidean ball. Let us denote by A the event $\{h_{\hat{K}_n}(u) > h_K(u) - \varepsilon, \forall u \in \mathcal{N}_\delta\}$. By (4) and the union bound,

$$\mathbb{P}(A) \geq 1 - \left(\frac{3}{\delta}\right)^d \exp(-Ln\varepsilon^\alpha). \tag{5}$$

Let A hold and $u \in \mathbb{S}^{d-1}$. By Lemma 1, we can write $u = u_0 + \sum_{j=1}^\infty \delta_j u_j$ where $(u_j)_{j \geq 0} \subseteq \mathcal{N}_\delta$ and $0 \leq \delta_j \leq \delta^j$ for all $j \geq 1$. Note that almost surely, $\hat{K}_n \subseteq K \subseteq B_d(0, 1)$, hence $h_{\hat{K}_n}(u) \leq h_K(u) \leq 1$, for all $u \in \mathbb{S}^{d-1}$. Thus,

$$
\begin{aligned}
h_{\hat{K}_n}(u) = h_{\hat{K}_n}\left(u_0 + \sum_{j=1}^\infty \delta_j u_j\right) &\geq h_{\hat{K}_n}(u_0) - \sum_{j=1}^\infty \delta^j h_{\hat{K}_n}(-u_j) \qquad \text{by (1) and (2)} \\
&> h_K(u_0) - \varepsilon - \frac{\delta}{1-\delta} \\
&\geq h_K(u) - \sum_{j=1}^\infty \delta^j h_K(u_j) - \varepsilon - \frac{\delta}{1-\delta} \qquad \text{again by (1) and (2)} \\
&\geq h_K(u) - \varepsilon - \frac{2\delta}{1-\delta} \geq h_K(u) - 2\varepsilon,
\end{aligned}
\tag{6}
$$

since $\delta = \varepsilon/4 \leq \varepsilon_0/4 \leq 1/2$.

Hence, by (5) and (6),

$$
\begin{aligned}
\mathbb{P}\left[\mathsf{d}_\mathsf{H}(\hat{K}_n, K) \geq 2\varepsilon\right] = \mathbb{P}\left[\exists u \in \mathbb{S}^{d-1}, h_{\hat{K}_n}(u) \leq h_K(u) - \varepsilon\right] \\
\leq \left(\frac{12}{\varepsilon}\right)^d \exp(-Ln\varepsilon^\alpha) \\
\leq 12^d \exp(-Ln\varepsilon^\alpha - d\ln\varepsilon).
\end{aligned}
\tag{7}
$$

By setting $\varepsilon = a_n + b_n x$ for some nonnegative number x, where a_n and b_n are defined in Theorem 1, we have that

$$\varepsilon^\alpha \geq C_\alpha \left(\frac{\tau_1 \ln n}{n} + \frac{x^\alpha}{n}\right),$$

and

$$\ln \varepsilon \geq \ln a_n \geq -\frac{\ln n}{\alpha}.$$

The conclusion of Theorem 1 follows. $\square$

If $a_n + b_n x > 1$, the upper bound in Theorem 1 can be replaced with 0, since K and $\hat{K}_n$ are both subsets of the unit ball, and so the Hausdorff distance between them is at most 2. When $\varepsilon_0 \leq a_n + b_n x \leq 1$, the upper bound in Theorem 1 can be replaced with its corresponding value at $x = (\varepsilon_0 - a_n) b_n^{-1}$, since the left-hand side of the inequality is a nonincreasing function of x. The constants in Theorem 1 do not depend on μ and K: The deviation inequality is uniform in all pairs (μ, K) that satisfy Assumption 1, with $K \subseteq B_d(0, 1)$. It is important to notice that the constant factors in Theorem 1 are much smaller than in [10], Theorem 1, when the dimension d becomes large. However, we believe that the factor 12^d cannot be replaced with a sub-exponential factor in general.

Theorem 1 yields the following moment inequalities.

Corollary 1. *Let the assumptions of Theorem 1 hold. For all real numbers $q \geq 1$,*

$$\mathbb{E}\big[\mathsf{d}_\mathsf{H}(\hat{K}_n, K)^q\big] = O\left(\left(\frac{\ln n}{n}\right)^{\frac{q}{\alpha}}\right),$$

with constant factors that do not depend on μ and K.

Proof. Let Z be a nonnegative random variable and q be a positive number. Then, by Fubini's theorem,

$$\mathbb{E}\big[Z^q\big] = q \int_0^\infty t^{q-1} \mathbb{P}[Z \geq t]\,\mathrm{d}t.$$

Let us apply this equality to $Z = \mathsf{d}_\mathsf{H}(\hat{K}_n, K)$ (which we denote by d_H for the sake of simplicity) and $q \geq 1$. Since $\mathsf{d}_\mathsf{H}(\hat{K}_n, K) \leq 2$ μ-almost surely,

$$
\begin{aligned}
\mathbb{E}\big[&\mathsf{d}_\mathsf{H}(\hat{K}_n, K)^q\big] \\
&= q \int_0^\infty t^{q-1} \mathbb{P}[\mathsf{d}_\mathsf{H} \geq t]\,\mathrm{d}t \\
&= q \int_0^{2a_n} t^{q-1} \mathbb{P}[\mathsf{d}_\mathsf{H} \geq t]\,\mathrm{d}t + q \int_{2a_n}^{2\varepsilon_0} t^{q-1} \mathbb{P}[\mathsf{d}_\mathsf{H} \geq t]\,\mathrm{d}t \\
&\quad + q \int_{2\varepsilon_0}^2 t^{q-1} \mathbb{P}[\mathsf{d}_\mathsf{H} \geq t]\,\mathrm{d}t.
\end{aligned}
\tag{8}
$$

In the first integral, let us bound the probability by 1. In the second and third integrals, we perform the change of variable $t = 2a_n + 2b_n x$. Then, by Theorem 1 and the remark that follows it, (8) becomes:

$$
\begin{aligned}
\mathbb{E}\big[\mathsf{d}_\mathsf{H}(\hat{K}_n, K)^q\big] &\leq (2a_n)^q + 2q b_n \int_0^{(\varepsilon_0 - a_n) b_n^{-1}} (2a_n + 2b_n x)^{q-1} \mathbb{P}[\mathsf{d}_\mathsf{H} \geq 2a_n + 2b_n x]\,\mathrm{d}x \\
&\quad + 2q b_n \int_{(\varepsilon_0 - a_n) b_n^{-1}}^{(1 - a_n) b_n^{-1}} (2a_n + 2b_n x)^{q-1} \mathbb{P}[\mathsf{d}_\mathsf{H} \geq 2a_n + 2b_n x]\,\mathrm{d}x
\end{aligned}
\tag{9}
$$

$$\leq (2a_n)^q + q b_n \int_0^{(\varepsilon_0 - a_n)b_n^{-1}} (2a_n + 2b_n x)^{q-1} 12^d \exp(-C_\alpha L x^\alpha) \, dx$$

$$+ q 2^{q-1} |1 - \varepsilon_0| b_n^{-1} 12^d \exp(-C_\alpha L |1 - \varepsilon_0|^\alpha b_n^{-\alpha}).$$

In the second term, we bound $(a_n + b_n x)^{q-1}$ by $2^{q-1} a_n^{q-1} + b_n^{q-1} x^{q-1}$, which yields

$$\int_0^{(\varepsilon_0 - a_n)b_n^{-1}} (2a_n + 2b_n x)^{q-1} 12^d \exp(-C_\alpha L x^\alpha) \, dx$$

$$\leq 12^d 4^{q-1} \left(a_n^{q-1} \int_0^\infty \exp(-C_\alpha L x^\alpha) \, dx + b_n^{q-1} \int_0^\infty x^{q-1} \exp(-C_\alpha L x^\alpha) \, dx \right). \tag{10}$$

Since $b_n = O(a_n)$, (9) and (10) yield Corollary 1. $\qquad\qquad\qquad\qquad\square$

Corollary 1 implies an upper bound for the minimax risk for estimation of $K \in \mathcal{K}_d^{(1)}$ under Assumption 1, of order $((\ln n)/n)^{1/\alpha}$ (see [48], Chapter 2, for a formal definition of minimax risks).

3.2. Estimation of functionals of K

In this section, we study a class of functionals of K that can be written in terms of integrals involving its support function. We bound the performances of the corresponding functionals of $\hat{K}_n$, which are also called *plug-in* estimators. What we call a functional is a mapping $T : \mathcal{K}_d \longrightarrow \mathbb{R}$ and the plug-in estimator of $T(K)$ is $T(\hat{K}_n)$.

For all real numbers $p \geq 1$ and all measurable functions $f : \mathbb{S}^{d-1} \to \mathbb{R}$, denote by $\|f\|_p = (\int_{\mathbb{S}^{d-1}} |f(u)|^p \, d\sigma(u))^{\frac{1}{p}}$, where σ is the uniform probability measure on the sphere and by $\|f\|_\infty = \sup_{u \in \mathbb{S}^{d-1}} |f(u)|$, provided these quantities are finite.

For $K \in \mathcal{K}_d$, denote by $w_K(u)$ its width in direction $u \in \mathbb{S}^{d-1}$: $w_K(u) = h_K(u) + h_K(-u)$. We are interested in the functionals $T_p(K) = \|h_K\|_p$ and $S_p(K) = \|w_K\|_p$, for $p \in [1, \infty]$ (extensions to broader classes of functionals would be possible). For instance, S_1 is the mean width and S_∞ is the diameter. Asymptotics of the expected mean width of $\hat{K}_n$ are known in the case of uniform measures in K. When μ is uniform in a smooth convex body K, it is well known that $S_1(K) - \mathbb{E}_K[S_1(\hat{K}_n)] \longrightarrow c_K n^{\frac{-2}{d+1}}$ as $n \to \infty$, where c_K is a positive number that depends on K, see [8,45] for explicit formulas. If μ is uniform in a polytope K, Schneider [42] proved that $S_1(K) - \mathbb{E}_K[S_1(\hat{K}_n)] \longrightarrow c_K' n^{\frac{-1}{d}}$ as $n \to \infty$, for some other positive constant c_K' that also depends on K. When μ is supported on the boundary of a smooth convex body K with a positive density g with respect to the surface area measure of ∂K, Müller [34] showed that $S_1(K) - \mathbb{E}_K[S_1(\hat{K}_n)] \longrightarrow c_{g,K}'' n^{\frac{-2}{d-1}}$ as $n \to \infty$, where the positive number $c_{g,K}''$ depends on g and K. Here, we prove moment inequalities for $S_1(\hat{K}_n)$ as well as for the other functionals $S_p, T_p, p \in [1, \infty]$ under Assumption 1.

Let $p \in [1, \infty]$ and $q \geq 1$. If K and K' are two convex bodies, then by the triangle inequality,

$$\left|T_p(K) - T_p(K')\right|^q \leq \|h_K - h_{K'}\|_p^q \tag{11}$$

and

$$\left|S_p(K) - S_p(K')\right|^q \leq \|w_K - w_{K'}\|_p^q. \tag{12}$$

The next theorem states uniform upper bounds for the moments $\mathbb{E}[\|h_K - h_{\hat{K}_n}\|_p^q]$ and $\mathbb{E}[\|w_K - w_{\hat{K}_n}\|_p^q]$ when (μ, K) satisfies Assumption 1. Note that the case $p = \infty$ is treated in Corollary 1 and we only treat here the case of finite p.

Theorem 2. *Let d be a positive integer, $\alpha, L > 0$, $0 < \varepsilon_0 \leq 1$ and $p, q \geq 1$ be real numbers. Assume Assumption 1 and that $K \subseteq B_d(0, 1)$. Then,*

$$\mathbb{E}[\|h_K - h_{\hat{K}_n}\|_p^q] = O(n^{\frac{-q}{\alpha}})$$

and

$$\mathbb{E}[\|w_K - w_{\hat{K}_n}\|_p^q] = O(n^{\frac{-q}{\alpha}}).$$

Proof. Let $u \in \mathbb{S}^{d-1}$. By (4), for all $\varepsilon \in [0, \varepsilon_0]$,

$$\mathbb{P}[h_{\hat{K}_n}(u) \leq h_K(u) - \varepsilon] \leq \exp(-Ln\varepsilon^\alpha).$$

For $\varepsilon \in (\varepsilon_0, 2]$,

$$\mathbb{P}[h_{\hat{K}_n}(u) \leq h_K(u) - \varepsilon] \leq \exp(-Ln\varepsilon_0^\alpha),$$

since the left hand side is a nonincreasing function of ε and if $\varepsilon > 2$,

$$\mathbb{P}[h_{\hat{K}_n}(u) \leq h_K(u) - \varepsilon] = 0.$$

Hence, by a similar argument as in the proof of Corollary 1, for all real numbers $k \geq 1$, there is some positive number c_k that depends neither on μ, K nor on u and such that

$$\mathbb{E}[|h_K(u) - h_{\hat{K}_n}(u)|^k] \leq c_k n^{\frac{-k}{\alpha}}. \tag{13}$$

If $q \leq p$, Jensen's inequality applied to the concave function $[0, \infty) \ni x \mapsto x^{\frac{q}{p}}$ yields

$$\mathbb{E}[\|h_{\hat{K}_n} - h_K\|_p^q] = \mathbb{E}\left[\left(\int_{\mathbb{S}^{d-1}} |h_K(u) - h_{\hat{K}_n}(u)|^p \, d\sigma(u)\right)^{q/p}\right]$$

$$\leq \mathbb{E}\left[\int_{\mathbb{S}^{d-1}} |h_K(u) - h_{\hat{K}_n}(u)|^p \, d\sigma(u)\right]^{q/p}. \tag{14}$$

By Fubini's theorem, which allows switching the expectation and the integral and by (13) with $k = p$, (14) yields

$$\mathbb{E}\big[\|h_{\hat{K}_n} - h_K\|_p^q\big] \le c_p^{q/p} n^{\frac{-q}{\alpha}},$$

which proves the theorem in the case $q \le p$. If now $q > p$, the mapping $[0, \infty) \ni x \mapsto x^{\frac{q}{p}}$ is convex and Jensen's inequality yields

$$\left(\int_{\mathbb{S}^{d-1}} \big|h_K(u) - h_{\hat{K}_n}(u)\big|^p\right)^{q/p} \le \int_{\mathbb{S}^{d-1}} \big|h_K(u) - h_{\hat{K}_n}(u)\big|^q,$$

so again by Fubini's theorem and using (13) with $k = q$,

$$\mathbb{E}\big[\|h_{\hat{K}_n} - h_K\|_p^q\big] \le c_q n^{\frac{-q}{\alpha}}. \qquad \square$$

It is not surprising that in Theorem 2, there is no logarithmic factor when $p \ne \infty$, as opposite to the case $p = \infty$ (see Corollary 1). This is due to the fact that L^∞ norms are more restrictive than L^p norms ($p \in [1, \infty)$), since they capture the largest values of a function, no matter the size of the subset where large values are achieved. When μ is the uniform probability measure in a smooth convex body, Theorem 3 in Section 4 shows that the logarithmic factor cannot be avoided for $p = \infty$. In [24], a convex body is estimated from noisy observations of its support function, and the risk is evaluated in terms of the L^2 norm of the support function, which exactly corresponds to $p = 2$ in Theorem 2. The risk of the least square estimator does not have logarithmic factors, and it is proven to be rate optimal, in a minimax sense. However, we believe that a logarithmic factor would be unavoidable if the risk was measured in terms of the L^∞ norm of the support function, i.e., in terms of the Hausdorff metric.

Another functional of interest, the *thickness* of K, exhibits a similar behavior as the diameter. The thickness of K, denoted by $\Delta(K)$, is its minimal width, i.e., $\Delta(K) = \min_{u \in \mathbb{S}^{d-1}} w_K(u)$. It is easy to see that $|\Delta(\hat{K}_n) - \Delta(K)| \le \|w_{\hat{K}_n} - w_K\|_\infty$. Hence, the plug-in estimator $\Delta(\hat{K}_n)$ converges at least at the same speed as $\hat{K}_n$ converges to K under the Hausdorff metric.

3.3. Special cases of Assumption 1

Before listing special cases, we recall the definition of the *reach* of a set (see, e.g., [47], Definition 11).

Definition 1. Let $A \subset \mathbb{R}^d$. The *reach* of A is the supremum of the set of real numbers $r \ge 0$ such that each $x \in A^{\complement}$ with $\rho(x, \partial A) \le r$ has a unique metric projection on ∂A.

A geometric interpretation of the reach is the radius of the smallest Euclidean ball that can be rolled up on the boundary of A, on the outside of A. If A has reach at least $r > 0$, we also say that the complement of A satisfies the r-*rolling ball condition*. In particular, if K is a convex body and its complement has reach $r > 0$, then for each point $x \in \partial K$, there exist $a \in K$ such that $x \in B_d(a, r) \subseteq K$.

For all real numbers $r \in (0, 1]$, we denote by $\mathcal{K}_{d,r}$ the class of all convex bodies in $\mathbb{R}^d$ whose complement has reach at least r and by $\mathcal{K}_{d,r}^{(1)}$ the class of all convex bodies included in $B_d(0, 1)$ whose complement has reach at least r.

Nearly uniform distributions in a convex body with positive reach complement

Let $K \in \mathcal{K}_{d,r}^{(1)}$ and μ be a probability measure with $\mathsf{supp}(\mu) = K$ and $\mu(A) \geq \lambda \mathsf{Vol}_d(A)$, for all Borel sets $A \subseteq K$, where λ is some given positive number. If μ is the uniform probability measure on K, take $\lambda = \kappa_d^{-1}$.

Proposition 1. *The pair (μ, K) satisfies Assumption* 1 *with $\alpha = (d+1)/2$, $L = \frac{2\lambda\kappa_{d-1}r^{\frac{d-1}{2}}}{d+1}$ and $\varepsilon_0 = r$.*

Proof. Let $u \in \mathbb{S}^{d-1}$ and $x^* \in \partial K$ such that $h_K(u) = \langle u, x^* \rangle$. Since K satisfies the r-rolling ball property, $x^* \in B_d(a, r) \subseteq K$ for some $a \in K$. Let $B = B_d(a, r)$. Then, for all $\varepsilon \in [0, r]$,

$$
\begin{aligned}
\mu\big(C_K(u, \varepsilon)\big) &\geq \lambda \mathsf{Vol}_d\big(C_K(u, \varepsilon)\big) \\
&\geq \lambda \mathsf{Vol}_d\big(C_B(u, \varepsilon)\big) \\
&= \lambda \int_0^\varepsilon \big(x(2r - x)\big)^{\frac{d-1}{2}} \kappa_{d-1}\, dx \\
&\geq \lambda \kappa_{d-1} r^{\frac{d-1}{2}} \int_0^\varepsilon x^{\frac{d-1}{2}}\, dx \\
&= \frac{2\lambda\kappa_{d-1}r^{\frac{d-1}{2}}}{d+1} \varepsilon^{\frac{d+1}{2}}.
\end{aligned}
$$
$\square$

Projection of a high dimensional uniform distribution in a convex body

Let $D > d$ be an integer and $\tilde{K} \in \mathcal{K}_{D,r}^{(1)}$. Denote by π_d the orthogonal projection in $\mathbb{R}^D$ onto the first d coordinates and let $K = \pi_d(\tilde{K})$. Then, if we identify $\mathbb{R}^d$ to $\mathbb{R}^d \times \{0\}^{D-d}$, it holds that $K \in \mathcal{K}_{d,r}^{(1)}$. Let μ be the image of the uniform distribution on $\tilde{K}$ by π_d, that is, the distribution of $\pi_d(\tilde{X})$, where $\tilde{X}$ is a uniform random variable in $\tilde{K}$.

Proposition 2. *The pair (μ, K) satisfies Assumption* 1 *with $\alpha = (D+1)/2$, $L = \frac{2\kappa_{D-1}r^{\frac{D-1}{2}}}{\kappa_D(D+1)}$ and $\varepsilon_0 = r$.*

Proof. This proposition follows from a similar argument as in the previous case. Indeed, let $u \in \mathbb{S}^{d-1}$, $x^* \in \partial K$ with $h_K(u) = \langle u, x^* \rangle$ and $a \in K$ such that $x^* \in B_d(a, r) \subseteq K$. Let $\tilde{u}$ be the

(unique) unit vector in R^D such that $u = \pi_d(\tilde{u})$. Let $\tilde{x}^* \in \partial\tilde{K}$ with $h_{\tilde{K}}(\tilde{u}) = \langle \tilde{u}, \tilde{x}^* \rangle$ and $\tilde{a} \in \tilde{K}$ such that $\tilde{x}^* \in B_D(\tilde{a}, r) \subseteq \tilde{K}$. Denote by $\tilde{\mu}$ the uniform distribution in $\tilde{K}$. Then, for all $\varepsilon \in [0, r]$, $\mu(C_K(u, \varepsilon)) = \tilde{\mu}(C_{\tilde{K}}(\tilde{u}, \varepsilon))$ and one can use the same argument as in the end of the proof of Proposition 1. □

More generally, one can consider distributions with densities that have a polynomial decay near the boundary of K.

Densities with polynomial decay near the boundary of K

Let $K \in \mathcal{K}_{d,r}^{(1)}$ and μ be a probability measure with $\mathrm{supp}(\mu) = K$. Assume that μ has a density f with respect to the Lebesgue measure and that for all $x \in K$ with $\rho(x, \partial K) \leq r$, $f(x) \geq C\rho(x, \partial K)^\gamma$, for some real numbers $C > 0$ and $\gamma \geq 0$. For instance, this is satisfied by the previous case with $\gamma = \frac{D-d}{2}$ and $C = r^{\frac{D-d}{2}} \kappa_{D-d}$. Note that the exponent α in Proposition 2, which we proved using geometric arguments, could also be deduced from this fact together with the following proposition.

Proposition 3. *The pair (μ, K) satisfies Assumption 1 where $\alpha = \frac{d+1+2\gamma}{2}$, $L > 0$ is a positive constant that only depends on r, d and γ and $\varepsilon_0 = r$.*

The explicit form of the constant L can be found in the following computation.

Proof. Let $u \in \mathbb{S}^{d-1}$ and $\varepsilon \in [0, r)$. Let $x^* \in \partial K$ with $h_K(u) = \langle u, x^* \rangle$ and $B = B_d(a, r)$ a ball of radius r containing x^* and included in K. Then,

$$\mu\big(C_K(u, \varepsilon)\big) \geq \mu\big(C_B(u, \varepsilon)\big) \geq C \int_{C_B(u,\varepsilon)} \rho(x, \partial K)^\gamma \, dx \geq C \int_{C_B(u,\varepsilon)} \rho(x, \partial B)^\gamma \, dx$$

$$= C(d-2)\omega_{d-2} \int_0^\varepsilon \left(\int_0^{\sqrt{2rt-t^2}} s^{d-2} \big(r - \sqrt{s^2 + (r-t)^2}\big)^\gamma \, ds \right) dt$$

$$\geq \frac{C(d-2)\omega_{d-2}}{2r} \int_0^\varepsilon \left(\int_0^A s^{d-2} \big(A^2 - s^2\big)^\gamma \, ds \right) dt$$

$$\geq \frac{C(d-2)\omega_{d-2}\mathcal{B}(d-1, \gamma+1)}{2r} \int_0^\varepsilon \big(\sqrt{2rt - t^2}\big)^{2\gamma+d-1} \, dt$$

$$\geq \frac{C(d-2)\omega_{d-2}\mathcal{B}(d-1, \gamma+1)r^{(2\gamma+d-3)/2}}{2} \varepsilon^{(2\gamma+d+1)/2},$$

where we identified the unit ball in the hyperplane that is orthogonal to u with $B_{d-1}(0, 1)$, we denoted by $A = \sqrt{2rt - t^2}$ and by $\mathcal{B}(\cdot, \cdot)$ the Beta function. □

Nearly uniform distributions on the boundary of a convex body with positive reach complement

Let $K \in \mathcal{K}_{d,r}^{(1)}$ and μ be a probability measure supported on ∂K with density f with respect to the $(d-1)$-dimensional Hausdorff measure on ∂K, satisfying $f(x) \geq \lambda$, for all $x \in \partial K$ and some $\lambda > 0$. Note that if μ is the uniform distribution on ∂K, then one can take $\lambda = 1/\omega_{d-1}$, since $K \subseteq B_d(0, 1)$.

Proposition 4. *The pair* (μ, K) *satisfies Assumption* 1 *with* $\alpha = (d-1)/2$, $L = \lambda \omega_{d-1} r^{(d-1)/2}$ *and* $\varepsilon_0 = r$.

Proof. Let $u \in \mathbb{S}^{d-1}$ and $0 \leq \varepsilon \leq r$. Then, $\mu(C_K(u, \varepsilon)) \geq \lambda \mathcal{H}_{d-1}(\partial K \cap C_K(u, \varepsilon))$, where $\mathcal{H}_{d-1}$ stands for the $(d-1)$-dimensional Hausdorff measure. Since K satisfies the r-rolling condition, $\mathcal{H}_{d-1}(\partial K \cap C_K(u, \varepsilon)) \geq \mathcal{H}_{d-1}(\partial B \cap C_B(u, \varepsilon))$, where B is any d-dimensional Euclidean ball with radius r. It follows that

$$\mathcal{H}_{d-1}(\partial K \cap C_K(u, \varepsilon)) \geq \mathcal{H}_{d-1}(\partial B \cap C_B(u, \varepsilon))$$

$$\geq \frac{1}{2} \omega_{d-1} r^{d-1} \int_0^{\frac{\varepsilon(2r-\varepsilon)}{r^2}} t^{\frac{d-3}{2}} (1-t)^{-1/2} \, dt$$

$$\geq \frac{1}{2} \omega_{d-1} r^{d-1} \int_0^{\frac{\varepsilon}{r}} t^{\frac{d-3}{2}} \, dt$$

$$\geq \omega_{d-1} r^{\frac{d-1}{2}} \varepsilon^{\frac{d-1}{2}}.$$

In addition, since K is convex and is included in $B_d(0, 1)$, the surface area of its boundary is bounded from above by that of $\mathbb{S}^{d-1}$, that is, ω_{d-1}. □

Nearly uniform distributions in a convex body with general boundary

Let $K \in \mathcal{K}_d^{(1)}$ and let $a \in \mathbb{R}^d$ and $\eta > 0$ such that $B_d(a, \eta) \subseteq K$. The existence of a and η is guaranteed by the fact that the convex body K must have nonempty interior. Let μ be a probability measure with $\mathrm{supp}(\mu) = K$ and that satisfies $\mu(A) \geq \lambda \mathrm{Vol}_d(A)$ for all Borel sets $A \subseteq K$, where $\lambda > 0$.

Proposition 5. *The pair* (μ, K) *satisfies Assumption* 1 *with* $\alpha = d$, $L = \frac{\lambda \kappa_{d-1}}{3\eta^{d-1}}$ *and* $\varepsilon_0 = \eta$.

Proof. Let $u \in \mathbb{S}^{d-1}$ and let $x^* \in \partial K$ such that $h_K(u) = \langle u, x^* \rangle$. Then, for all $\varepsilon \in [0, \eta]$, $C_K(u, \varepsilon) \supseteq C_{\mathsf{cone}}(u, \varepsilon)$, where cone is the smallest cone with apex x^* containing the $(d-1)$-dimensional ball with center a, radius η, in the affine hyperplane containing a, orthogonal to u (cf. Figure 1). Hence, $\mu(C_K(u, \varepsilon)) \geq \mu(C_{\mathsf{cone}}(u, \varepsilon)) \geq \frac{\lambda \varepsilon^d \kappa_{d-1}}{3\eta^{d-1}}$. □
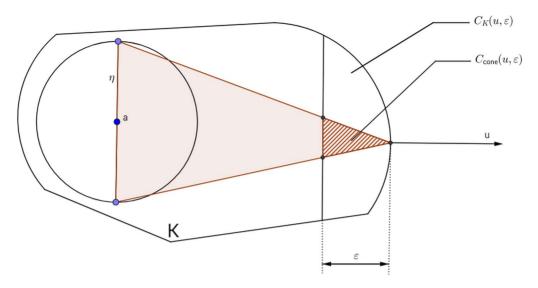
Figure 1. The case of general convex bodies.

4. Rate-optimality of the random polytope

Let n be a positive integer and $X_1, \ldots, X_n$ be random variables mapping the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ onto $\mathbb{R}^D$, for some integer $D \geq 1$. A set-valued statistic is a set of the form $\tilde{K}_n(X_1, \ldots, X_n)$, where $\tilde{K}_n$ maps $(\mathbb{R}^D)^n$ onto the collection of all compact subsets of $\mathbb{R}^d$, such that the function that maps $(\omega, x) \in \Omega \times \mathbb{R}^d$ to 1 if $x \in \tilde{K}_n(X_1(\omega), \ldots, X_n(\omega))$ and is zero otherwise is measurable with respect to the product of $\mathcal{F}$ and the Borel σ-algebra of $\mathbb{R}^d$. For the sake of brevity, we write $\tilde{K}_n$ instead of $\tilde{K}_n(X_1, \ldots, X_n)$. For instance, if $X_1, \ldots, X_n$ are i.i.d. random points, then $\hat{K}_n$ is a set-valued statistic. If a set-valued statistic $\tilde{K}_n$ is aimed to estimate a compact set K which may be unknown in practice, it is called an estimator of K and $\mathsf{d}_\mathsf{H}(\tilde{K}_n, K)$ is a (measurable) random variable. More general definitions of set-valued random variables can be found in [32], but they are not needed for our purposes.

Uniform distribution in a convex body with positive reach complement

Next theorem, together with Corollary 1, establishes rate minimax optimality of $\hat{K}_n$ on the class $\mathcal{K}_{d,r}^{(1)}$, under the Hausdorff metric.

Theorem 3. *Let $r \in (0, 1)$. For all real numbers $q \geq 1$,*

$$\inf_{\tilde{K}_n} \sup_{K \in \mathcal{K}_{d,r}^{(1)}} \mathbb{E}_K\big[\mathsf{d}_\mathsf{H}(\tilde{K}_n, K)^q\big] = \omega\left(\left(\frac{\ln n}{n}\right)^{\frac{2q}{d+1}}\right),$$

where the infimum is taken over all set-valued statistics $\tilde{K}_n$.

Note that for $r = 1$, the class $\mathcal{K}_{d,1}^{(1)}$ reduces to the singleton $\{B_d(0,1)\}$ and the trivial estimator $\tilde{K}_n = B_d(0,1)$ satisfies $\sup_{K \in \mathcal{K}_{1,d}^{(1)}} \mathbb{E}_K[\mathsf{d}_\mathsf{H}(\tilde{K}_n, K)] = 0$.

Proof. Here, we extend the arguments of the proof of [30], Theorem 3, to higher dimensions. This standard method was also exploited in [27]. We fix $R \in (r, 1]$ and set $G_0 = B_d(0, R)$. Let δ be a positive and small enough real number. For $u \in \mathbb{S}^{d-1}$, we define the set $G(u)$ as follows. Let $\eta : \mathbb{R} \to \mathbb{R}$ be a given nonnegative function, twice continuously differentiable, such that $\eta(x) = 0$ for all x of norm at least one and $\max_{x \in \mathbb{R}} \eta(x) = \eta(0) = 1$. For instance, one can take $\eta(x) = e^4 g(2x - 1)g(2 - 2x)$, for all $x \in \mathbb{R}$, where $g(x) = \exp(-1/x)\mathbb{1}_{x>0}$. Let H be the supporting hyperplane of G at the point $Ru \in \partial G_0$. Identify this hyperplane with $\mathbb{R}^{d-1}$, with origin at the point Ru. Then, a parametrization of G_0 is

$$G_0 = \left\{ (t, y) \in H \times \mathbb{R} : \|t\| \le R, \, R - \sqrt{R^2 - \|t\|^2} \le y \le R + \sqrt{R^2 - \|t\|^2} \right\}.$$

We define the set $G(u)$ by modifying the parametrisation of ∂G_0 in a small neighborhood of Ru:

$$G(u) = \left\{ (t, y) \in H \times \mathbb{R} : \|t\| \le R, \right.$$

$$\left. R - \sqrt{R^2 - \|t\|^2} + \alpha \delta^2 \eta\left(\frac{2\|t\|}{R\delta}\right) \le y \le R + \sqrt{R^2 - \|t\|^2} \right\},$$

where α is a positive number that can be tuned independently of δ so that $G(u) \in \mathcal{K}_{d,r}$, for small enough δ. Note that for all $u \in \mathbb{S}^{d-1}$, $G(u) \subseteq G_0$ and

$$\begin{aligned}
\mathsf{Vol}_d\big(G_0 \backslash G(u)\big) &= \alpha \delta^2 \int_{\mathbb{R}^d} \eta\left(\frac{2\|t\|}{R\delta}\right) \mathrm{d}t \\
&= \frac{\alpha \delta^{d+1} R^{d-1}}{2^{d-1}} \int_{\mathbb{R}^d} \eta\big(\|t\|\big) \mathrm{d}t \\
&= c\delta^{d+1},
\end{aligned} \tag{15}$$

for some positive constant c that depends on R and d only. In addition, if $(u, u') \in \mathbb{S}^{d-1} \times \mathbb{S}^{d-1}$ is such that $\|u - u'\| \ge \delta$,

$$\mathsf{d}_\mathsf{H}\big(G(u), G(u')\big) = \alpha \delta^2. \tag{16}$$

Let now $\mathcal{N}$ be a maximal δ-packing of $\mathbb{S}^{d-1}$, that is, a finite subset of $\mathbb{S}^{d-1}$ with maximal cardinality, satisfying that for all $(u, v) \in \mathcal{N}$ with $u \ne v$, $\|u - v\| \ge \delta$. Let us denote by N its cardinality and by $u_1, \dots, u_N$ its elements and, for $j = 1, \dots, N$, set $G_j = G(u_j)$. By a standard argument (see, e.g., [26]), $N \ge c\delta^{-d}$, for some positive constant c. Now, by setting $\delta = (\frac{\ln n}{n})^{\frac{1}{d+1}}$, we apply [27], Lemma 1, to the sets $G_0, G_1, \dots, G_N$, whose assumptions are satisfied due to (15) and (16) and this proves Theorem 3 for $q = 1$. The lower bound for larger values of q is then obtained by Hölder's inequality. $\qquad\square$

Projection of a high dimensional uniform distribution in a convex body with positive reach complement

The same idea as for the previous proof can be carried over in order to prove that $n^{-2/(D+1)}$ is the minimax rate of estimation of the projection of a D-dimensional convex support with positive reach ($D > d$).

Let $D > d$ be an integer. Let $\tilde{K} \in \mathcal{K}_D$ and $K = \pi_d(\tilde{K})$ be the orthogonal projection of $\tilde{K}$ on the first d coordinates. Let $\tilde{X}_1, \ldots, \tilde{X}_n$ be i.i.d. uniform random variables in $\tilde{K}$ and $X_1, \ldots, X_n$ be their orthogonal projections on the first d coordinates. We denote by $\mathbb{E}_{\tilde{K}}^{\mathsf{proj}_d}$ the expectation operator associated with the distribution of $X_1, \ldots, X_n$.

Theorem 4. *Let $D > d$ be a positive integer and $0 < r < 1$. Then, for all real numbers $q \geq 1$,*

$$\inf_{\tilde{K}_n} \sup_{\tilde{K} \in \mathcal{K}_{d,r}^{(1)}} \mathbb{E}_{\tilde{K}}^{\mathsf{proj}_d}\big[\mathsf{d}_{\mathsf{H}}\big(\tilde{K}_n, \pi_d(\tilde{K})\big)^q\big] = \omega\!\left(\left(\frac{\ln n}{n}\right)^{\frac{2q}{D+1}}\right),$$

where the infimum is taken over all set-valued statistics $\tilde{K}_n$.

Proof. Let $R \in (r, 1]$ and $\delta > 0$. Consider a δ-packing $\mathcal{N}$ of the $(d-1)$-dimensional sphere $\mathbb{S}^{d-1} \times \{0\}^{D-d}$. Let $\tilde{K} = B_D(0, R)$ and for $u \in \mathcal{N}$, let $\tilde{K}(u)$ be defined exactly as $G(u)$ in the proof of Theorem 3. Let $K = \pi_d(\tilde{K})$ and $K(u) = \pi_d(\tilde{K}(u))$ for all $u \in \mathcal{N}$. Then, similarly to (15) and (16), $\mathsf{Vol}_d(\tilde{K} \setminus \tilde{K}(u)) = c\delta^{D+1}$ and $\mathsf{d}_{\mathsf{H}}(K(u), K(u')) = \alpha\delta^2$ for all $u, u' \in \mathcal{N}$ with $u \neq u'$. Denote by KL the Kullback–Leibler divergence between probability measures (cf. [48], Definition 2.5). Then, by the data processing inequality (cf. [21], Lemma 2), for all $u \in \mathcal{N}$,

$$\mathsf{KL}(\pi_d\mathbb{P}_{\tilde{K}}, \pi_d\mathbb{P}_{\tilde{K}(u)}) \leq \mathsf{KL}(\mathbb{P}_K, \mathbb{P}_{\tilde{K}(u)}) = \ln\!\left(\frac{\mathsf{Vol}_d(\tilde{K})}{\mathsf{Vol}_d(\tilde{K}(u))}\right)$$

$$\leq \frac{\mathsf{Vol}_d(\tilde{K} \setminus \tilde{K}(u))}{\mathsf{Vol}_d(\tilde{K}(u))} \leq C\delta^{D+1},$$

for some positive constant C that depends on d only and where we denoted by $\pi_d\mathbb{P}_{\tilde{K}}$ (resp. $\pi_d\mathbb{P}_{\tilde{K}(u)}$) the image of the distribution $\mathbb{P}_{\tilde{K}}$ (resp. $\mathbb{P}_{\tilde{K}(u)}$) by the orthogonal projection π_d. Then, Fano's lemma (cf. [48], Corollary 2.6) yields Theorem 4 for $q = 1$, and the result can be obtained for larger values of q by using Hölder's inequality. $\qquad\square$

Uniform distribution in a general convex body

Here, we prove that in the case of a uniform distribution in some convex body $K \in \mathcal{K}_d$ and under no assumption on the boundary of K, the rate that we achieved for polytopes in Section 5, that is, $n^{-1/d}$, cannot be improved in a minimax sense.

Theorem 5. *For all real numbers $q \geq 1$,*

$$\inf_{\tilde{K}_n} \sup_{K \in \mathcal{K}_d^{(1)}} \mathbb{E}_K\big[d_H(\tilde{K}_n, K)^q\big] = \omega\big(n^{-q/d}\big),$$

where the infimum is taken over all set-valued statistics $\tilde{K}_n$.

Proof. This proof is an extension of the proof of [30], Theorem 2. Let $K_0 = \{x = (x_1, \ldots, x_d) \in B_d(0, 1) : x_j \geq 0, \forall j = 1, \ldots, d\}$ and $K_1 = \{x = (x_1, \ldots, x_d) \in B_d(0, 1) : x_j \geq 0, \forall j = 1, \ldots, d, x_1 + \cdots + x_d \geq h\}$, where $h > 0$. It is easy to see that $d_H(K_0, K_1) = h$ and $\mathsf{Vol}_d(G_0 \triangle G_1) = ch^d$, where c is the volume of the d-dimensional simplex. The same argument as in [30], Theorem 2, yields Theorem 5. $\square$

5. Uniform distributions with polytopal supports

Let K be a polytope and denote by V its vertex set. Next, the lemma follows from the fact that a convex function defined on a compact and convex set achieves its maximum at an extreme point of that convex set.

Lemma 2. *Let L be a convex set included in K. Then,*

$$d_H(K, L) = \max_{v \in V} \rho(v, L).$$

It follows that, for $\varepsilon > 0$, $d_H(K, \hat{K}_n) \geq \varepsilon$ if and only if $X_i \notin B_d(v, \varepsilon)$, for some vertex $v \in V$, and for all $i = 1, \ldots, n$. Hence, by a union bound, if $\mathsf{Vol}_d(K) \neq 0$,

$$
\begin{aligned}
\mathbb{P}_K\big[d_H(K, \hat{K}_n) \geq \varepsilon\big] &\leq \sum_{v \in V} \mathbb{P}_K\big[X_i \notin B_d(v, \varepsilon), \forall i = 1, \ldots, n\big] \\
&\leq \sum_{v \in V} \left(1 - \frac{\mathsf{Vol}_d(B_d(v, \varepsilon) \cap K)}{\mathsf{Vol}_d(K)}\right)^n.
\end{aligned}
\tag{17}
$$

In order to get a uniform inequality on a certain class of polytopes, it is natural to restrict the polytope K to have a bounded number a vertices. Otherwise, any convex body could be approximated arbitrarily well by a polytope in the class. Let $p \geq d + 1$ be an integer and let K have no more than p vertices. In addition, suppose that the vertices of K are not too peaked: Assume the existence of $\nu \in (0, 1)$ and $\varepsilon_0 > 0$ such that $\mathsf{Vol}_d(B_d(v, \varepsilon) \cap K) \geq \nu \mathsf{Vol}_d(B_d(v, \varepsilon)) = \nu \varepsilon^d \kappa_d$, for all $v \in V$ and $\varepsilon \in (0, \varepsilon_0)$. Finally, also suppose that $K \subseteq [0, 1]^d$. Then, $\mathsf{Vol}_d(K) \leq 1$ and for $\varepsilon \in (0, \varepsilon_0)$, (17) becomes

$$\mathbb{P}_K\big[d_H(K, \hat{K}_n) \geq \varepsilon\big] \leq p \exp\big(-\nu \kappa_d n \varepsilon^d\big). \tag{18}$$

Denote by $\mathcal{P}(p, \nu, \varepsilon_0)$ the class of all polytopes $K \subseteq [0, 1]^d$ that have no more than p vertices and such that $\mathsf{Vol}_d(B(v, \varepsilon) \cap K) \geq \nu \varepsilon^d \kappa_d$, for all $v \in V$ and $\varepsilon \in (0, \varepsilon_0)$. This last condition is called a *standardness condition*, see [15]. We have the following theorem.

Theorem 6. *Let d and $p \geq d+1$ be positive integers, $\nu \in (0,1)$ and ε_0 be positive real numbers. The following uniform deviation inequality holds*:

$$\sup_{K \in \mathcal{P}(p,\nu,\varepsilon_0)} \mathbb{P}_K\left[n^{1/d}\mathsf{d}_{\mathsf{H}}(K, \hat{K}_n) \geq x \right] \leq p e^{-\nu \kappa_d x^d},$$

for all $x \in [0, \epsilon_0 n^{1/d}]$.

Note that if $x \in (\varepsilon_0 n^{1/d}, \sqrt{d} n^{1/d})$, the right hand side can be replaced by $p e^{-L \varepsilon_0^d n}$ and if $x > \sqrt{d} n^{1/d}$, it can be replaced by zero, since $\hat{K}_n \subseteq K \subseteq [0,1]^d$, yielding $\mathsf{d}_{\mathsf{H}}(\hat{K}_n, K) \leq \sqrt{d}$ almost surely. The next corollary follows directly.

Corollary 2. *Let d and $p \geq d+1$ be positive integers, and $\nu \in (0,1)$ and $\varepsilon_0 > 0$. Then, for all real numbers $q \geq 1$,*

$$\sup_{K \in \mathcal{P}(p,\nu,\varepsilon_0)} \mathbb{E}_K\left[\mathsf{d}_{\mathsf{H}}(\hat{K}_n, K)^q \right] = O\left(n^{\frac{-q}{d}} \right).$$

It is easy to see from its proof that Theorem 5 still holds if the condition "$B_d(a, \eta) \subseteq K$ for some $a \in \mathbb{R}^d$" is added under the supremum, for some small fixed $\eta < 1$. Moreover, together with Proposition 5 in Section 3.3, Corollary 1 entails that

$$\sup_{K \in \mathcal{K}_d^{(1)}: \exists a \in \mathbb{R}^d, B_d(a,\eta) \subseteq K} \mathbb{E}_K\left[\mathsf{d}_{\mathsf{H}}(\tilde{K}_n, K)^q \right] = O\left(\left(\frac{\ln n}{n} \right)^{q/d} \right).$$

The lower and upper bounds differ only by a logarithmic factor. Moreover, by adapting the proof of Theorem 5, one can see that the rate in Corollary 2 is tight and it is the same rate as in Theorem 5. This yields that when μ is a uniform distribution, polytopal supports are the hardest convex bodies to estimate in a minimax sense, with respect to the Hausdorff metric. The contrary happens under the Nikodym metric: polytopal supports are the ones that can be estimated with the highest accuracy and convex supports with smooth boundary are the hardest to estimate, see [10]. In statistical terms, the estimator $\hat{K}_n$ does not adapt to polytopal supports under the Hausdorff metric, unlike under the Nikodym metric [11]. This is very surprising, since estimating K under the Hausdorff metric is the same as estimating h_K under the L^∞ norm. Yet, when K is a polytope, its support function is piecewise linear, so one would expect it to be easier to estimate.

6. Uniform case: A rescaled version of the Hausdorff metric

In [11] and [10], the risk of $\hat{K}_n$ is measured using the Nikodym distance rescaled by the volume of K, namely, $\frac{\mathsf{Vol}_d(\hat{K}_n \triangle K)}{\mathsf{Vol}_d(K)}$. Rescaling with the volume of K allows to consider convex bodies K that are not necessarily included in some given bounded set, see, for example, $B_d(0,1)$. In this section, we introduce a scaled version of d_{H} which allows to remove the assumption $K \subseteq B_d(0,1)$. In order to avoid unnecessary technicalities, we only deal with uniform distributions.

Define the pseudo-distance d_L between two sets K and K' as $d_L(K, K') = \inf\{\varepsilon \geq 0 : \exists a \in \mathbb{R}^d, a + (1 + \varepsilon)^{-1}(K - a) \subseteq K' \subseteq a + (1 + \varepsilon)(K - a)\}$. It is easy to see that $\ln(1 + d_L)$ is a proper distance. If $d_L(K, K')$ is small, then $\ln(1 + d_L(K, K')) \approx d_L(K, K')$ and it is simpler, in our next results, to deal with $d_L(K, K')$. One advantage of d_L over d_H is that it is invariant under the action of any invertible affine map T: $d_L(K, K') = d_L(T(K), T(K'))$, for all compact sets K, K'. Moreover, d_L the well studied Banach–Mazur multiplicative distance. The Banach–Mazur distance between two convex bodies K and K' is defined as $d_{BM}(K, K') = \inf\{\lambda \geq 1 : \exists u, v \in \mathbb{R}^d, \exists T \in GL(d), K + u \subseteq T(K' + v) \subseteq \lambda(K + u)\}$, where $GL(d)$ stands for the set of all linear isomorphisms in $\mathbb{R}^d$ (see, e.g., [44]). It is a multiplicative pseudo-distance, in the sense that $\ln d_{BM}$ satisfies the triangle inequality. The following lemma holds.

Lemma 3. *For all* $K, K' \in \mathcal{K}_d$, $d_{BM}(K, K') \leq (1 + d_L(K, K'))^2$.

Proof. Let $K, K' \in \mathcal{K}_d$. Let $\varepsilon > d_L(K, K')$. Then, there exists $a \in \mathbb{R}^d$ for which $(1 + \varepsilon)^{-1}(K - a) \subseteq K' - a \subseteq (1 + \varepsilon)(K - a)$, yielding $K - a \subseteq (1 + \varepsilon)(K' - a) \subseteq (1 + \varepsilon)^2(K - a)$, that is, $(1 + \varepsilon)^2 \geq d_{BM}(K, K')$. Lemma 3 follows by letting ε go to $d_L(K, K')$. $\qquad\square$

In particular, $\ln d_{BM}(K, K') \leq 2d_L(K, K')$, for all convex bodies K, K'. The following lemma, where we denote by conv the convex hull, is straightforward.

Lemma 4. *Let* $T : \mathbb{R}^d \to \mathbb{R}^d$ *be an invertible affine map. Let* $K \in \mathcal{K}_d$, $X_1, \ldots, X_n$ *be i.i.d. uniform random variables in* K *and* $Y_1, \ldots, Y_n$ *be i.i.d. uniform random variables in* TK. *Then,* $d_L(K, \text{conv}(X_1, \ldots, X_n))$ *and* $d_L(TK, \text{conv}(Y_1, \ldots, Y_n))$ *have the same distribution.*

As a consequence, working with d_L instead of d_H allows to remove the assumption $K \subseteq B_d(0, 1)$.

6.1. Convex supports with positive reach complement

For $0 < r \leq 1$, denote by $\tilde{\mathcal{K}}_{d,r}$ the class of all convex bodies $K \in \mathcal{K}_d$ satisfying the $\tilde{r}$-rolling condition, with $\tilde{r} = r(\frac{\text{Vol}_d(K)}{\kappa_d})^{1/d}$. By introducing the number $\tilde{r}$ in this definition, we ensure that all convex bodies that are similar to K, that is, obtained by rescaling, translating, rotating or reflecting K, will be in $\tilde{\mathcal{K}}_{d,r}$ as soon as K is.

Theorem 7. *Let* d *and* n *be positive integers and* $r \in (0, 1]$. *Set* $c = \frac{2d}{(d+1)\tau}$, $\tau_3 = \max(1, \frac{2^{(d+1)/2}d\kappa_d}{r^d\kappa_{d-1}})$, $a_n = (\frac{\tau_3 \ln n}{n})^{\frac{2}{d+1}}$ *and* $b_n = n^{\frac{-2}{d+1}}$. *Then, the random polytope* $\hat{K}_n$ *satisfies*

$$\sup_{K \in \tilde{\mathcal{K}}_{d,r}} \mathbb{P}_K\big[d_L(\hat{K}_n, K) \geq 4a_n + 4b_n x\big] \leq 6^d \exp\big(-cx^{\frac{d+1}{2}}\big),$$

for all $x \geq 0$ *with* $a_n + b_n x \leq 1/4$.

Proof. The proof of this theorem follows the same lines as that of Theorem 1. Let $d \geq 1$, $r \in (0, 1]$ and $K \in \tilde{\mathcal{K}}_{d,r}$. Let $\tilde{r} = r(\frac{\mathsf{Vol}_d(K)}{\kappa_d})^{1/d}$. Let n be a positive integer and $X, X_1, X_2, \ldots, X_n$ be $n + 1$ i.i.d. random points uniformly distributed in K. Without loss of generality, one can assume that 0 is an interior point of K and $B_d(0, \tilde{r}) \subseteq K$. Indeed, since $K \in \tilde{\mathcal{K}}_{d,r}$, there exists a point $a \in K$ such that $B_d(a, \tilde{r}) \subseteq K$ and K could be replaced with $K - a$ without changing the distribution of $\mathsf{d}_\mathsf{L}(\hat{K}_n, K)$. For all $u \in \mathbb{S}^{d-1}$, $h_K(u) \geq \tilde{r}$ and

$$
\begin{aligned}
\mathsf{d}_\mathsf{L}(\hat{K}_n, K) &\leq \inf\{\varepsilon > 0 : (1+\varepsilon)^{-1} K \subseteq \hat{K}_n \subseteq (1+\varepsilon)K\} \\
&= \inf\{\varepsilon > 0 : (1+\varepsilon)^{-1} K \subseteq \hat{K}_n\} \\
&= \inf\{\varepsilon > 0 : \forall u \in \mathbb{S}^{d-1}, (1+\varepsilon)^{-1} h_K(u) \subseteq h_{\hat{K}_n}(u)\}.
\end{aligned}
\tag{19}
$$

Let $\varepsilon \in (0, 1/4]$ and $u \in \mathbb{S}^{d-1}$. By the $\tilde{r}$-rolling condition, we have

$$
\begin{aligned}
\mathsf{Vol}_d\big(C_K(u, \varepsilon\tilde{r}/2)\big) &= \int_0^{\varepsilon\tilde{r}/2} \sqrt{x(2\tilde{r} - x)}^{d-1} \kappa_{d-1} \, \mathrm{d}x \geq \kappa_{d-1} \tilde{r}^{\frac{d-1}{2}} \int_0^{\varepsilon\tilde{r}/2} x^{\frac{d-1}{2}} \, \mathrm{d}x \\
&\geq \frac{2\kappa_{d-1}\tilde{r}^d}{2^{\frac{d+1}{2}}(d+1)} \varepsilon^{\frac{d+1}{2}} = \frac{2\kappa_{d-1} r^d \mathsf{Vol}_d(K)}{2^{\frac{d+1}{2}} \kappa_d (d+1)} \varepsilon^{\frac{d+1}{2}}.
\end{aligned}
$$

Hence, if $\eta = \frac{\kappa_{d-1} r^d}{2^{\frac{d-1}{2}} \kappa_d (d+1)}$,

$$
\begin{aligned}
\mathbb{P}\left[\frac{h_{\hat{K}_n}(u)}{h_K(u)} \leq (1+\varepsilon)^{-1}\right] &\leq \mathbb{P}\left[\frac{h_{\hat{K}_n}(u)}{h_K(u)} \leq 1 - \frac{\varepsilon}{2}\right] \\
&= \mathbb{P}\left[h_{\hat{K}_n}(u) \leq h_K(u) - \frac{\varepsilon h_K(u)}{2}\right] \\
&\leq \mathbb{P}\left[h_{\hat{K}_n}(u) \leq h_K(u) - \frac{\varepsilon\tilde{r}}{2}\right] \\
&= \left(1 - \frac{\mathsf{Vol}_d(C_K(u, \varepsilon\tilde{r}/2))}{\mathsf{Vol}_d(K)}\right)^n \\
&\leq \exp\left(-\eta n \varepsilon^{\frac{d+1}{2}}\right).
\end{aligned}
\tag{20}
$$

By positive homogeneity of support functions (see (1)), (20) holds for all $u \in \mathbb{R}^d \setminus \{0\}$.

Let $\delta = \varepsilon/2$. Recall that $K^\circ \in \mathcal{K}_d$, since we have assumed that $0 \in \mathrm{int}(K)$, and that $0 \in \mathrm{int}(K^\circ)$ since K is bounded. It is easy to adapt Lemma 1 and prove the existence of $\mathcal{N}_\delta \subseteq \partial(K^\circ)$, with cardinality at most $(3/\delta)^d$, such that every $u \in \partial(K^\circ)$ can be written as $u = u_0 - \sum_{j=1}^\infty \delta_j u_j$ with $0 \leq \delta_j \leq \delta^j$ and $u_j \in \mathcal{N}_\delta$, for all $j \geq 1$. Then, using (20) together with the union bound yields that with probability at least $1 - (\frac{6}{\varepsilon})^d \exp(-\eta n \varepsilon^{\frac{d+1}{2}})$,

$$
\frac{h_{\hat{K}_n}(u)}{h_K(u)} > (1+\varepsilon)^{-1}
\tag{21}
$$

simultaneously for all $u \in \mathcal{N}_\delta$. Assume that (21) holds and let $u \in \partial(K^\circ)$. Write $u = u_0 - \sum_{j=1}^{\infty} \varepsilon_j u_j$, as above. By (1), (2) and (3),

$$h_{\hat{K}_n}(u) \geq h_{\hat{K}_n}(u_0) - \sum_{j=1}^{\infty} \delta_j h_{\hat{K}_n}(u_j). \tag{22}$$

In addition, for all $j \geq 1$, $(1+\varepsilon)^{-1} h_K(u_j) \leq h_{\hat{K}_n}(u_j) \leq h_K(u_j)$. Since $0 \in \text{int}(K)$ and $u_j \in \partial(K^\circ)$, $h_K(u_j) = \|u_j\|_{K^\circ} = 1$ and the following is true, for all $u \in \partial(K^\circ)$:

$$h_{\hat{K}_n}(u) > (1+\varepsilon)^{-1} - \sum_{j=1}^{\infty} \delta^j$$

$$\geq (1+\varepsilon)^{-1} - \frac{\delta}{1-\delta} \geq 1 - 2\varepsilon. \tag{23}$$

Since $0 \in \text{int}(K^\circ)$, for all $u \in \mathbb{S}^{d-1}$, there exists $\lambda > 0$ for which $\lambda u \in \partial(K^\circ)$, yielding

$$\frac{h_{\hat{K}_n}(u)}{h_K(u)} = \frac{h_{\hat{K}_n}(\lambda u)}{h_K(\lambda u)} > 1 - 2\varepsilon \geq (1 + 4\varepsilon)^{-1},$$

since $\varepsilon \leq 1/4$. Hence, with probability at least $1 - (\frac{6}{\varepsilon})^d \exp(-\eta n \varepsilon^{\frac{d+1}{2}})$, $\frac{h_{\hat{K}_n}(u)}{h_K(u)} > (1 + 4\varepsilon)^{-1}$ simultaneously for all $u \in \mathbb{S}^{d-1}$, yielding $\mathsf{d}_\mathsf{L}(\hat{K}_n, K) < 4\varepsilon$. Therefore, $\mathbb{P}[\mathsf{d}_\mathsf{L}(\hat{K}_n, K) \geq 4\varepsilon] \leq 6^d \exp(-\eta n \varepsilon^{\frac{d+1}{2}} - d \ln \varepsilon)$, which yields Theorem 7 by setting $\varepsilon = a_n + b_n x$. $\square$

6.2. General convex supports

Now, we consider any $K \in \mathcal{K}_d$ and let μ be the uniform probability measure in K. By John's ellipsoid theorem (see [31] for instance), there exists an invertible affine map $T : \mathbb{R}^d \to \mathbb{R}^d$ and $a \in \mathbb{R}^d$ such that $B_d(a, 1/(2d)) \subseteq TK \subseteq B_d(0, 1/2)$. Necessarily, $a \in B_d(0, 1/2)$, hence, up to a translation, one can assume that $B_d(0, 1/(2d)) \subseteq TK \subseteq B_d(0, 1)$ without loss of generality. Therefore, by Lemma 4, we can even assume that $B_d(0, 1/(2d)) \subseteq K \subseteq B_d(0, 1)$, without loss of generality. Hence, using Proposition 5, one can adapt Theorem 1 with $\alpha = d$ in order to prove the following result.

Theorem 8. *Let* $a_n = (\frac{\ln n}{n})^{1/d}$, $b_n = n^{-1/d}$ *and* $L = \frac{(2d)^{d-1}\kappa_{d-1}}{3\kappa_d}$. *Then, the random polytope* $\hat{K}_n$ *satisfies*

$$\sup_{K \in \mathcal{K}_d} \mathbb{P}_K\big[\mathsf{d}_\mathsf{L}(\hat{K}_n, K) \geq 8da_n + 8db_n x\big] \leq 12^d \exp\big(-Lx^d\big),$$

for all $x \geq 0$ *with* $a_n + b_n x \leq 1/(16d)$.

Proof. Let $K \in \mathcal{K}_d$. As mentioned above, one can assume $B_d(0, 1/(2d)) \subseteq K \subseteq B_d(0, 1)$, without loss of generality. By (19), for all $\varepsilon \in (0, 1/2]$,

$$\mathbb{P}_K\big[\mathsf{d}_\mathsf{L}(\hat{K}_n, K) \geq \varepsilon\big] \leq \mathbb{P}_K\big[\exists u \in \mathbb{S}^{d-1}, h_{\hat{K}_n} \leq (1+\varepsilon)^{-1} h_K(u)\big]$$

$$\leq \mathbb{P}_K\big[\exists u \in \mathbb{S}^{d-1}, h_{\hat{K}_n} \leq h_K(u) - \varepsilon/(4d)\big]$$

$$\leq \mathbb{P}_K\big[\mathsf{d}_\mathsf{H}(K, \hat{K}_n) \geq \varepsilon/(4d)\big].$$

Take $\varepsilon = 8d(a_n + b_n x)$, where a_n and b_n are defined in the statement of Theorem 8. Then, the conclusion follows directly from Theorem 1 and Proposition 5 with $\eta = 1/(2d)$, $\alpha = d$, $L = \frac{(2d)^{d-1}\kappa_{d-1}}{3\kappa_d}$ and by noting that $L \geq 1$ (so $\tau_1 = 1$), by [36], Lemma 2.2. $\qquad \square$

7. Random polytopes in high dimension

In high dimension, the random polytope $\hat{K}_n$ is hard to compute. Indeed, it requires $O(n^{d/2})$ operations (see, e.g., [14]), which becomes very costly when both n and d are large. The random polytope can be written using infinitely many constraints: $\hat{K}_n = \{x \in \mathbb{R}^d : \langle u, x \rangle \leq \max_{1 \leq i \leq n} \langle u, X_i \rangle, \forall u \in \mathbb{S}^{d-1}\}$. In this section, we show that if the parameter α in Assumption 1 is known, then it is possible to significantly reduce the computational cost of $\hat{K}_n$ by taking only a finite subset of the linear constraints that define $\hat{K}_n$.

For an integer $M \geq 1$, let $U_1, \ldots, U_M$ be i.i.d. random unit vectors and let

$$\hat{K}_{n,M} = \Big\{x \in \mathbb{R}^d : \langle U_j, x \rangle \leq \max_{1 \leq i \leq n} \langle U_j, X_i \rangle, \forall j = 1, \ldots, M\Big\}.$$

Then, the following holds.

Theorem 9. *Let $K \in \mathcal{K}_d$ and $X_1, \ldots, X_n$ be i.i.d. random points in K, with some probability measure μ. Let Assumption 1 hold. Set $M = Cn^{\frac{d-1}{\alpha}} (\ln n)^{\frac{\alpha-d+1}{\alpha}}$, where $C > 2\alpha^{-1}d^2 8^{\frac{d-1}{2}}$. Then,*

$$\mathsf{d}_\mathsf{H}(\hat{K}_{n,M}, K) = O_{\mathbb{P}}\bigg(\Big(\frac{\ln n}{n}\Big)^{\frac{1}{\alpha}}\bigg).$$

Proof. For $j = 1, \ldots, M$, let $\hat{h}_j = \max_{1 \leq i \leq n} \langle U_j, X_i \rangle = h_{\hat{K}_n}(U_j)$ and $h_j = h_K(U_j)$. Since $K \in \mathcal{K}_d$, K has nonempty interior and it is bounded. Thus, let $R > r > 0$ and $a \in \mathbb{R}^d$ such that $B(a, r) \subseteq G \subseteq B(a, R)$. Let $\eta \in (0, r)$ and $\varepsilon \in (0, 1)$. Let $\mathcal{A}$ be the event: "$|\hat{h}_j - h_j| \leq \eta$, for all $j = 1, \ldots, M$" and $\mathcal{B}$ be the event "The collection $\{U_1, \ldots, U_M\}$ is an ε-net of $\mathbb{S}^{d-1}$".

Using (4) together with the union bound,

$$\mathbb{P}[\mathcal{A}^\complement] \leq M e^{-Ln\eta^\alpha}. \tag{24}$$

By [12], Lemma 10,

$$\mathbb{P}[\mathcal{B}^\complement] \leq 6^d \exp\bigg(-\frac{M\varepsilon^{d-1}}{2d8^{\frac{d-1}{2}}} + d \ln\Big(\frac{1}{\varepsilon}\Big)\bigg). \tag{25}$$

Assume that both $\mathcal{A}$ and $\mathcal{B}$ are satisfied. Then, by [12], Lemma 7, $\mathsf{d}_\mathsf{H}(\hat{K}_n, K) \leq \frac{\eta R}{r}\frac{1+\eta/r}{1-\eta/r} + \frac{2R\varepsilon}{1-\varepsilon}$. As a consequence of (24) and (25), this yields

$$
\begin{aligned}
&\mathbb{P}\left[\mathsf{d}_\mathsf{H}(\hat{K}_{n,M}, K) > \frac{\eta R}{r}\frac{1+\eta/r}{1-\eta/r} + \frac{2R\varepsilon}{1-\varepsilon}\right] \\
&\qquad \leq M e^{-Ln\eta^\alpha} + 6^d \exp\left(-\frac{M\varepsilon^{d-1}}{2d8^{\frac{d-1}{2}}} + d\ln\left(\frac{1}{\varepsilon}\right)\right).
\end{aligned}
\tag{26}
$$

Setting $\varepsilon = (\frac{\ln n}{n})^{1/\alpha}$ and $\eta = (\frac{x\ln n}{n})^{1/\alpha}$, for some positive number x, ensuring that $\eta < r$, completes the proof of Theorem 9. $\qquad\square$

As a corollary, the random polytope $\hat{K}_n$ can be approximated by $\hat{K}_{n,M}$:

Corollary 3. *Let the assumptions of Theorem* 9 *hold. Then,*

$$
\mathsf{d}_\mathsf{H}(\hat{K}_{n,M}, K) = O_\mathbb{P}\left(\left(\frac{\ln n}{n}\right)^{\frac{1}{\alpha}}\right).
$$

Computing $\hat{K}_{n,M}$ requires sampling M uniform random unit vectors (which can be done by sampling standard Gaussian vectors and rescaling them) and checking nM linear inequalities. Hence, the overall computation cost of computing $\hat{K}_{n,M}$ is $O(nMd)$. For instance, when μ is the uniform distribution in a general convex body K, we saw previously that α can be chosen as $\alpha = d$. Hence, Corollary 3 shows that $\hat{K}_n$ can approximated in $O(d^2 8^{\frac{d-1}{2}} n^{2-1/d}(\ln n)^{1/d})$ operations, without affecting the rate of its statistical accuracy. If μ is the uniform distribution in a convex body whose complement has positive reach, then $\hat{K}_n$ can be approximated in $O(d^2 8^{\frac{d-1}{2}} n^{\frac{3d-1}{d+1}}(\ln n)^{-\frac{d-3}{d+1}})$ operations. In both cases, the computation cost is significantly better than that of $\hat{K}_n$, which is $O(n^{d/2})$. Note that the computation cost of $\hat{K}_{n,M}$ is still exponential in the dimension, through the multiplicative constants.

References

[1] Affentranger, F. and Wieacker, J.A. (1991). On the convex hull of uniform random points in a simple *d*-polytope. *Discrete Comput. Geom.* **6** 291–305. MR1098810
[2] Baldin, N. and Reiß, M. (2016). Unbiased estimation of the volume of a convex body. *Stochastic Process. Appl.* **126** 3716–3732. MR3565474
[3] Bárány, I. (1989). Intrinsic volumes and *f*-vectors of random polytopes. *Math. Ann.* **285** 671–699. MR1027765
[4] Bárány, I. (2005). Corrigendum: "Random polytopes in smooth convex bodies" [*Mathematika* **39** (1992) 81–92 MR1176473]. *Mathematika* **51** 31.
[5] Bárány, I. and Buchta, C. (1993). Random polytopes in a convex polytope, independence of shape, and concentration of vertices. *Math. Ann.* **297** 467–497. MR1245400
[6] Bárány, I., Fodor, F. and Vígh, V. (2010). Intrinsic volumes of inscribed random polytopes in smooth convex bodies. *Adv. in Appl. Probab.* **42** 605–619. MR2779551

[7] Bárány, I. and Larman, D.G. (1988). Convex bodies, economic cap coverings, random polytopes. *Mathematika* **35** 274–291. MR0986636

[8] Böröczky, K.J., Fodor, F., Reitzner, M. and Vígh, V. (2009). Mean width of random polytopes in a reasonably smooth convex body. *J. Multivariate Anal.* **100** 2287–2295. MR2560369

[9] Böröczky, K.J. Jr., Hoffmann, L.M. and Hug, D. (2008). Expectation of intrinsic volumes of random polytopes. *Period. Math. Hungar.* **57** 143–164. MR2469601

[10] Brunel, V.-E. Uniform deviation and moment inequalities for random polytopes with general densities in arbitrary convex bodies. Preprint. Available at arXiv:1704.01620.

[11] Brunel, V.-E. (2016). Adaptive estimation of convex and polytopal density support. *Probab. Theory Related Fields* **164** 1–16. MR3449384

[12] Brunel, V.-E. (2018). Concentration of the empirical level sets of Tukey's halfspace depth. Probability and Related Fields. To appear. Available at arXiv:1605.09456.

[13] Buchta, C. and Müller, J. (1984). Random polytopes in a ball. *J. Appl. Probab.* **21** 753–762. MR0766813

[14] Chazelle, B. (1993). An optimal convex hull algorithm in any fixed dimension. *Discrete Comput. Geom.* **10** 377–409. MR1243335

[15] Cuevas, A. (2009). Set estimation: Another bridge between statistics and geometry. *Bol. Estad. Investig. Oper.* **25** 71–85. MR2750781

[16] Dümbgen, L. and Walther, G. (1996). Rates of convergence for random approximations of convex sets. *Adv. in Appl. Probab.* **28** 384–393. MR1387882

[17] Dwyer, R.A. (1988). On the convex hull of random points in a polytope. *J. Appl. Probab.* **25** 688–699. MR0967853

[18] Efron, B. (1965). The convex hull of a random set of points. *Biometrika* **52** 331–343. MR0207004

[19] Fresen, D.J. and Vitale, R.A. (2014). Concentration of random polytopes around the expected convex hull. *Electron. Commun. Probab.* **19** 8. MR3254738

[20] Gayraud, G. (1997). Estimation of functionals of density support. *Math. Methods Statist.* **6** 26–46. MR1456645

[21] Gerchinovitz, S., Ménard, P. and Stoltz, G. (2017). Fano's inequality for random variables. Preprint. Available at arXiv:1702.05985.

[22] Glasauer, S. and Schneider, R. (1996). Asymptotic approximation of smooth convex bodies by polytopes. *Forum Math.* **8** 363–377. MR1387701

[23] Groemer, H. (1974). On the mean value of the volume of a random polytope in a convex set. *Arch. Math. (Basel)* **25** 86–90. MR0341286

[24] Guntuboyina, A. (2012). Optimal rates of convergence for convex set estimation from support functions. *Ann. Statist.* **40** 385–411. MR3014311

[25] Hug, D. (2013). Random polytopes. In *Stochastic Geometry, Spatial Statistics and Random Fields. Lecture Notes in Math.* **2068** 205–238. Springer, Heidelberg. MR3059649

[26] Kolmogorov, A.N. and Tihomirov, V.M. (1959). ε-entropy and ε-capacity of sets in function spaces. *Uspekhi Mat. Nauk* **14** 3–86. MR0112032

[27] Korostelëv, A.P., Simar, L. and Tsybakov, A.B. (1995). Efficient estimation of monotone boundaries. *Ann. Statist.* **23** 476–489. MR1332577

[28] Korostelev, A.P., Simar, L. and Tsybakov, A.B. (1995). On estimation of monotone and convex boundaries. *Publ. Inst. Stat. Univ. Paris* **39** 3–18. MR1744393

[29] Korostelëv, A.P. and Tsybakov, A.B. (1993). *Minimax Theory of Image Reconstruction. Lecture Notes in Statistics* **82**. New York: Springer. MR1226450

[30] Korostelëv, A.P. and Tsybakov, A.B. (1994). Asymptotic efficiency of the estimation of a convex set. *Problemy Peredachi Informatsii* **30** 33–44. MR1310060

[31] Leichtweiss, K. (1959). Über die affine Exzentrizität konvexer Körper. *Arch. Math.* **10** 187–199. MR0107843

[32] Molchanov, I. (2013). Foundations of stochastic geometry and theory of random sets. In *Stochastic Geometry*, *Spatial Statistics and Random Fields*. *Lecture Notes in Math.* **2068** 1–20. Springer, Heidelberg. MR3059643

[33] Moore, M. (1984). On the estimation of a convex set. *Ann. Statist.* **12** 1090–1099. MR0751296

[34] Müller, J.S. (1989). On the mean width of random polytopes. *Probab. Theory Related Fields* **82** 33–37. MR0997430

[35] Pateiro-Lopez, B. (2008). Set estimation under convexity type restrictions. Ph.D. Thesis, University of Santiago de Compostela.

[36] Reisner, S., Schütt, C. and Werner, E. (2001). Dropping a vertex or a facet from a convex polytope. *Forum Math.* **13** 359–378. MR1831090

[37] Reitzner, M. (2003). Random polytopes and the Efron–Stein jackknife inequality. *Ann. Probab.* **31** 2136–2166. MR2016615

[38] Reitzner, M. (2004). Stochastic approximation of smooth convex bodies. *Mathematika* **51** 11–29. MR2220208

[39] Reitzner, M. (2010). Random polytopes. In *New Perspectives in Stochastic Geometry* 45–76. Oxford Univ. Press, Oxford. MR2654675

[40] Rényi, A. and Sulanke, R. (1963). Über die konvexe Hülle von *n* zufällig gewählten Punkten. *Z. Wahrsch. Verw. Gebiete* **2** 75–84. MR0156262

[41] Rényi, A. and Sulanke, R. (1964). Über die konvexe Hülle von *n* zufällig gewählten Punkten. II. *Z. Wahrsch. Verw. Gebiete* **3** 138–147. MR0169139

[42] Schneider, R. (1987). Approximation of convex bodies by random polytopes. *Aequationes Math.* **32** 304–310. MR0900710

[43] Schneider, R. (2008). Recent results on random polytopes. *Boll. Unione Mat. Ital.* (9) **1** 17–39. MR2387995

[44] Schneider, R. (2014). *Convex Bodies*: *The Brunn–Minkowski Theory*, expanded ed. *Encyclopedia of Mathematics and Its Applications* **151**. Cambridge: Cambridge Univ. Press. MR3155183

[45] Schneider, R. and Wieacker, J.A. (1980). Random polytopes in a convex body. *Z. Wahrsch. Verw. Gebiete* **52** 69–73. MR0568260

[46] Schütt, C. (1994). Random polytopes and affine surface area. *Math. Nachr.* **170** 227–249. MR1302377

[47] Thäle, C. (2008). 50 years sets with positive reach – a survey. *Surv. Math. Appl.* **3** 123–165. MR2443192

[48] Tsybakov, A.B. (2009). *Introduction to Nonparametric Estimation*. *Springer Series in Statistics*. New York: Springer. Revised and extended from the 2004 French original, Translated by Vladimir Zaiats. MR2724359

[49] Vu, V.H. (2005). Sharp concentration of random polytopes. *Geom. Funct. Anal.* **15** 1284–1318. MR2221249