Discrete Comput Geom (2011) 46: 29–47
DOI 10.1007/s00454-011-9328-x

A Polynomial Number of Random Points Does Not Determine the Volume of a Convex Body

Ronen Eldan

Received: 9 March 2009 / Revised: 28 November 2010 / Accepted: 17 December 2010 /
Published online: 19 February 2011

Abstract We show that there is no algorithm which, provided a polynomial number of random points uniformly distributed over a convex body in $\mathbb{R}^n$, can approximate the volume of the body up to a constant factor with high probability.

Keywords Convex bodies · Random point oracle · Volume computing algorithms · Sampling convex bodies · Distribution of mass on convex bodies

1 Introduction

Volume-related properties of high-dimensional convex bodies is one of the main topics of convex geometry in research today. Naturally, calculating or approximating the volume of a convex body is an important problem. Starting from the 1980s, several works have been made in the area of finding a fast algorithm for computing the volume of a convex body (see, for example, [1–3, 8, 9], and references therein).

These algorithms usually assume that the convex body $K \subset \mathbb{R}^n$ is given by a certain oracle. An oracle is a "black box" which provides the algorithm some information about the body. One example of an oracle is the *membership oracle*, which, given a point $x \in \mathbb{R}^n$, answers either "$x \in K$" or "$x \notin K$". Another example is the *random point oracle*, which generates random points uniformly distributed over K.

All volume computing algorithms, known to the author, that appear in the literature use the membership oracle. This note deals with a question asked by L. Lovász about

Communicated by L. Lovasz.

The author is partially supported by the Israel Science Foundation and by a Farajun Foundation Fellowship.

R. Eldan (✉)
School of Mathematical Sciences, Tel-Aviv University, Tel-Aviv 69978, Israel
e-mail: roneneldan@gmail.com

the random point oracle. It has been an open problem for a while whether or not it is possible to find a fast algorithm which computes the volume of K, provided that there is access to the random point oracle [4, 7].

We answer this question negatively. In order to formulate our main result, we begin with some definitions.

An algorithm which uses the random point oracle is a (possibly randomized) function whose input is a finite sequence of random points generated according to the uniform measure on K and whose output is a number, which is presumed to be an approximation for the volume of K. The complexity of the algorithm will be defined by the length of the sequence of random points. We are interested in the existence of algorithms with complexity which depends polynomially on the dimension n.

We say that an algorithm is correct up to C with probability p if for any $K \subset \mathbb{R}^n$, given a sequence of random points from K, the output of the algorithm is between $\frac{Vol(K)}{C}$ and $C\,Vol(K)$, with probability at least p.

We prove the following theorem:

Theorem 1 *There do not exist constants $C, p, \kappa > 0$ such that for any dimension n, there exists an algorithm with complexity $O(n^\kappa)$ which is correct in estimating the volume of convex bodies in $\mathbb{R}^n$ up to C with probability p.*

It is important to emphasize that this result is not a result in complexity theory. In this note we show that a polynomial number of points actually does not contain enough information to estimate the volume, regardless of the number of calculations, and hence, it is of information-theoretical nature.

For convex geometers, the main point in this study may be the additional information on volume distribution in convex bodies it provides. We suggest the reader to look this result in view of the recent results concerning the distribution of mass in convex bodies. In particular, results regarding thin-shell concentration and the Central Limit Theorem for Convex bodies, proved in the general case by B. Klartag, show that essentially all of the mass of an isotropic convex body K is contained in a very thin-shell around the origin and that almost all of the marginals are approximately Gaussian. This may suggest that, in some way, all convex bodies, when neglecting a small portion of the mass, behave more or less the same as a Euclidean ball in many senses. Philosophically, one can also interpret these results as follows: provided a small number of points from a logarithmically concave measure, one cannot distinguish it from a spherically symmetric measure. For definitions and results, see [5]. One of the main stages of our proof is to show that one cannot distinguish between the uniform distribution over certain convex bodies, which are geometrically far from a Euclidean ball, and some spherically symmetric distribution, when the number of sample points is at most polynomially large.

Here is a more quantitative formulation of what we prove:

Theorem 2 *There exist a constant $\varepsilon > 0$ and a number $N \in \mathbb{N}$ such that for all $n > N$, there does not exist an algorithm whose input is a sequence of length e^{n^ε} of points generated randomly according to the uniform measure in a convex body*

$K \subset \mathbb{R}^n$, which determines Vol(K) up to e^{n^ε} with probability more than e^{-n^ε} to be correct.

Remark After showing that the volume of a convex body cannot be approximated, one may further ask: what about an algorithm that estimates the *volume radius* of a convex body, defined by $VolRad(K) = Vol(K)^{\frac{1}{n}}$? A proof which shows that it is also impossible has to be far more delicate than our proof. For example, under the hyperplane conjecture, it is easy to estimate the volume radius of a convex body up to some $C > 0$.

One may also compare this result to the two following related results: a recent result of Goyal and Rademacher [4] shows that in order to *learn* a convex body, one needs at least $2^{c\sqrt{\frac{n}{\varepsilon}}}$ random points. Learning a convex body roughly means finding a set having at most ε relative symmetric difference with the actual body (see [4]). Klivans, O'Donnel, and Servedio [6] show that any convex body can be agnostically learned with respect to the Gaussian distribution using $2^{O(\sqrt{n})}$ labeled Gaussian samples.

The general idea of the proof is as follows. Let $\{K_\alpha\}_{\alpha \in I_1}$ and $\{K_\alpha\}_{\alpha \in I_2}$ be two families of convex bodies. For $i = 1, 2$, a probability measure μ_i on the set of indices I_i induces a random construction of convex bodies, which in turn induces a probability measure P_i on the set of sequences of points in $\mathbb{R}^n$ in the following simple way: first generate an index α according to μ_i, and then generate a sequence of N uniformly distributed random samples from K_α.

In the proof we will define two distinct random constructions of convex bodies $K_i = (\{K_\alpha\}_{\alpha \in I_i}, \mu_i),\ i = 1, 2$, such that:

1. For every $\alpha_1 \in I_1$ and $\alpha_2 \in I_2$, the ratio between $Vol(K_{\alpha_1})$ and $Vol(K_{\alpha_2})$ is large.
2. If N is not too large, both distributions P_1, P_2 are close in total variation distance to some distributions of samples in which the samples are independent and have a spherically symmetric law.
3. The radial profiles (hence the distribution of the Euclidean norm of a random sample) of typical random bodies K_1, K_2 are very close to each other.

In other words, we will define two constructions of random convex bodies for which: 1. The typical volumes of the bodies they produce will be far from equal. 2. They will be both indistinguishable from spherically symmetric constructions for a polynomial number of samples. 3. The radial profiles they produce are indistinguishable from each other for a polynomial number of samples.

To go on with the proof, a simple application of Yao's lemma will help us assume that the algorithm is deterministic. A deterministic algorithm is actually a function $F : \mathbb{R}^{n^{\kappa+1}} \to \mathbb{R}$ which takes a sequence of points and returns the volume of the body. If the total variation distance between the probabilities P_1 and P_2 defined above is small, then there exists a set $A \subset \mathbb{R}^{n^{\kappa+1}}$ which has a high probability with respect to both P_1 and P_2. Obviously, for all $x \in A$, $F(x)$ is wrong in approximating the volume of at least one of the families.

In Sect. 2, we will describe how we build these families of bodies, $\{K_\alpha\}$, using a random construction which starts from a Euclidean ball, to which deletions which cut out parts of it, generated by some Poisson process, are applied. Then, using elementary properties of the Poisson process and some concentration of measure properties of the ball, we will see that the correlation between different points in polynomially long sequence of random points generated uniformly from the body will be very weak (with respect to the generation of the body itself). Using this fact, we will only have to inspect the distribution of a single random point. The construction will have a spherically symmetric nature, so the density of a single random point will only depend on its distance from the origin, and therefore we will only have to care about the distribution of the distance of a point from the origin in the generated bodies. The role of the following section, which is more technical but fairly delicate, will be to calibrate this construction so that these families have different volumes, yet, approximately the same distribution of distance from the origin.

Before we proceed to the proof, let us introduce some notation. In this note the number n will always denote a dimension. For an expression $f(n)$ which depends on n, by $f(n) = \mathbf{SE}(n)$ we mean: there exists some $n_0 \in \mathbb{N}$ and $\epsilon > 0$ such that for all $n > n_0$, $|f(n)| < e^{-n^{\epsilon}}$. Also write $f(n) = g(n)(1 + \mathbf{SE}(n))$ for $|\frac{f(n)}{g(n)} - 1| = \mathbf{SE}(n)$ and $f(n) = g(n) + \mathbf{SE}(n)$ for $|f(n) - g(n)| = \mathbf{SE}(n)$. The notation $f(n) \lesssim g(n)$ and $f(n) \gtrsim g(n)$ will be interpreted as $f(n) < g(n)$ and $f(n) > g(n)$ for n large enough.

Moreover, we decide that $N = N(n)$ denotes the length of the sequence of random points. All throughout this note we assume that there exists a universal constant $\varepsilon > 0$ such that $N(n) < e^{n^{\varepsilon}}$.

2 The Deletion Process

In this section we will describe the construction of random bodies which will later be used as counterexamples. Our goal, after describing the actual construction, will be to prove, using some simple properties of the Poisson distribution, a weak-correlation property between different points generated from the body.

Denote by D_n the n-dimensional Euclidean ball of unit radius, centered at the origin, and by ω_n its Lebesgue measure.

Recall that for two probability measures P_1, P_2 on a set Ω, the total-variation distance between the two measures is defined by

$$d_{\mathrm{TV}}(P_1, P_2) = \sup_{A \subseteq \Omega} \bigl| P_1(A) - P_2(A) \bigr|.$$

One can easily check that if these measures are absolutely continuous with respect to some third measure Q, then it is also equal half the $L_1(Q)$ distance between the two densities.

Define $r_0 = n^{-\frac{1}{3}}$ and

$$T_0(\theta) = D_n \cap \bigl\{ x; \langle x, \theta \rangle \le r_0 \bigr\}.$$

Let T be a function from the unit sphere to the set of convex bodies such that for every $\theta \in S^{n-1}$, $T(\theta)$ satisfies $T_0(\theta) \subseteq T(\theta) \subseteq D_n$. (Recall that most of the mass of the Euclidean ball is contained in $\{x_1 \in [-1, Cn^{-\frac{1}{2}}]\}$. So $T(\theta)$ contains almost all the mass of the Euclidean ball.) Moreover, let $m > 0$. We will now describe our construction of a random convex body $K_{T,m}$. First, suppose that $m \in \mathbb{N}$. Let $\Theta = (\theta_1, \theta_2, \ldots, \theta_m)$ be m independent random directions distributed according to the uniform measure on S^{n-1}. We define $K_{T,m}$ as

$$K_{T,m} = D_n \bigcap_i T(\theta_i).$$

Finally, instead of taking a fixed $m \in \mathbb{N}$, we take ζ to be a Poisson random variable with expectation m, independent of the above. We can now define $K_{T,\zeta}$ in the same manner.

Let us denote the probability measure on the set of convex bodies induced by the process described above by μ. After generating the body $K_{T,m}$, which, from now on will be denoted just by K wherever there is no confusion caused, we consider the following probability space: let $\Omega = (D_n)^N$ be the set of sequences of length N of points from D_n. Denote by λ the uniform probability measure on Ω, and for a convex body K, denote by λ_K the uniform probability measure on $K^N = \prod_{1 \le i \le N} K \subseteq \Omega$. Finally, define the probability measure $P = P_{T,m}$ on Ω as follows: for $A \subseteq \Omega$,

$$P(A) = \int \lambda_K(A)\, d\mu(K) = \int \frac{Vol(K^N \cap A)}{Vol(K^N)}\, d\mu(K).$$

(The measure P describes the following process: first, generate the random set K according to construction described above, and then generate N i.i.d. random points, independent of the above, according to the uniform measure on K.) Moreover, for $p = (x_1, \ldots, x_N) \in \Omega$, define $\pi_i(p) = x_i$, the projections onto the ith copy of the Euclidean ball.

It easy to check that P is absolutely continuous with respect to λ. We define the following function on Ω:

$$f_{T,m}(p) = \mathbb{P}\big(p \in K_{T,m}^N\big) = \mathbb{P}\big(\forall 1 \le i \le N, \pi_i(p) \in K_{T,m}\big). \tag{1}$$

As we will see later, the function f is related in a simple way to $\frac{dP}{d\lambda}$. Namely, we will have

$$\frac{dP}{d\lambda}(p) = \big(1 + \mathbf{SE}(n)\big)\frac{f(p)}{\int_\Omega f}$$

for all p in some subset of Ω with measure close to 1. For convenience, from now on $f_{T,m}$ will be denoted by f.

We start with some simple geometric observations regarding Ω. Denote by σ the rotation-invariant probability measure on S^{n-1}. Define, for $p \in \Omega$ and $1 \le i \le N$,

$$A_i(p) = \big\{\theta \in S^{n-1} : \pi_i(p) \notin T(\theta)\big\}. \tag{2}$$

For $1 \le i, j \le N$, let $F_{i,j} \subset \Omega_N$ be the event defined by

$$F_{i,j} = \left\{ p : \frac{\sigma(A_i(p) \cap A_j(p))}{\sigma(A_i(p))} < e^{-n^{0.1}} \right\} \tag{3}$$

and let

$$F = \bigcap_{1\le i\neq j\le N} F_{i,j} \tag{4}$$

(which should be understood as "no two points are too close to each other" and, as we will see, will imply that points are weakly correlated). We start with the following simple lemma.

Lemma 3 *Under the above notation*:

(i) $\lambda(F) = 1 + \mathbf{SE}(n)$.
(ii) *There exists some* $\varepsilon_0 > 0$ *such that if we assume that*

$$\mathbb{P}_\mu\bigl(Vol(K) < \omega_n e^{-n^{\varepsilon_0}}\bigr) < e^{-n^{\varepsilon_0}} \tag{5}$$

(*that is*, *the volume of* K *is typically not much smaller than the volume of* D_n), *then* $P(F) = 1 + \mathbf{SE}(n)$.

Proof (i) Let p be uniformly distributed in Ω. Denote $x_i = \pi_i(p)$, so that x_1, x_2 are independent points uniformly distributed in D_n. Let us calculate $\lambda(F_{1,2})$.

First, for a fixed $\theta \in S^{n-1}$, one has

$$\mathbb{P}\bigl(x_1 \notin T(\theta)\bigr) \le \mathbb{P}\bigl(x_1 \notin T_0(\theta)\bigr) = \mathbb{P}\bigl(\bigl\{\langle x_1, \theta\rangle \ge r_0\bigr\}\bigr).$$

Recalling that $r_0 = n^{-\frac{1}{3}} \gg n^{-\frac{1}{2}}$, by elementary calculations regarding marginals of the Euclidean ball, one gets

$$\mathbb{P}\bigl(x_1 \notin T(\theta)\bigr) \lesssim e^{-n^{0.2}}.$$

Now, fix $x_2' \in D_n$. Define $A_i := A_i(p)$. One has

$$\mathbb{E}\bigl(\sigma(A_1 \cap A_2)\big| x_2 = x_2'\bigr) = \int_{A_2} \mathbb{P}(\theta \in A_1)\, d\sigma(\theta) = \int_{A_2} \mathbb{P}\bigl(x_1 \notin T(\theta)\bigr)\, d\sigma(\theta)$$
$$\lesssim \sigma(A_2) e^{-n^{0.2}},$$

and so,

$$\mathbb{E}\left(\frac{\sigma(A_1 \cap A_2)}{\sigma(A_2)}\middle| x_2 = x_2'\right) \lesssim e^{-n^{0.2}}. \tag{6}$$

Now, this is true for every choice of x_2', and so integrating over x_2' gives

$$\mathbb{E}\frac{\sigma(A_1 \cap A_2)}{\sigma(A_2)} \lesssim e^{-n^{0.2}}.$$

Now we use Markov's inequality to get

$$\lambda\bigl(F_{1,2}^C\bigr) = \lambda\left(\left\{\frac{\sigma(A_1 \cap A_2)}{\sigma(A_2)} > e^{-n^{0.1}}\right\}\right) = \mathbf{SE}(n). \tag{7}$$

A union bound completes the proof of (i).

(ii) First, we can condition on the event $\{Vol(K) > \omega_n e^{\varepsilon_0}\}$ (with ε_0 to be chosen later). Equation (5) ensures us that it will happen with probability $= 1 - \mathbf{SE}(n)$. Observe that for any event $E \subset \Omega$ which is measurable by the σ-field generated by π_1, π_2, we have

$$\lambda_K(E) = \frac{\omega_n^2 \lambda((K \times K \times D_n \times \cdots \times D_n) \cap E)}{Vol(K)^2} \le \frac{\omega_n^2 \lambda(E)}{Vol(K)^2}. \tag{8}$$

Now, taking $E = F_{1,2}^C$, choosing ε_0 to be small enough, and using (7) and (8) along with (5), one gets

$$P(F_{1,2}) = 1 + \mathbf{SE}(n).$$

Applying a union bound finishes the proof. □

We can now turn to the lemma which contains the main ideas of this section:

Lemma 4 *There exist $\varepsilon_0, \varepsilon_1 > 0$ and n_0 such that for every $n > n_0$, the following holds: Whenever m is small enough such that the following condition is satisfied:*

$$\mathbb{P}\big(\{\theta \in K\}\big) > e^{-n^{\varepsilon_0}} \quad \forall \theta \in S^{n-1} \tag{9}$$

(that is, we are not removing too much volume, in expectation, even from the outer shell), then:

(i) *We have*

$$P\big(\big|Vol(K) - \mathbb{E}\big(Vol(K)\big)\big| > e^{-n^{\varepsilon_1}}\mathbb{E}\big(Vol(K)\big)\big) = \mathbf{SE}(n), \tag{10}$$

and also (5) holds.

(ii) *For all $p \in F$, we have*

$$f(p) = \big(1 + \mathbf{SE}(n)\big) \prod_{j=1}^{N} \mathbb{P}\big(\pi_j(p) \in K\big).$$

In other words, if we define $\tilde{f} : D_n \to \mathbb{R}$ as

$$\tilde{f}(x) = \mathbb{P}(x \in K), \tag{11}$$

then

$$f(p) = \big(1 + \mathbf{SE}(n)\big) \prod_i \tilde{f}\big(\pi_i(p)\big) \quad \forall p \in F. \tag{12}$$

(iii)

$$\frac{\mathbb{E}(Vol(K^N \cap F))}{(\mathbb{E}Vol(K))^N} - 1 = \mathbf{SE}(n).$$

Proof We begin by proving (ii).

Fix $p \in F$. Define $x_i = \pi_i(p)$ and $A_i = A_i(p) \subset S^{n-1}$ as in (2). Also define $G_j = \bigcap_{i \le j} \{x_i \in K\}$. Fix $2 \le j \le N$. Let us try to estimate $P(G_j | G_{j-1})$.

When conditioning on the event G_{j-1}, we can consider our Poisson process as a superposition of three "disjoint" Poisson processes: the first one, with intensity λ_s, only generates deletions that cut x_j but leave all the x_i's for $i < j$ intact. The second one, with intensity λ_u, deletes x_j along with one of the other x_i's, and the third one is the complement (hence, deletions that do not affect x_j). We have, recalling that the expectation of the number of deletions is m,

$$\lambda_s\big(S^{n-1}\big) + \lambda_u\big(S^{n-1}\big) = m\sigma(A_j). \tag{13}$$

Moreover,

$$\lambda_u\big(S^{n-1}\big) \le m \sum_{i<j} \sigma(A_i \cap A_j) \tag{14}$$

(in the above formula we are including, multiple times, deletions that cut more than two points, and hence it is the inequality rather than equality).

Now, using the definition of F, one gets

$$\frac{\lambda_u(S^{n-1})}{\lambda_s(S^{n-1}) + \lambda_u(S^{n-1})} = \mathbf{SE}(n). \tag{15}$$

Note that (9) implies

$$e^{-(\lambda_s(S^{n-1})+\lambda_u(S^{n-1}))} \ge e^{-m\sigma(\{\theta; \frac{x_j}{|x_j|} \notin T(\theta)\})} \ge e^{-n^{\varepsilon_0}} \tag{16}$$

(the first inequality follows from the fact that $T(\theta)$ are star-shaped). The last two inequalities give

$$\lambda_u\big(S^{n-1}\big) = \mathbf{SE}(n). \tag{17}$$

It follows that

$$\left| \frac{P(G_j|G_{j-1})}{P(\{x_j \in K\})} - 1 \right| = \frac{e^{-\lambda_s(S^{n-1})}}{e^{-(\lambda_s(S^{n-1})+\lambda_u(S^{n-1}))}} - 1 = \mathbf{SE}(n). \tag{18}$$

Moreover, one has

$$P(G_N) = \prod_j P(G_j|G_{j-1}) = \prod_j \left(\frac{P(G_j|G_{j-1})}{P(\{x_j \in K\})} P\big(\{x_j \in K\}\big) \right). \tag{19}$$

Using (18) and (19), we get

$$f(p) = P(G_N) = \big(1 + \mathbf{SE}(n)\big) \prod_j P\big(\{x_j \in K\}\big). \tag{20}$$

This proves (ii).

(i) Showing that (5) holds is just a matter of noticing that $\mathbb{P}(x \in K)$ is decreasing with respect to $|x|$ and taking ε_0 small enough. We turn to estimate $\mathbb{E}(Vol(K)^2)$. We have

$$\mathbb{E}\big(Vol(K)^2\big) = \int_{D_n \times D_n} \mathbb{P}\big(\{x_1 \in K\} \cap \{x_2 \in K\}\big)\, dx_1\, dx_2 \tag{21}$$

$$= \int_{(D_n \times D_n) \cap F_{1,2}} \mathbb{P}\big(\{x_1 \in K\} \cap \{x_2 \in K\}\big)\, dx_1\, dx_2 \tag{22}$$

$$+ \int_{(D_n \times D_n) \cap F_{1,2}^C} \mathbb{P}\big(\{x_1 \in K\} \cap \{x_2 \in K\}\big)\, dx_1\, dx_2$$

(we will later see that the second summand is negligible). Now, (20) gives

$$\int_{(D_n \times D_n) \cap F_{1,2}} \mathbb{P}\big(\{x_1 \in K\} \cap \{x_2 \in K\}\big)\, dx_1\, dx_2$$

$$= \big(1 + \mathbf{SE}(n)\big) \int_{(D_n \times D_n) \cap F_{1,2}} \mathbb{P}\big(\{x_1 \in K\}\big)\mathbb{P}\big(\{x_2 \in K\}\big)\, dx_1\, dx_2, \tag{23}$$

which also implies that

$$\int_{(D_n \times D_n) \cap F_{1,2}} \mathbb{P}\big(\{x_1 \in K\} \cap \{x_2 \in K\}\big)\, dx_1\, dx_2 > \frac{1}{2} e^{-2n^{\varepsilon_0}}.$$

Recall that $\lambda(F_{1,2}^C) = \mathbf{SE}(n)$ (as a result of the previous lemma). Taking ε_0 to be small enough, we will get

$$\mathbb{E}\big(Vol(K)^2\big) = \big(1 + \mathbf{SE}(n)\big) \int_{(D_n \times D_n) \cap F_{1,2}} \mathbb{P}\big(\{x_1 \in K\} \cap \{x_2 \in K\}\big)\, dx_1\, dx_2$$

$$= \big(1 + \mathbf{SE}(n)\big) \int_{(D_n \times D_n) \cap F_{1,2}} \mathbb{P}\big(\{x_1 \in K\}\big)\mathbb{P}\big(\{x_2 \in K\}\big)\, dx_1\, dx_2.$$

On the other hand,

$$\mathbb{E}\big(Vol(K)\big)^2 = \int_{(D_n \times D_n)} \mathbb{P}\big(\{x_1 \in K\}\big)\mathbb{P}\big(\{x_2 \in K\}\big)\, dx_1\, dx_2. \tag{24}$$

Using the same considerations as above, the part of the integral over $F_{1,2}^C$ can be ignored; hence,

$$\mathbb{E}\big(Vol(K)\big)^2 = \big(1 + \mathbf{SE}(n)\big) \int_{(D_n \times D_n) \cap F_{1,2}} \mathbb{P}\big(\{x_1 \in K\}\big)\mathbb{P}\big(\{x_2 \in K\}\big)\, dx_1\, dx_2. \tag{25}$$

So, we finally get

$$\mathbb{E}\big(Vol(K)^2\big) = \big(1 + \mathbf{SE}(n)\big)\mathbb{E}\big(Vol(K)\big)^2. \tag{26}$$

Recalling that we assume (9) and using Chebyshev's inequality, this easily implies (i), which finishes (ii).

For the proof of (iii),

$$\begin{aligned}\mathbb{E}\big(Vol\big(K^N \cap F\big)\big) &= \int_F \mathbb{P}\big(p \in K^N\big)\,dp = \big(1+\mathbf{SE}(n)\big)\int_F \prod_i \mathbb{P}\big(\pi_i(p) \in K\big)\\ &\le \big(\mathbb{E}Vol(K)\big)^N. \qquad\square\end{aligned}$$

Consider the density $\frac{dP}{d\lambda}$. Our next goal is to find a connection between this density and the function f. Let $A \subseteq F \subset \Omega$. Using the concentration properties of $Vol(K)$, we will prove that

$$P(A) = \frac{\int_A f(p)\,dp}{(\int_{D_n} \tilde{f}(x))^N} + \mathbf{SE}(n), \tag{27}$$

where f, $\tilde{f}$ are defined in (1) and (11).

We have

$$P(A) = \mathbb{E}_\mu\left(\frac{Vol(K^N \cap A)}{Vol(K^N)}\right) = \mathbb{E}_\mu\left(\frac{Vol(K^N \cap A)}{Vol(K)^N}\right). \tag{28}$$

By Fubini,

$$\mathbb{E}_\mu Vol\big(K^N \cap A\big) = \int_A f(p)\,dp. \tag{29}$$

Consider the event

$$G := \left\{\left|\frac{Vol(K)^N}{\mathbb{E}(Vol(K))^N} - 1\right| < e^{-n^{\frac{\varepsilon_1}{2}}}\right\}$$

(where ε_1 is the constant from Lemma 4). We have, by the definition of G,

$$\int_G \frac{Vol(K^N \cap A)}{Vol(K)^N}\,d\mu(K) = \frac{\int_G Vol(K^N \cap A)\,d\mu(K)}{\mathbb{E}(Vol(K))^N} + \mathbf{SE}(n). \tag{30}$$

It follows from part (i) of Lemma 4 that

$$\begin{aligned}\mu(G) &= \mathbb{P}\left(\left|\left(\frac{Vol(K)}{\mathbb{E}(Vol(K))}\right)^N - 1\right| \le e^{-n^{\frac{\varepsilon_1}{2}}}\right)\\ &\ge \mathbb{P}\left(\left|\frac{Vol(K)}{\mathbb{E}(Vol(K))} - 1\right| \le 2Ne^{-n^{\frac{\varepsilon_1}{2}}}\right)\\ &\ge \mathbb{P}\left(\left|\frac{Vol(K)}{\mathbb{E}(Vol(K))} - 1\right| \le e^{-n^{\varepsilon_1}}\right)\\ &= 1 + \mathbf{SE}(n).\end{aligned}$$

So $\mu(G) = 1 + \mathbf{SE}(n)$, which gives

$$\int_{G^C} \frac{Vol(K^N \cap A)}{Vol(K)^N}\, d\mu(K) \le \mu\big(G^C\big) = \mathbf{SE}(n). \tag{31}$$

We will also need that

$$\frac{\int_{G^C} Vol(K^N \cap A)\, d\mu(K)}{(\mathbb{E}Vol(K))^N} = \mathbf{SE}(n). \tag{32}$$

To prove this, first recall that $A \subseteq F$. This gives

$$\begin{aligned} \frac{\int_{G^C} Vol(K^N \cap A)\, d\mu(K)}{(\mathbb{E}Vol(K))^N} &\le \frac{\int_{G^C} Vol(K^N \cap F)\, d\mu(K)}{(\mathbb{E}Vol(K))^N} \\ &= \frac{\mathbb{E}_\mu Vol(K^N \cap F)}{(\mathbb{E}Vol(K))^N} - \frac{\int_G Vol(K^N \cap F)\, d\mu(K)}{(\mathbb{E}Vol(K))^N}. \end{aligned} \tag{33}$$

Now,

$$\int_G \frac{Vol(K^N \cap F)}{Vol(K^N)}\, d\mu(K) = \mathbb{E}_\mu \frac{Vol(K^N \cap F)}{Vol(K^N)} + \mathbf{SE}(n) = P(F) + \mathbf{SE}(n) = 1 + \mathbf{SE}(n),$$

and so,

$$\frac{\int_G Vol(K^N \cap F)\, d\mu(K)}{(\mathbb{E}Vol(K))^N} = 1 + \mathbf{SE}(n). \tag{34}$$

Using part (iii) of Lemma 4, along with (33) and (34), proves (32).

Plugging together (28), (30), (31), and (32) implies

$$\begin{aligned} P(A) &= \mathbb{E}_\mu \frac{Vol(K^N \cap A)}{Vol(K)^N} = \int_G \frac{Vol(K^N \cap A)}{Vol(K)^N}\, d\mu(K) + \mathbf{SE}(n) \\ &= \frac{\int_G Vol(K^N \cap A)\, d\mu(K)}{\mathbb{E}(Vol(K))^N} + \mathbf{SE}(n) = \frac{\mathbb{E}_\mu Vol(K^N \cap A)}{\mathbb{E}(Vol(K))^N} + \mathbf{SE}(n). \end{aligned} \tag{35}$$

Recall that, as a result of Fubini's theorem,

$$\mathbb{E}_\mu\big(Vol(K)\big) = \int_{D_n} \tilde{f}(x)\, dx. \tag{36}$$

Plugging (35), (36), and (29) proves (27). We would now like to use the result of Lemma 4, to replace f with $\tilde{f}$. Let $A' \subseteq \Omega$. Define $A = A' \cap F$,

$$P(A') = P(A) + P\big(A' \cap F^C\big).$$

Part (ii) of Lemma 3 with (27) gives

$$P(A') = P(A) + \mathbf{SE}(n) = \frac{\int_A f(p)\, dp}{(\int_{D_n} \tilde{f}(x)\, dx)^N} + \mathbf{SE}(n).$$

We can now plug in (12) to get

$$P(A') = \frac{\int_A \prod_i \tilde{f}(\pi_i(p))\,dp}{(\int_{D_n} \tilde{f}(x))^N} + \mathbf{SE}(n).$$

So, finally defining

$$\frac{d\tilde{P}}{dp} = \frac{\mathbf{1}_{\{p\in F\}} \prod_i \tilde{f}(\pi_i(p))}{(\int_{D_n} \tilde{f}(x))^N} = \mathbf{1}_{\{p\in F\}} \prod_i \frac{\tilde{f}(\pi_i(p))}{\int_{D_n} \tilde{f}(x)\,dx},$$

we have proved the following lemma:

Lemma 5 *Suppose that condition* (9) *from Lemma* 4 *holds. Then one has*

$$d_{\mathrm{TV}}(P, \tilde{P}) = \mathbf{SE}(n).$$

Note that the measure $\tilde{P}$ is not, in general, a probability measure. The lemma, however, ensures that $\tilde{P}(\Omega)$ is very close to 1.

Recall that our plan is to find two families of convex bodies, achieved by two pairs (T_1, m_1) and (T_2, m_2), such that $d_{\mathrm{TV}}(P_1, P_2)$ is small, even though their volumes differ.

The above lemma motivates us to try to find such pairs with $\frac{\tilde{f}_1}{\int \tilde{f}_1} = \frac{\tilde{f}_2}{\int \tilde{f}_2} + \mathbf{SE}(n)$. We formulate this accurately in the following lemma.

Lemma 6 *Suppose that there exist two pairs* (T_i, m_i) *for* $i = 1, 2$ *such that* (9) *is satisfied and in addition, defining* $\tilde{f}_1$ *and* $\tilde{f}_2$ *as in* (11),

$$\left\| \frac{\tilde{f}_1}{\int_{D_n} \tilde{f}_1} - \frac{\tilde{f}_2}{\int_{D_n} \tilde{f}_2} \right\|_{L_1(D_n)} = \mathbf{SE}(n). \tag{37}$$

Then $d_{\mathrm{TV}}(P_1, P_2) = \mathbf{SE}(n)$.

Proof Using the previous lemma, it is enough to show that $d_{\mathrm{TV}}(\tilde{P}_1, \tilde{P}_2) = \mathbf{SE}(n)$. Define $g_i = \frac{\tilde{f}_i}{\int_{D_n} \tilde{f}_i}$. We have

$$\begin{aligned} d_{\mathrm{TV}}(\tilde{P}_1, \tilde{P}_2) &\le \int_\Omega \Bigl| \prod_{1\le i\le N} g_1\bigl(\pi_i(p)\bigr) - \prod_{1\le i\le N} g_2\bigl(\pi_i(p)\bigr) \Bigr| \\ &\le \sum_{1\le j\le N} \int_\Omega \Bigl| \prod_{1\le i\le j} g_1\bigl(\pi_i(p)\bigr) \prod_{j+1\le i\le N} g_2\bigl(\pi_i(p)\bigr) \\ &\qquad - \prod_{1\le i\le j+1} g_1\bigl(\pi_i(p)\bigr) \prod_{j+2\le i\le N} g_2\bigl(\pi_i(p)\bigr) \Bigr| \\ &= N \int_{D_n} \bigl| g_1(x) - g_2(x) \bigr| = \mathbf{SE}(n). \end{aligned}$$

□

In the next section we deal with how to calibrate T_i and m_i so that (37) holds.

3 Building the Two Profiles

Our goal in this section is to build convex bodies with a prescribed radial profile.

For a measurable body $L \subset \mathbb{R}^n$, define

$$g_L(r) = 1 - \sigma\left(\frac{1}{r} L \cap S^{n-1}\right). \tag{38}$$

This function should be understood as the "profile" of mass of the complement of L, which will eventually be the ratio of mass which a single deletion removes, in expectation, as a function of the distance from the origin. Define $g_i(r) = g_{T_i}(r)$.

Let us try understand exactly what kind of construction we require. Fix $x \in D_n$. Keeping in mind that the function $T_i(\theta)$ commutes with orthogonal transformations, we learn that the probability that x is removed in a single deletion of T_i is exactly $g_i(|x|)$. By elementary properties of the Poisson process, this gives

$$\mathbb{P}(x \in K_i) = \exp\bigl[-m_i g_i\bigl(|x|\bigr)\bigr]. \tag{39}$$

In view of (37), we would like the ratio $\frac{\mathbb{P}(x \in K_1)}{\mathbb{P}(x \in K_2)}$ to be (approximately) constant. Using (39), we see that the latter follows from

$$m_1 g_1\bigl(|x|\bigr) - m_2 g_2\bigl(|x|\bigr) = C.$$

If we choose to pick $m_2 = 2m_1$, this equality will be implied by the following requirements on T_1, T_2:

$$g_1(1) = g_2(1) \neq 0, \quad \text{and} \quad g_1'(r) = 2g_2'(r), \quad r \in [0, 1]. \tag{40}$$

Assuming that (5) holds and making use of the concentration of the radial profile of D_n, we will actually only be required to make sure that the derivatives are proportional for $r \in [1 - n^{-0.99}, 1]$.

Note that when (40) is attained, by picking different values of m_1, the ratio between the expected volumes of K_1 and K_2 can be made arbitrarily large while the expected radial profiles remain about as close. Lemma 6 will then ensure us that this is enough for the distributions to be indistinguishable.

The above is established in the main lemma of this section:

Lemma 7 *For every dimension n, there exist two convex bodies $T_1, T_2 \subset \mathbb{R}^n$ satisfying the following*:

(i)

$$D_n \supseteq T_i \supseteq D_n \cap \bigl\{x; \langle x, e_1\rangle \leq n^{-\frac{1}{3}}\bigr\}, \quad i = 1, 2. \tag{41}$$

(ii) *The radial profiles satisfy*,

$$g_1(1) = g_2(1) \neq 0, \quad \textit{and} \quad g_1'(r) = 2g_2'(r) \quad \forall r \in \bigl[1 - n^{-0.99}, 1\bigr]. \tag{42}$$

To achieve this, we begin by describing the following construction: Define $\delta_0 = n^{-\frac{1}{4}}$ and $\delta_1 = n^{-0.99}$. For every two constants a, b such that $a \in [2, 200]$ and $b \in [-1000, 1000]$, let $f = f_{a,b}$ be the linear function with negative slope which satisfies

$$f\big(\delta_0(1+\delta_1 b)\big) = \sqrt{1-(\delta_0(1+\delta_1 b))^2} \tag{43}$$

and

$$\min_{x\in\mathbb{R}} \sqrt{x^2+f^2(x)} = a\delta_0 \tag{44}$$

(hence, it is a line of distance $a\delta_0$ from the origin which meets the unit circle at $x = \delta_0(1+b\delta_1)$. Note that there exists such a linear function with negative slope since $a\delta_0 \gg \delta_0(1+b\delta_1)$). We define the convex body $T_{a,b}$ by

$$T_{a,b} = D_n \cap \big\{(x, \vec{y}) \in \mathbb{R}\times\mathbb{R}^{n-1} = \mathbb{R}^n; |y| \le f(x)\big\} \tag{45}$$

(an intersection of the ball with a cone defined by a linear equation the coefficients of which depend on a, b).

Recall that we require that $a > 2$ and $b > -1000$. First of all, from requirement (43) and from the fact that the slope of f is negative it follows directly that $T_{a,b}$ satisfies (41) (since $\delta_0 \gg n^{-1/3}$).

Define $g_{a,b}(r) = g_{T_{a,b}}(r)$ as in (38). Let us find an expression for $g_{a,b}(r)$. First, a simple calculation shows that (44) implies that the function $f_{a,b}$ intersects the x axis at $x < 2a\delta_0$. This shows that $T_{a,b} \cap rS^{n-1}$ has only one connected component for all $r > \frac{1}{2}$ (hence, the vertex of the cone is inside the sphere).

Consider the intersection $\frac{1}{r}T_{a,b} \cap S^{n-1}$. If $r > \frac{1}{2}$, it must be a set of the form $S^{n-1} \cap \{x_1 < x(a,b,r)\}$ for some function $x(a,b,r)$. Let us try to find the expression for this function. Equation (44) shows that $T_{a,b}$ is an intersection of D_n with halfspaces at distance $a\delta_0$ from the origin. This implies that $x(a,b,r)$ must satisfy

$$x(a,b,r) = \sin\left(\arcsin\left(\frac{a\delta_0}{r}\right) + c\right)$$

for some constant c (draw a picture). To find the value of c, we use (43) to get $x(a,b,1) = \delta_0(1+b\delta_1)$, and so

$$x(a,b,r) = \sin\left(\arcsin\left(\frac{a\delta_0}{r}\right) - \arcsin(a\delta_0) + \arcsin\big(\delta_0(1+b\delta_1)\big)\right). \tag{46}$$

Next, define

$$\Psi(x) = \frac{1}{\omega_n}\int_{\min(x,1)}^{1} \big(1-t^2\big)^{\frac{n-3}{2}}\, dt,$$

the surface area measure of a cap the base of which has distance x from the origin. We finally have

$$g_{a,b}(r) = \sigma\big(S^{n-1} \cap \big\{x_1 \ge x(a,b,r)\big\}\big) = \Psi\big(x(a,b,r)\big). \tag{47}$$

Given a subset $I' \subseteq \mathbb{R} \times \mathbb{R}$, we define

$$K_{I'} = \bigcap_{(a,b)\in I'} T_{a,b}. \tag{48}$$

Clearly,

$$g_{I'}(r) := g_{K_{I'}}(r) = \sup_{(a,b)\in I'} g_{a,b}(r).$$

Our goal is to choose such a subset so that (42) is fulfilled. We will use the following elementary result:

Lemma 8 *Let $c > 0$, and let $\{f_\alpha\}_{\alpha\in I}$ be a family of twice-differentiable functions defined on $[x_1, x_2]$ such that for every triplet $(x, y, y') \in [x_1, x_2] \times [y_1, y_2] \times [y'_1, y'_2]$, there exists $\alpha \in I$ such that*

$$f_\alpha(x) = y, \qquad f_\alpha(x)' = y', \qquad f''(t) \le c \quad \forall t \in [x_1, x_2]. \tag{49}$$

Then for every twice differentiable function $g : [x_1, x_2] \to [y_1, y_2]$ with

$$g'(x) \in [y'_1, y'_2], \qquad g''(x) > c, \tag{50}$$

there exists a subset $I' \subset I$ such that

$$g(x) = \sup_{\alpha\in I'} f_\alpha(x). \tag{51}$$

In view of the above lemma, we would like to show that by choosing appropriate values of a, b, one can attain functions $g_{a,b}$ which, for a fixed r_0, have prescribed values $g_{a,b}(r_0)$, $g'_{a,b}(r_0)$, and a small enough second derivative.

Define $r(u) = 1 - \delta_1 u$. Note that substituting $r \to u$, almost all of the mass of the Euclidean ball is contained in $u \in [0, 1]$ (the thin shell of the Euclidean ball). We now turn to prove the following lemma:

Lemma 9 *Suppose that (u, g_0, g'_0) satisfy $0 \le u \le 1$,*

$$\Psi(\delta_0) - 100\delta_0\delta_1\Psi'(\delta_0) \le g_0 \le \Psi(\delta_0) + 100\delta_0\delta_1\Psi'(\delta_0),$$
$$10\delta_0\delta_1\Psi'(\delta_0) \le g'_0 \le 100\delta_0\delta_1\Psi'(\delta_0).$$

There exist constants $a \in [2, 200], b \in [-1000, 1000]$ such that $g_{a,b}(r(u)) = g_0$, $(g_{a,b}(r(u)))' = g'_0$, and $g_{a,b}(r(t))'' \le \delta_0\delta_1\Psi'(\delta_0)$ $\forall 0 \le t \le 1$.

Proof Throughout this proof we always assume $u \in [0, 1], a \in [2, 200]$, and $b \in [-1000, 1000]$.

Let us inspect the function $x(a, b, r))$ defined in (46). Differentiating it twice, while recalling that $a\delta_0 \ll \frac{1}{2}$, gives us the following fact: there exists $C > 0$, independent of n, such that $|\frac{\partial^2}{\partial r^2} x(a, b, r)| < C$. Consider $x(a, b, u) := x(a, b, r(u))$. One has

$$x_{uu}(a, b, u) = O\big(\delta_1^2\big) \tag{52}$$

(here and afterwards, by "O" we mean that the term is smaller than some universal constant times the expression inside the brackets, which is valid as long as u, a, b attain values in the intervals defined above). This implies that for all $u \in [0, 1]$,

$$\begin{aligned} x_u(a,b,u) &= x_u(a,b,0) + O\big(\delta_1^2\big) \\ &= a\delta_0\delta_1 \sin'\big(\arcsin\big(\delta_0(1+b\delta_1)\big)\big)\big(1+O(\delta_0)\big) + O\big(\delta_1^2\big) \\ &= a\delta_0\delta_1\big(1+O(\delta_0)\big), \end{aligned}$$

and so,

$$x(a,b,u) = x(a,b,0) + a\delta_0\delta_1 u\big(1+O(\delta_0)\big) = \delta_0 + \delta_0\delta_1(au+b)\big(1+O(\delta_0)\big).$$

Let us now define $w(a,b,u) = \frac{1}{\delta_1}(\frac{x(a,b,r(u))}{\delta_0} - 1)$. So,

$$w(a,b,u) = (au+b)\big(1+O(\delta_0)\big) \tag{53}$$

and

$$w_u(a,b,u) = a\big(1+O(\delta_0)\big), \qquad w_{uu}(a,b,u) = O\bigg(\frac{\delta_1}{\delta_0}\bigg). \tag{54}$$

Next, we consider $g_{a,b}(r(u)) = \Psi(x(a,b,u)) = \Psi(\delta_0(1+\delta_1 w(a,b,u))$. We have

$$\Psi\big(x(a,b,u)\big) = \Psi(\delta_0) + \delta_0\delta_1\Psi'(\delta_0)w(a,b,u) + \frac{\delta_0^2\delta_1^2}{2}\Psi''(t)w(a,b,u)^2 \tag{55}$$

for some $t \in [\delta_0, x(a,b,u)]$. But note that the following holds:

$$\big(\log\Psi'(v)\big)' = \frac{\Psi''(v)}{\Psi'(v)} = -\frac{2v\frac{n-3}{2}(1-v^2)^{\frac{n-5}{2}}}{(1-v^2)^{\frac{n-3}{2}}} = -\frac{v(n-3)}{(1-v^2)}, \tag{56}$$

and, for all $v \in [\frac{\delta_0}{2}, 2\delta_0]$,

$$\big(\log\Psi'(v)\big)' = O(n\delta_0).$$

Integration of this inequality yields that for t such that $t - \delta_0 = O(\delta_0\delta_1)$, one has

$$\log\Psi'(t) - \log\Psi'(\delta_0) = O\big(n\delta_0^2\delta_1\big)$$

or

$$\Psi'(t) = \Psi'(\delta_0)\big(1+O\big(n\delta_0^2\delta_1\big)\big). \tag{57}$$

Combining (56) and (57) gives

$$\delta_0^2\delta_1^2\Psi''(t) = O\big(\Psi'(\delta_0)\delta_1^2 n\delta_0^3\big) = o\big(\Psi'(\delta_0)\delta_0\delta_1\big). \tag{58}$$

This finally gives

$$\begin{aligned} g_{a,b}\big(r(u)\big) = \Psi\big(x(a,b,u)\big) &= \Psi(\delta_0) + \big(\delta_0\delta_1\Psi'(\delta_0)w(a,b,u)\big)\big(1+o(1)\big) \\ &= \Psi(\delta_0) + \delta_0\delta_1\Psi'(\delta_0)(au+b)\big(1+o(1)\big). \end{aligned} \tag{59}$$

Next we try to estimate the derivative of $\Psi(x(a,b,u))$. We have

$$\frac{\partial}{\partial u} g_{a,b}\big(r(u)\big) = \frac{\partial}{\partial u}\Psi\big(x(a,b,u)\big)$$
$$= \Psi'\big(x(a,b,u)\big)x_u(a,b,u) = \Psi'\big(x(a,b,u)\big)\delta_0\delta_1 w_u(a,b,u) \qquad (60)$$

and, using (57),

$$\frac{\partial}{\partial u}\Psi\big(x(a,b,u)\big) = \Psi'(\delta_0)\big(1+o(1)\big)\delta_0\delta_1 w_u(a,b,u)$$
$$= \big(a\delta_0\delta_1\Psi'(\delta_0)\big)\big(1+o(1)\big). \qquad (61)$$

Using the continuity and Ψ and $x(a,b,u)$, we can now conclude the following: for any fixed $b \in [-1000, 1000]$ and $u \in [0,1]$, an inspection of (61) teaches us that when a varies in $[2, 200]$, $\frac{\partial}{\partial u}\Psi(x(a,b,u))$ can attain all values in the range $[3\delta_0\delta_1\Psi'(\delta_0), 100\delta_0\delta_1\Psi'(\delta_0)]$. An inspection of (59) shows that afterwards, by letting b vary in $[-1000, 1000]$, $g_{a,b}(r(u))$ will attain all values in $[\Psi(\delta_0) - 100\delta_0\delta_1\Psi'(\delta_0), \Psi(\delta_0) + 100\delta_0\delta_1\Psi'(\delta_0)]$. To estimate the second derivative, $g''_{a,b}$, we write

$$\frac{\partial^2}{\partial u^2}\Psi\big(x(a,b,u)\big) = \Psi''\big(x(a,b,u)\big)\delta_0^2\delta_1^2 w_u^2(a,b,u)$$
$$+ \delta_0\delta_1\Psi'\big(x(a,b,u)\big)w_{uu}(a,b,u)w_u(a,b,u)$$

(using (54) and (58))

$$= o\big(\delta_0\delta_1\Psi'(\delta_0)\big) + O\big(\delta_1^2\Psi'(\delta_0)\big) = o\big(\delta_0\delta_1\Psi'(\delta_0)\big).$$

This completes the proof of the lemma. □

We are now ready to prove the main lemma of the section.

Proof of Lemma 7 Define

$$f_i(r) = \Psi(\delta_0) + C_i\delta_0\delta_1\Psi'(\delta_0)(u+1)^2$$

with $C_1 = 20$, $C_2 = 40$. Lemmas (9) and (8) show that there exist two subsets I_1, I_2 of $[2, 200] \times [-1000, 1000]$ such that the bodies $T_i = T_{I_i}$ that we constructed in (48) satisfy (42). Also, (41) is satisfied, since it is satisfied for $T_{a,b}$ for all $(a,b) \in [2, 200] \times [-1000, 1000]$, as we have seen. □

4 Tying Up Loose Ends

Proof of Theorem 2 Use Lemma 7 two build the two bodies T_i. Let U_θ be an orthogonal transformation which sends e_1 to θ. Define $T_i(\theta) = U_\theta(T_i)$ (note that the choice of orthogonal transformation does not matter because T_i are bodies of revolution around e_1). Define the functions $g_i = g_{T_i}$ as in (38). Let $m_1 = \frac{n^\varepsilon}{g_1(1)}$ with $\varepsilon > 0$ to be chosen later. Define $m_2 = 2m_1$. So, (42) implies that

$$m_2 g_2(r) = m_1 g_1(r) + m_1 g_1(1) \quad \forall r \in \left[1 - n^{-0.99}, 1\right]. \tag{62}$$

Now, let $K_i = K_{T_i, m_i}$ be the random bodies we constructed in Sect. 2.

For a fixed $x \in D_n$, as in (39), we have

$$\tilde{f}_i(x) = P(x \in K_i) = e^{-m_i \sigma(\{x \notin T_i(\theta)\})} = e^{-m_i g_i(|x|)}. \tag{63}$$

Now, (62) and (63) give

$$\frac{\tilde{f}_1(x)}{\tilde{f}_2(x)} = e^{m_1(g(1))} = e^{n^\varepsilon} \tag{64}$$

for all x with $|x| \in [1 - n^{-0.99}, 1]$.

Let us choose ε small enough so that

$$m_2 g_2(1) < n^{\varepsilon_0},$$

where ε_0 is the constant from (9). Clearly, this ensures that (9) holds for both random bodies K_i. Now, ε can be made further smaller, so that concentration properties of the Euclidean ball will give us

$$\int_{D_n} \tilde{f}_i = \big(1 + \mathbf{SE}(n)\big) \int_{D_n \setminus (1 - n^{-0.99}) D_n} \tilde{f}_i \tag{65}$$

for $i = 1, 2$. Clearly, the above can still be satisfied for some universal constant $\varepsilon > 0$ as long as n is large enough. Next, (64) and (65) imply that

$$\frac{\int_{D_n} \tilde{f}_1}{\int_{D_n} \tilde{f}_2} = \big(1 + \mathbf{SE}(n)\big) e^{n^\varepsilon},$$

and so (again, taking ε small enough) one gets

$$\int_{D_n} \left| \frac{\tilde{f}_1}{\int_{D_n} \tilde{f}_1} - \frac{\tilde{f}_2}{\int_{D_n} \tilde{f}_2} \right| dx = \int_{(1 - n^{-0.99}) D_n} \left| \frac{\tilde{f}_1}{\int_{D_n} \tilde{f}_1} - \frac{\tilde{f}_2}{\int_{D_n} \tilde{f}_2} \right| dx + \mathbf{SE}(n) = \mathbf{SE}(n).$$

Now use Lemma 6 to get that

$$d_{\mathrm{TV}}(P_1, P_2) = \mathbf{SE}(n). \tag{66}$$

Denote $R = \frac{1}{2} e^{n^\varepsilon}$. Then

$$\mathbb{E}\big(Vol(K_1)\big) = \big(1 + \mathbf{SE}(n)\big) 2R\mathbb{E}\big(Vol(K_2)\big).$$

Suppose by negation that there exists a classification function $F : \omega \to \mathbb{R}$ that determines the volume of a body K up to a constant $e^{n^{\varepsilon_2}}$ with probability 0.52. Denote

$L = [\frac{\mathbb{E}(Vol(K_1))}{R}, R\mathbb{E}(Vol(K_1))]$. Note that using (10), the "correctness" of the function implies that

$$P_1\big(F(p) \in L\big) \geq 0.51.$$

Denote $A \subset \Omega$ as $A = \{p \in \Omega : F(p) \in L\}$. Then $P_1(A) > 0.51$, and (66) implies that also $P_2(A) > 0.51$. This means that

$$P_2\big(F(p) \in L\big) > 0.5.$$

But clearly, again, (10) implies that with probability $= 1 + \mathbf{SE}(n)$, the volume of K_2 is not in L. This contradicts the existence of such a function F.

We still have to generalize the above in two aspects: for an even smaller probability of estimating the volume and the possibility that the algorithm is nondeterministic. Upon inspection of the proof above, we notice that it can be easily extended in the following way: instead of taking just two families of random bodies, K_1 and K_2, one may take $d > 2$ different families that are all indistinguishable by the algorithm and have different volumes. The proof can be stretched as far as $d = e^{n^{\frac{\varepsilon}{2}}}$. To deal with nondeterministic algorithms, we will use Yao's lemma (see [10], Lemma 11). Let us generate an index i uniformly distributed in $\{1, \dots, d\}$, then a body K from the family K_i, and then a sequence of uniformly distributed random points on K. Following the lines of the above proof, we see that every deterministic algorithm, given this sequence, will be incorrect in estimating the volume of K with probability (at least) $= 1 - \frac{1}{d} + \mathbf{SE}(n)$. It follows from Yao's lemma that every nondeterministic algorithm will be incorrect with the same probability for at least one of the families K_i. This finishes the proof of the theorem. □

Acknowledgements I am deeply grateful to my supervisor Prof. Boäz Klartag for very useful discussions and encouragement all along my work on the subject. I would also like to express my thanks to my supervisor, Prof. Vitali Milman for introducing me to this question and encouraging me to work on it, to the referee for providing useful comments and insights, and to Prof. Oded Regev for reviewing a preliminary version of the paper and providing numerous insightful comments and remarks.

References

1. Bárány, I., Füredi, Z.: Computing the volume is difficult. Discrete Comput. Geom. **2**, 319–326 (1987)
2. Bollobás, B.: Volume Estimates and Rapid Mixing. Flavors of Geometry, vol. 31. MSRI Publications, Berkeley (1997)
3. Dyer, M.E., Frieze, A.M., Kannan, R.: A random polynomial time algorithm for approximating the volume of convex bodies. J. ACM **38**(1), 1–17 (1991)
4. Goyal, N., Rademacher, L.: Learning convex bodies is hard (submitted manuscript, 2008). doi:10.1.1.162.3154
5. Klartag, B.: Power-law estimates for the central limit theorem for convex sets. J. Funct. Anal. **245**, 284–310 (2007)
6. Klivans, A.R., O'Donnell, R., Servedio, R.A.: Learning geometric concepts via Gaussian surface area. In: Proc. 49th IEEE Symposium on Foundations of Computer Science (2008)
7. Lovász, L.: Private communication
8. Lovász, L., Simonovits, M.: Random walks in a convex body and an improved volume algorithm. Random Struct. Algorithms **4**, 359–412 (1993)
9. Lovász, L., Vempala, S.: Simulated annealing in convex bodies and an $0^*(n^4)$ volume algorithm. FOCS 650–672 (2003)
10. Rademacher, L., Vempala, S.: Dispersion of mass and the complexity of randomized geometric algorithms. In: Proc. 47th IEEE Symposium on Foundations of Computer Science (2006)