# RANDOM WALKS AND AN $O^*(n^5)$ VOLUME ALGORITHM FOR CONVEX BODIES

Ravi Kannan[1]

László Lovász[2]

Miklós Simonovits[3]

January 1997 [4]

**Abstract**

Given a high dimensional convex body $K \subseteq \mathbb{R}^n$ by a separation oracle, we can approximate its volume with relative error $\varepsilon$, using $O^*(n^5)$ oracle calls. Our algorithm also brings the body into isotropic position.

As all previous randomized volume algorithms, we use "rounding" followed by a multiphase Monte-Carlo (product estimator) technique. Both parts rely on sampling (generating random points in $K$), which is done by random walk. Our algorithm introduces three new ideas:

- the use of the isotropic position (or at least an approximation of it) for rounding,
- the separation of global obstructions (diameter) and local obstructions (boundary problems) for fast mixing, and
- a stepwise interlacing of rounding and sampling.

## 1. Introduction

For a variety of geometric objects, classical results characterize various geometric parameters. Many of these results are useful even in practical situations: they can easily be transformed into efficient algorithms. Some other theorems do not yield fast algorithms. One of the most challenging examples is the problem of calculating the volume of a high dimensional body $K$. Even in the simplest cases, when $K$ is convex, no classical formula seems to translate into an efficient algorithm.

[1]Dept. of Computer Science, Carnegie-Mellon University, Pittsburgh, PA 15213; kannan@theory.cs.cmu.edu

[2]Dept. of Computer Science, Yale University, New Haven CT 06510; lovasz@cs.yale.edu

[3]Mathematical Institute, Hungarian Academy of Sciences, Budapest, Reáltanoda u. 13-15, H-1053 Hungary; miki@math-inst.hu

[4]This research was supported by the Discrete Mathematics program of Yale University and NSF grants CCR 9208597 and 9402916.

The reader may find powerful *negative* results in this direction in Elekes (1986), Bárány and Füredi (1986), Dyer and Frieze (1988), Khachian (1988), (1989), (1993), Lawrence (1989) and Lovász and Simonovits (1992). These results show that the diameter or the volume of convex bodies cannot be computed, and in certain models not even approximated, by any deterministic polynomial algorithm. (These problems can easily be solved in polynomial time in any *fixed dimension* $n$. However the degree of those polynomials estimating the running time of these algorithms increases fast with $n$.)

The situation changes dramatically if we allow *randomization*. Dyer, Frieze and Kannan (1989) gave the first polynomial randomized algorithm to calculate the volume of convex bodies. Their algorithm contained the main ingredients in all subsequent improvements: a multiphase Monte-Carlo algorithm (using the so-called product estimator) to reduce volume computation to sampling; the use of Markov chain techniques for sampling, and the use of the conductance bound on the mixing time, due to Sinclair and Jerrum (1988). Unfortunately, the degree of the time bound ($O^*(n^{23})$) prohibited practical applications. (We often use the "soft-O" notation ($O^*$), indicating that we suppress factors of $\ln n$ as well as factors depending on other parameters like the error bound $\varepsilon$.)

In all *randomized volume algorithms* we are given a convex body (a compact and full-dimensional convex set) $K \subseteq \mathbb{R}^n$; we assume $n \geq 3$ for convenience. We are also given two small positive numbers $\varepsilon$ (the required precision of our volume estimates) and $\eta$ (an upper bound on the probability of error).

Dyer, Frieze and Kannan (1989) (and its subsequent improvements) establish randomized algorithms returning a nonnegative number $\zeta$ such that

$$(1-\varepsilon)\zeta < \mathrm{vol}(K) < (1+\varepsilon)\zeta$$

with probability at least $1-\eta$. The running time of the algorithm is polynomial in $n$, $1/\varepsilon$ and $\ln(1/\eta)$.

This algorithm was subsequently improved by Lovász and Simonovits (1990) (to $O^*(n^{16})$), Applegate and Kannan (1990) (to $O^*(n^{10})$), Lovász (1992) (to $O^*(n^{10})$), Dyer and Frieze (1992) (to $O^*(n^8)$), and Lovász and Simonovits (1992), (1993) (to $O^*(n^7)$). The main result of this paper improves the number of oracle calls used by the algorithm to $O^*(n^5)$. This is, we hope, on the borderline of allowing practical implementations.

### 1.1 Computational model

The convex body is given by a *separation oracle*: a subroutine that, for a given point, tells us whether the point is in $K$, and if not, it gives a hyperplane separating the point from $K$ (see Grötschel, Lovász and Schrijver (1988) or Lovász (1986) for a discussion of this and other oracles for convex bodies).

We also assume that $K$ contains the unit ball $B$ and is contained in the ball of radius $n^{\text{const}}$ about the origin. Every convex body can be transformed into such a position (in fact with radius $n^{3/2}$) by a standard application of the ellipsoid method (or, quite often, by inspection), in $O^*(n^4)$ steps.

In order not to obscure our arguments by numerical considerations, we assume throughout that we can do exact real arithmetic, and that the separation oracle gives an exact answer. Neither of these assumptions is essential; we could do all our computations with a precision (say) $\varepsilon^* = \varepsilon/n^{10}$, and it would suffice to use a *weak separation oracle*, allowing an invalid answer if the queried point is closer to the surface of $K$ than $\varepsilon^*$, or a hyperplane that does not quite separate: the queried point may be on its wrong side, but by at most $\varepsilon^*$. The proof that this does not cause any harm is not hard and can be carried out on a quite general level; since it is virtually identical to the proof given by Lovász and Simonovits (1993), we do not reproduce it here.

The complexity of the algorithm can be measured by

- the number of queries to the oracle,
- the number of random bits used, and
- the number of arithmetic operations.

In the above sketch of the history of the problem we considered the number of *oracle calls* (and we shall do so in the sequel); the number of random bits used is larger by a factor of $O^*(n)$ and the number of arithmetic operations is larger by a factor of $O^*(n^2)$.

In our algorithms, the number of oracle calls is a random variable itself, and for various algorithms we will give slightly different kinds of bounds on it. Sometimes, we are able to give absolute bounds; sometimes, we can bound the expected number of calls; and sometimes, we can only bound the expected number of calls conditional on the non-occurance of a certain rare but very bad event. For the final algorithm, we can then conclude (applying Markov's inequality to the "good" cases) that with probability at least $1-\eta$, the number of oracle calls is at most a given bound. One may keep the number of oracle calls below this bound in all cases by interrupting the algorithm if the number of oracle calls gets too large, and include this in the probability of failure. But it will be necessary to use the more complicated information on the distribution of the number of oracle calls for certain subroutines.

We note that the role of the error probability $\eta$ is unimportant, at least as long as we are computing a single numerical parameter (like the volume): as Jerrum, Valiant and Vazirani (1986) pointed out, if one can solve such a problem for some $\eta_0 < 1/2$ in time $T$, then – iterating the algorithm and using the median – one achieves a reliability $\eta$ in time $c(\eta_0)\ln(1/\eta)\cdot T$. (For a more detailed explanation see e.g. Lovász and Simonovits (1993).) However, for certain other algorithmic problems, we'll have to deal with the error probability explicitly.

## 2. The main results

In this paper we shall improve upon previous randomized algorithms estimating the volume of an $n$-dimensional convex body $K$, given by a separation oracle. We also obtain improved algorithms for sampling and rounding. Let us assume throughout that $K$ satisfies

$$B \subseteq K \subseteq dB$$

for some $d < n^{\mathrm{const}}$. The main results of this paper can be summarized as follows.

**Theorem 2.1** *There is a (randomized) algorithm that, given $\varepsilon, \eta > 0$, returns a real number $\zeta$ for which*

$$(1-\varepsilon)\zeta < \mathrm{vol}(K) < (1+\varepsilon)\zeta$$

*with probability at least $1-\eta$. The algorithm uses*

$$O\left(\frac{n^5}{\varepsilon^2}(\ln\frac{1}{\varepsilon})^3(\ln\frac{1}{\eta})\ln^5 n\right) = O^*(n^5)$$

*oracle calls.*

The proof of this Theorem is given at the end of Section 6.

As in all previous volume algorithms, the main technical tool is *sampling* from $K$, i.e., generating (approximately) uniformly distributed and (approximately) independent random points in $K$. We in fact make use of several sampling algorithms, working under slightly different assumptions. A result that has a simple statement is the following.

**Theorem 2.2** *Given a convex body $K$ satisfying $B \subseteq K \subseteq dB$, a positive integer $N$ and $\varepsilon > 0$, we can generate a set of $N$ random points $\{v_1, \dots, v_N\}$ in $K$ that are*

(a) *almost uniform in the sense that the distribution of each one is at most $\varepsilon$ away from the uniform in total variation distance, and*

(b) *almost (pairwise) independent in the sense that for every $1 \le i < j \le N$ and every two measurable subsets $A$ and $B$ of $K$,*

$$|\mathsf{P}(v_i \in A, v_j \in B) - \mathsf{P}(v_i \in A)\mathsf{P}(v_j \in B)| \le \varepsilon.$$

*The algorithm uses only $O^*\left(n^3 d^2 + N n^2 d^2\right)$ calls on the oracle.*

This running time represents an improvement of $O^*(n)$ over previous algorithms (see Lovász and Simonovits (1993) - Theorem 3.7) for this problem.

To make the sampling algorithm as efficient as possible, we have to find an affine transformation that minimizes the parameter $d$ . Finding an affine transformation $A$ such that

$$B \subseteq AK \subseteq d'B \tag{1}$$

for some small $d'$ is called *rounding* or *sandwiching*. For every convex $K$, the sandwiching ratio $d' = n$ can be achieved (using the so called the Löwner-John ellipsoid) but it is not known how to find the corresponding $A$ in polynomial time. For related references we again recommend Grötschel–Lovász–Schrijver (1988) and the Handbook of Convex Geometry, (1993). For our purposes "approximate sandwiching" is sufficient, where $d'B$ is required to contain most of $K$ but not the whole body. The theorem below will imply that that one can approximately well-round $K$ with $d' = O(\sqrt{n}/\ln(1/\varepsilon))$ using $O^*(n^5)$ oracle calls.

The approximate sandwiching will be done using an important auxiliary result, which may be of interest in its own: an algorithm to find an affine transformation to bring the body to a particularly nice position. To formulate it we need the following definitions.

**Definition 2.3.** Let $K$ be a convex body in $\mathbb{R}^n$, and let $b(K)$ denote its center of gravity. We say that $K$ is in *isotropic position* if its center of gravity is in the origin, and for each $i, j$, $1 \le i \le j \le n$, we have

$$\frac{1}{\mathrm{vol}(K)} \int_K x_i x_j = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j, \end{cases} \tag{2}$$

or equivalently, for every vector $v \in \mathbb{R}^n$,

$$\frac{1}{\mathrm{vol}(K)} \int_K (v^\mathsf{T} x)^2 \, dx = \|v\|^2. \tag{3}$$

(Here $x_i$ denotes the $i$th coordinate of $x$. Note that we normalize differently from, e.g., Milman and Pajor (1987); their definition corresponds to applying a homothetical transformation to get $\mathrm{vol}(K) = 1$.) The isotropic position has many interesting features. Among others, it minimizes $\int_K \|x\|^2/\mathrm{vol}(K)$ (see Milman–Pajor 1987).

If $K$ is in isotropic position, then

$$\frac{1}{\mathrm{vol}(K)} \int_K \|x\|^2 \, dx = n,$$

from which it follows that "most" (i.e., all but a fraction of $\varepsilon$) of $K$ is contained in a ball of radius $\sqrt{\frac{n}{\varepsilon}}$. Using a result of Borell (1975), one can show that the radius of the ball could be replaced by $2\sqrt{2n}\log(1/\varepsilon)$. Also, if $K$ is in isotropic position, it contains the unit ball (cf. Lemma 5.1). It is well known that for every convex body, there is an affine transformation to map it on a body in isotropic position, and this transformation is unique up to an isometry fixing the origin.

In our case, we have to allow an error $\vartheta > 0$, and want to find an affine transformation bringing $K$ into *nearly isotropic position*:

**Definition 2.4.** We say that $K$ is in $\vartheta$-nearly isotropic position ($0 < \vartheta \le 1$), if

$$\|b(K)\| \le \vartheta,$$

and for every vector $v \in \mathbb{R}^n$,

$$(1-\vartheta)\|v\|^2 \le \frac{1}{\mathrm{vol}(K)} \int_{K-b(K)} (v^\mathsf{T} x)^2 \, dx \le (1+\vartheta)\|v\|^2. \tag{4}$$

Our main "sandwiching" result is the following:

**Theorem 2.5** *Given $0 < \eta, \vartheta < 1$, there exists a randomized algorithm finding an affine transformation $A$ for which $AK$ is in $\vartheta$-nearly isotropic position with probability at least $1-\eta$. The number of oracle calls is*

$$O(n^5 \ln(\vartheta\eta) \ln n).$$

In particular, we obtain, with probability at least $1-\eta$, a body $AK$ that is almost contained in $2\sqrt{2n}\log(1/\varepsilon)B$ (for any $0<\varepsilon<1$):

$$\mathrm{vol}(AK \setminus 2\sqrt{2n}\log(1/\varepsilon)B) < \varepsilon \mathrm{vol}(AK).$$

The number of oracle calls is

$$O\left(n^5 \ln\frac{1}{\eta}\ln n\right).$$

In the rest of this section we give an informal description of our algorithms.

## 2.1. Sampling

Our sampling algorithm, like all previous work mentioned above, uses random walks (Markov chains) to sample. We do a random walk on the points of $K$, moving at each step to a uniformly selected random point in a ball of radius $\delta$ about the current point (if this remains inside $K$). If the new point is outside $K$, we stay where we were, and consider the step "wasted". The step-size $\delta$ will be chosen appropriately, but typically it is about $1/\sqrt{n}$.

It follows by elementary Markov chain theory that the distribution of the point after $t$ steps tends to the uniform distribution as $t$ tends to infinity. The crucial issue is, *how long to walk before the walking point becomes nearly uniformly distributed?*

There are two reasons for needing a long walk: we have to get to the "distant" parts of $K$, and we may get stuck in "corners" of $K$. The first reason suggests that we choose a step-size that is large relative to the diameter of the body, while the probability of the second can be reduced by choosing a small step-size.

A main ingredient in our improvement in sampling is to formally separate the two reasons of slowness mentioned above. In Section 4, we show that disregarding wasted steps at the boundary results in another random walk – we call it the *speedy walk* – and that the number of steps this speedy walk needs to get close to its own stationary distribution can be bounded in terms of the diameter/step-size ratio alone. The proof is based on the well-established techniques of *isoperimetric inequalities*, but needs some harder geometric work which is contained in Section 3.

There are two points where we have to pay for the improvement in the running time bounds for the speedy walk:

— First, the stationary distribution for the speedy walk is not uniform. Ideally, we would like to generate uniformly distributed points; but the speedy walk generates points from another distribution, which we call the *speedy distribution*. The density of this distribution $\hat{Q}$ is proportional to the probability of *not* jumping out from $K$ from a given point. So $\hat{Q}$ depends on the step-size. Clearly, deep inside the body the density of $\hat{Q}$ is close to 1, and approaching to a fairly flat portion of the boundary this density drops to (roughly) 1/2. Where $K$ has sharp corners, this density may become very small.

The probability $\lambda$ that making one step from a uniformly distributed random point in $K$ does not take us out of the body (and so the step is not wasted) will be called the *average local conductance*, and will be an important parameter throughout. For example, the total variation distance of the speedy distribution from the uniform is bounded by $(1-\lambda)/\lambda$. If the step-size is not too large, the average local conductance is close to 1, and the speedy distribution is not too far from the uniform. We can either replace the uniform distribution by it (with some care, we can make sure that the errors don't accumulate), or we can use a rejection sampling trick to generate uniformly distributed points. In fact, both of these ideas will be used in our paper.

— Second, while we are ignoring wasted steps at the boundary, they do take time, and we have to bound their number separately. We show that the fraction of steps that are wasted at the boundary can be estimated in terms of the average local conductance; this in turn can be estimated by the surface/volume ratio of the body. To keep this value small, it will suffice to have a large inscribed ball (say, the unit ball). But the key parameter turns out to be the average local conductance, and for our sampling algorithm to work efficiently, it suffices to guarantee that this is not too small.

### 2.2. Rounding

To apply our sampling algorithms, we would like to have $K$ satisfy $B \subseteq K \subseteq dB$ for a small $d$. (Recall that $K \subseteq dB$ implies the upper bound for the number of steps of the speedy walk, while $B \subseteq K$ implies that the average local conductance is large enough so that at most a constant fraction of steps are wasted at the boundary.) With the ellipsoid algorithm, we can achieve $d = O(n^{3/2})$, and quite often a bound $d = O(n^{\text{const}})$ can be achieved by inspection, and we assume such a bound.

Our aim is to bring $K$ into isotropic position, for which, as remarked earlier, we get $d = O^*(\sqrt{n})$ (after discarding a small part of $K$). So, we develop an algorithm (in Section 5) to put $K$ into nearly isotropic position. We have to do this in several steps.

We show in Theorem 5.11 that if we are given $m = O(n^2/\vartheta\eta)$ pairwise "nearly" independent samples, each "nearly" uniformly drawn from a convex body $K$ in $\mathbb{R}^n$, then we can bring $K$ into $\vartheta$-nearly isotropic position with probability at least $1-\eta$, by applying an affine transformation that brings this discrete set of points into isotropic position. (We conjecture that the $O^*(n^2)$ bound for the number of points can be improved to $O^*(n)$.)

Using our sampling algorithm to generate these points, we would need too much time to generate so many sample points, unless we had $K \subseteq O(\sqrt{n})B$ and an average local conductance at least a constant to begin with (where the stepsize is $\delta \approx 1/\sqrt{n}$). So we have to improve the shape of $K$ by other means. We use "bootstrapping" for this, improving the shape of the body $K$ in two different ways.

Assume first that we already have $B \subseteq K \subseteq 10nB$. Scaling down we can achieve $K \subseteq 10\sqrt{n}B$; this cheap fix for making the circumscribed ball small, however, leaves us with the average local conductance possibly very small. Using Algorithm 5.16, we improve the average local conductance to a value close to 1, without increasing the average square distance from the origin. This improvement allows us to generate $O^*(n^2)$ random points in $K$ in the allotted time.

Now consider a general convex body $K$, satisfying only $B \subseteq K \subseteq dB$ with $d = O(n^2)$. To bring it into nearly isotropic position, we need several phases. In each phase, we consider the part of $K$ that is inside a ball with radius $10n$, and apply an affine transformation bringing this part into nearly isotropic position. We prove that $O(\ln n)$ phases of this suffice.

## 2.3. Volume

To calculate the volume of a body $K$, which is already in a *nearly isotropic position*, we apply a multiphase Monte-Carlo algorithm (Algorithm 6.1) as in essentially all previous algorithms. Informally, we consider the convex bodies $K_i = (2^{i/n}B) \cap K$ (figure 1), and estimate the ratios $\mathrm{vol}(K_{i-1})/\mathrm{vol}(K_i)$ by generating $O^*(n)$ random points in $K_i$ and (essentially) counting how often we hit $K_{i-1}$. For technical reasons, we work with the speedy distribution, which means that we count the points with appropriate weights.

## 2.4. Preliminaries

**a. Notation.** The following seven parameters will be used throughout:

$n$ the dimension of the space,

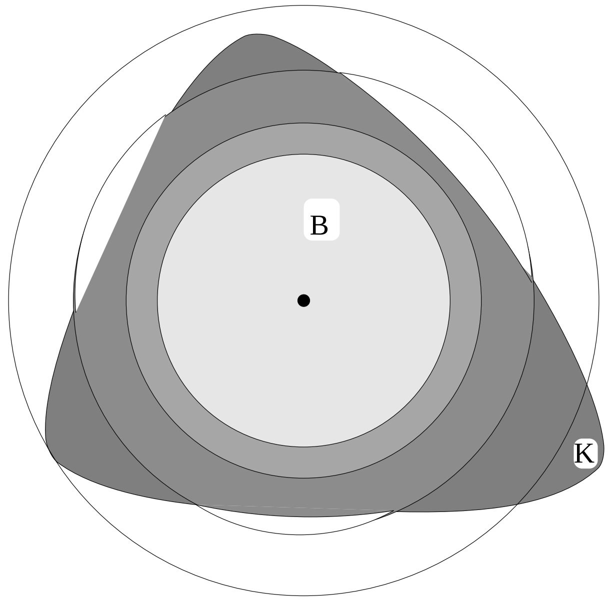


Figure 1: Multi-phase Monte-Carlo

| | |
|---|---|
| $d$ | the radius of a ball containing $K$, |
| $\varepsilon > 0$ | required precision of our volume estimates, |
| $\eta > 0$ | the probability of failure of our estimates, |
| $\vartheta > 0$ | error bound for various rounding algorithms |
| $0 < \lambda < 1$ | the average local conductance |
| $\delta > 0$ | the "step-size" - the radius of the ball in which we move |

The notation $\log x$ means logarithm with base 2. Of course, $\ln x$ denotes the logarithm with base $e$.

**b. Preliminaries from probability.** Let $P$ and $Q$ be two probability distributions on the same $\sigma$-algebra $(\Omega, \mathcal{A})$. The *total variation distance* of $P$ and $Q$ is defined by

$$|P - Q|_{tv} = \max_A (P(A) - Q(A)).$$

It is easy to see that we might as well write

$$|P - Q|_{tv} = \max_A |P(A) - Q(A)|,$$

which implies that the total variation distance of two distributions is exactly half of their $\ell_1$ distance.

It is a very simple but useful fact in our error-estimations that if $X$ is any random variable with distribution $P$ (with values in $\Omega$), then we can construct a random variable $Y$ with distribution $Q$ such that $X = Y$ with probability $1 - |P - Q|_{tv}$.

Throughout, we also will use another measure of how close $P$ and $Q$ are, denoted by $M(P, Q)$ (which we call informally the "$M$"-distance, although it is not a distance, for example, it is not symmetric) defined by

$$M(P,Q) = \sup_S \frac{|P(S) - Q(S)|}{\sqrt{Q(S)}},$$

where $S$ ranges over all $P$- and $Q$-measurable sets with $Q(S) > 0$. While $M(P, Q)$ may be infinite, this will not be the case in our applications of this notion. In fact, in our applications the density $dP/dQ$ of $P$ with respect to $Q$ will be bounded by some $c \geq 1$; then clearly $M \leq c - 1$.

The reason for introducing the $M$-distance is that we need Theorem 2.6 below from the theory of rapid mixing of Markov chains.

A Markov chain is given by a $\sigma$-algebra $(\Omega, \mathcal{A})$ together with a probability measure $P_u$ for every $u \in \Omega$. We fix an initial distribution $Q_0$ and choose an initial element $w_0$ according to this distribution. This generates a random sequence $w_0, w_1, \ldots, w_t, \ldots$ of elements of $\Omega$: given $w_t$ we choose $w_{t+1}$ with probability $P_{w_t}(A)$ from $A$. The distribution of the $t$th element $w_t$ will be denoted by $Q_t$.

The Markov chains we consider will always have a *stationary distribution*, i.e., a probability measure $Q$ on $(\Omega, \mathcal{A})$ such that

$$\int_\Omega P_u(A)\, dQ(u) = Q(A).$$

A Markov chain is *time-reversible* if (roughly speaking) for any two sets $A, B \in \mathcal{A}$, it steps from $A$ to $B$ as often as from $B$ to $A$. Formally, this means that

$$\int_B P_u(A)\, dQ(u) = \int_A P_u(B)\, dQ(u).$$

We call a Markov chain *lazy* if $P_u(u) \geq 1/2$ at each $u \in \Omega$. This condition is technical. Every Markov chain can be made lazy by simply tossing a coin at each step and making a move only if it is tails. This way we eliminate effects of periodicity or almost-periodicity (technically this amounts to a positive semidefinite kernel) at the cost of a slowdown by a factor of 2.

Put

$$\Phi(A) = \int_A P_u(\Omega \setminus A)\, dQ(u).$$

This value is the probability of the event that choosing $w_0$ from the stationary distribution, we have $w_0 \in A$ but $w_1 \notin A$. The *conductance* of the Markov chain is

$$\Phi = \inf_{0<Q(A)<1/2} \frac{\Phi(A)}{Q(A)}.$$

It is well-known (and will also follow from our results below) that if $\Phi > 0$, and the Markov Chain is lazy, then $Q_k \to Q$ in the $\ell_1$ distance. To bound the rate of convergence is a central problem of this field. Of the many results in this direction we will use the following (Lovász and Simonovits, 1993).

**Theorem 2.6.** *If $Q_t$ is the distribution after $t$ steps of our random walk and $Q$ the stationary distribution, then*

$$M(Q_t, Q) \le \left(1 - \frac{\Phi^2}{2}\right)^t M(Q_0, Q).$$

Such a simple exponential convergence does not hold for the total variation distance. This $M$-distance has some similarity to the chi-squared distance, for which Fill (1991) proves a nice exponential convergence as above.

Let $X$ and $Y$ be two random variables with values in possibly different $\sigma$-algebras. We say that $X$ and $Y$ are *$\varepsilon$-independent* if

$$|\mathsf{P}(X \in A, Y \in B) - \mathsf{P}(X \in A)\mathsf{P}(Y \in B)| \le \varepsilon$$

for every two measurable sets $A$ and $B$. This is a rather weak measure of independence, but the following simple lemma gives a convenient way to apply it.

**Lemma 2.7.** *Let $X$ and $Y$ be $\varepsilon$-independent real-valued random variables such that $|X| \le a$ and $|Y| \le b$. Then*

$$|\mathsf{E}(XY) - \mathsf{E}(X)\mathsf{E}(Y)| \le 4\varepsilon ab.$$

**Proof.** Trivial, using that

$$\mathsf{E}(XY) = \int_{-a}^{a} \int_{-b}^{b} \mathsf{P}(X \ge s, Y \ge t)\, ds\, dt,$$

and

$$\mathsf{E}(X)\mathsf{E}(Y) = \int_{-a}^{a} \int_{-b}^{b} \mathsf{P}(X \ge s)\mathsf{P}(Y \ge t)\, ds\, dt. \qquad \square$$

We also introduce a definition used later.

**Definition 2.8.** We call a set $u_1, \ldots, u_k \in K$ of samples $\varepsilon$-*good* for a distribution $\Phi$ if

(a) for the distribution $\Phi_i$ of $u_i$ we have $|\Phi_i - \Phi|_{tv} < \varepsilon$, and

(b) $u_i$ and $u_j$ are $\varepsilon$-independent for all $1 \le i < j \le k$.

**Preliminaries from geometry.** We denote by $\mathrm{vol}_k$ the $k$-dimensional Lebesgue measure, and put $\mathrm{vol} = \mathrm{vol}_n$.

The volume of the unit ball in $\mathbb{R}^n$ is

$$\pi_n = \mathrm{vol}(B) = \frac{\pi^{n/2}}{\Gamma(1 + n/2)}.$$

Therefore $\pi_n/\pi_{n-1} \sim \sqrt{2\pi/n}$. Further, for $n \ge 3$, $\pi_n/\pi_{n-1} > 2/\sqrt{n}$.

Let us also quote the following well-known fact.

**Lemma 2.9** *Let $H$ be a halfspace in $\mathbb{R}^n$ not containing the center of the unit ball $B$. If the distance of $H$ from the center is $1/\sqrt{n} < t \le 1$, then*

$$\frac{1}{10t\sqrt{n}} \frac{\mathrm{vol}(H \cap B)}{\mathrm{vol}(B)} < \frac{1}{t\sqrt{n}}(1 - t^2)^{(n+1)/2} < e^{-nt^2/2}.$$

# 3. An isoperimetric inequality

Let $K$ be a convex body in $\mathbb{R}^n$ and $\delta > 0$. We denote by $B'$ the ball $\delta B$. (Recall that we typically choose $\delta \approx 1/\sqrt{n}$.) We denote the uniform distribution on $x + B'$ by $P_x$. So

$$P_x(U) = \frac{\mathrm{vol}(U \cap (x + B'))}{\mathrm{vol}(B')}$$

for each measurable $U \subseteq \mathbb{R}^n$.

The *local conductance* at $x \in K$ is defined by

$$\ell(x) = P_x(K) = \frac{\mathrm{vol}(K \cap (x + B'))}{\mathrm{vol}(B')}.$$

The local conductance is the probability that if we make a random step in the ball $B'$ around $x \in K$, we stay in $K$. The local conductance plays a crucial role in our paper since one of the main issues throughout will be to keep down the number of wasted steps. One new point in this paper is that we keep only the *average local conductance*

$$\lambda = \frac{1}{\mathrm{vol}(K)} \int_K \ell(x)\, dx$$

close to 1 (and not the pointwise local conductance), which allows us to choose a larger step-size.

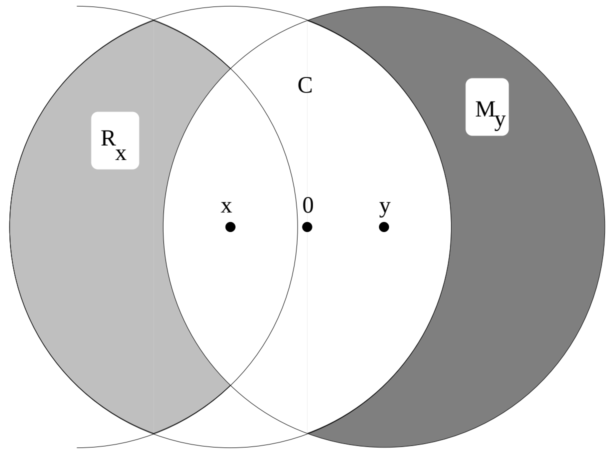


Figure 2: Half-Moon domains

We note that the local conductance is a log-concave function, by the Brunn–Minkowski Theorem. (A function is log-concave if it is non-negative valued, its support is convex, and its logarithm, wherever the function is non-zero, is concave.)

The following theorem is the key to obtaining improved bounds on the mixing time of the random walk.

**Theorem 3.1** *Let $K \subseteq dB$ be a convex body, $d > 32\delta$, and let $K = S_1 \cup S_2$ be a partition of $K$ into two measurable subsets. Then*

$$\int_{S_1} P_x(S_2)\,dx \geq \frac{\delta}{10\sqrt{n}d} \min\left(\int_{S_1} \ell(x)\,dx, \int_{S_2} \ell(x)\,dx\right).$$

Dividing both sides by $\int \ell(x)\,dx$, the above theorem says that the conductance of the "speedy walk" is large.

The proof of Theorem 3.1 needs several lemmas, which aim at showing that if a convex body $K$ has a "large" intersection with each of two "nearby" congruent balls, then it also has a "large" intersection with their intersection.

Let $x, y$ be two points with $\|x - y\| < \delta/\sqrt{n}$. Set

$$C = (x + B') \cap (y + B').$$

and consider the "moons"

$$M_x = (x + B') \setminus (y + B'), \qquad M_y = (y + B') \setminus (x + B').$$

(figure 2). Also set

$$R_x = M_x \cap (x - y + C), \qquad R_y = M_y \cap (y - x + C).$$

Let $C'$ be obtained by blowing up $C$ from its center $\frac{1}{2}(x+y)$ by a factor of $\left(1+\frac{4}{4n-1}\right)$.

**Lemma 3.2.** *$M_x \setminus R_x \subseteq C'$.*

**Proof.** For convenience, assume that $x = -y$. Let $z \in M_x \setminus R_x$. Write $z = \mu x + w$ where $w$ is orthogonal to $x$. It is easy to see that $\|z - x\| \le \delta$, and $\|z - y\|, \|z - 2x + y\| > \delta$ from the hypothesis. It follows that $\mu \in (0, 2)$. Let

$$\alpha = \left(1 + \frac{4}{4n-1}\right)^{-1} = \frac{4n-1}{4n+3} = 1 - \frac{4}{4n+3}.$$

We wish to show that $\|\alpha z - y\| \le \delta$. Clearly,

$$\|\alpha z - y\|^2 = (\alpha\mu + 1)^2\|x\|^2 + \alpha^2\|w\|^2.$$

Here $\|w\|^2 \le \delta^2 - (\mu-1)^2\|x\|^2$, $|\mu| \le 2$, and $\|x\|^2 \le \delta^2/(4n)$. So

$$\begin{aligned}
\|\alpha z - y\|^2 &\le (\alpha\mu+1)^2\|x\|^2 + \alpha^2(\delta^2 - (\mu-1)^2\|x\|^2) \\
&= \left((\alpha\mu+1)^2 - \alpha^2(\mu-1)^2\right)\|x\|^2 + \alpha^2\delta^2 \\
&\le \left((4n+3)\alpha^2 + 4\alpha + 1\right)\frac{\delta^2}{4n}
\end{aligned}$$

Using the actual value of $\alpha$ the coefficient of $\delta^2$ becomes 1. $\square$

**Lemma 3.3.** *For every convex body $K$ containing $x$ and $y$,*

$$\mathrm{vol}(K \cap (M_x \setminus R_x)) \le (e-1)\mathrm{vol}(K \cap C).$$

**Proof.** By Lemma 3.2, blowing up $C$ by a factor of $(1 + \frac{4}{4n-1})$ we cover both $K \cap C$ and $K \cap (M_x \setminus R_x)$. Hence

$$\mathrm{vol}(K \cap (C \cup (M_x \setminus R_x))) \le \mathrm{vol}(K \cap C')$$

$$\le \left(1 + \frac{4}{4n-1}\right)^n \mathrm{vol}(K \cap C) \le e\mathrm{vol}(K \cap C).$$

From this the inequality follows easily. $\square$

**Lemma 3.4.** *For every convex body $K$,*

$$\mathrm{vol}(K \cap C)^2 \ge \mathrm{vol}(K \cap R_x)\mathrm{vol}(K \cap R_y).$$

**Proof.** Consider the function

$$g(u) = \mathrm{vol}(K \cap (u + C)).$$

By the Brunn-Minkowski theorem, this function is log-concave, and so

$$\begin{aligned} g(0)^2 &\geq g(x-y)g(y-x) \\ &= \mathrm{vol}(((x-y)+C)\cap K)\mathrm{vol}(((y-x)+C)\cap K) \\ &\geq \mathrm{vol}(R_x \cap K)\mathrm{vol}(R_y \cap K). \end{aligned}$$

□

**Lemma 3.5.** *For every convex body $K$ containing $x$ and $y$,*

$$\mathrm{vol}(K \cap C) \geq \frac{1}{e+1} \min\{\mathrm{vol}(K \cap (x + B')), \mathrm{vol}(K \cap (y + B'))\}.$$

**Proof.** We have

$$\begin{aligned} \mathrm{vol}(K \cap R_x) &= \mathrm{vol}(K \cap M_x) - \mathrm{vol}(K \cap (M_x \setminus R_x)) \\ &\geq \mathrm{vol}(K \cap (x + B')) - \mathrm{vol}(K \cap C) - (e-1)\mathrm{vol}(K \cap C) \end{aligned}$$

by Lemma 3.3. We also get a symmetric lower bound for $\mathrm{vol}(K \cap R_y)$. Then, by Lemma 3.4, we have

$$\begin{aligned} \mathrm{vol}(K \cap C) &\geq \min\{\mathrm{vol}(K \cap R_x), \mathrm{vol}(K \cap R_y)\} \\ &\geq \min\{\mathrm{vol}(K \cap (x + B')), \mathrm{vol}(K \cap (y + B'))\} - e\mathrm{vol}(K \cap C). \end{aligned}$$

The lemma now follows. □

**Lemma 3.6.** *Let $K$ be a convex body and let $K = S_1 \cup S_2$ be a partition of $K$ into two measurable subsets. Let $x \in S_1$ and $y \in S_2$ be such that $\|x - y\| < \delta/\sqrt{n}$. Then*

$$P_x(S_2) + P_y(S_1) \geq \frac{1}{e+1} \min\{\ell(x), \ell(y)\}.$$

**Proof.** We have

$$P_x(S_2) = \frac{1}{\mathrm{vol}(B')}\mathrm{vol}(S_2 \cap (x + B')) \geq \frac{1}{\mathrm{vol}(B')}\mathrm{vol}(S_2 \cap C),$$

and similarly,

$$P_y(S_1) \geq \frac{1}{\mathrm{vol}(B')} \mathrm{vol}(S_1 \cap C).$$

Hence

$$P_x(S_2) + P_y(S_1) \geq \frac{1}{\mathrm{vol}(B')} \mathrm{vol}(K \cap C).$$

By Lemma 3.5,

$$\begin{aligned} & P_x(S_2) + P_y(S_1) \\ \geq\ & \frac{1}{(e+1)\mathrm{vol}(B')} \min\{\mathrm{vol}(K \cap (x + B')), \mathrm{vol}(K \cap (y + B'))\} \\ =\ & \frac{1}{e+1} \min\{\ell(x), \ell(y)\}. \end{aligned}$$

□

We need two further technical lemmas. The first is an elementary inequality involving the exponential function.

**Lemma 3.7.** *For reals $u_0, u_1$ with $0 < u_0 \leq u_1$ and integer $m \geq 0$, we have*

$$\int_{u_0}^{u_1} e^{-u} u^m \, du \geq \frac{1}{4\sqrt{m}} (1 - e^{u_0 - u_1}) \min \left[ \int_0^{u_0} e^{-u} u^m \, du, \ \int_{u_1}^{\infty} e^{-u} u^m \, du \right].$$

**Proof.** We have trivially

$$\int_{u_0}^{u_1} e^{-u} u^m \, du \geq u_0^m (e^{-u_0} - e^{-u_1}).$$

Thus it suffices to prove that

$$u_0^m e^{-u_0} \geq \frac{1}{4\sqrt{m}} \min \left[ \int_0^{u_0} e^{-u} u^m \, du, \ \int_{u_1}^{\infty} e^{-u} u^m \, du \right].$$

For a fixed $u_0$, this inequality is strongest if $u_1 = u_0$, so we may assume this. Assume first that $u_0 \geq m$, then we claim the second term in the inequality satisfies the condition. Rearranging, we claim that

$$\int_0^{\infty} \left(1 + \frac{x}{u_0}\right)^m e^{-x} \, dx \leq 4\sqrt{m}.$$

It suffices to prove this inequality for the worst case when $u_0 = m$. Returning to the previous variables, we have to prove that

$$\int_m^{\infty} e^{-u} u^m \, du \leq 4\sqrt{m} m^m e^{-m}.$$

Here

$$\int_m^{\infty} e^{-u} u^m \, du < \int_0^{\infty} e^{-u} u^m \, du = m! < \left(\frac{m}{e}\right)^m \cdot 4\sqrt{m}.$$

This proves the assertion when $u_0 \geq m$.

The case $u_0 \leq m$ follows similarly. (As a matter of fact, the two integrals

$$\int_0^m e^{-u} u^m \, du, \quad \int_m^\infty e^{-u} u^m \, du$$

are asymptotically equal to $m!/2$, so the constant 4 could be replaced by $\sqrt{\pi/2}$.) $\square$

The second technical lemma generalizes the last inequality to log-concave functions.

**Lemma 3.8.** *Let $a < x < y < b$ be reals, let $F$ be a log-concave function defined on $[a,b]$, and let $g$ be a non-negative linear function defined on $[a,b]$. Assume that $F(x) \geq F(y)$. Then*

$$\begin{aligned}
&\int_x^y F(t)(g(t))^{n-1} \, dt \\
&\geq \frac{1}{4\sqrt{n}} \frac{F(x) - F(y)}{F(x)} \min\left\{ \int_a^x F(t)(g(t))^{n-1} \, dt, \int_y^b F(t)(g(t))^{n-1} \, dt \right\}.
\end{aligned}$$

**Proof.** Let $\mu = \frac{\ln(F(x)/F(y))}{y-x}$. Let $h(t)$ be the function obtained by linearly interpolating $\ln F(t)$ over $[x,y]$, i.e., let

$$h(t) = \mu(y-t) + \ln F(y).$$

Let $H(t) = e^{h(t)}$. It is easy to see that it suffices to prove the lemma with $F(t)$ replaced by $H(t)$. Let $g(t) = \alpha t + \beta$.

**Case 1.** $\alpha \leq 0$. Then

$$\begin{aligned}
\int_y^b (g(t))^{n-1} H(t) dt &\leq (g(y))^{n-1} \int_y^\infty H(t) dt \\
&= \frac{1}{\mu} (g(y))^{n-1} F(y).
\end{aligned}$$

But,

$$\begin{aligned}
\int_x^y (g(t))^{n-1} H(t) \, dt &\geq (g(y))^{n-1} \int_x^y H(t) \, dt \\
&= \frac{1}{\mu} (g(y))^{n-1} (F(x) - F(y)).
\end{aligned}$$

From these two inequalities and the fact that $F(x) \geq F(y)$, the lemma follows.

**Case 2**. $\alpha > 0$. It is easy to see that for any two reals $A$ and $B$, we have

$$\int_A^B (g(t))^{n-1} H(t) dt = D \int_{(A+\frac{\beta}{\alpha})\mu}^{(B+\frac{\beta}{\alpha})\mu} u^{n-1} e^{-u} du,$$

where $D$ is a constant (independent of $A, B$).

We now apply Lemma 3.7 with $u_0 = (x + \frac{\beta}{\alpha})\mu$ and $u_1 = (y + \frac{\beta}{\alpha})\mu$. $\square$

**Proof of Theorem 3.1.** Define

$$h(x) = \begin{cases} P_x(S_1), & \text{if } x \in S_2, \\ P_x(S_2), & \text{if } x \in S_1. \end{cases}$$

(the probability that a random step from $x$ crosses over). Clearly $h(x) \le \ell(x)$. An easy computation shows that

$$\int_{S_1} h(x)\,dx = \int_{S_2} h(x)\,dx$$

(a long walk steps as often from $S_1$ to $S_2$ as vice versa). Therefore it suffices to show that

$$\int_K h(x)\,dx \ge \frac{\delta}{5d\sqrt{n}} \min\left\{\int_{S_1} \ell(x)\,dx, \int_{S_2} \ell(y)\,dy\right\}$$

Suppose this is false. Setting

$$f_i(x) = \begin{cases} \frac{\delta}{5d\sqrt{n}}\ell(x) - h(x), & \text{if } x \in S_i, \\ -h(x), & \text{if } x \in S_{3-i}, \\ 0 & \text{otherwise} \end{cases}$$

we have

$$\int_{\mathbb{R}^n} f_i(x)\,dx > 0 \tag{5}$$

for $i = 1, 2$.

By the Localization Lemma (2.5 of Lovász–Simonovits 1993), there exist points $a, b$ and a linear function $g(t) \ge 0$ for $t \in [0,1]$ such that setting $x(t) = (1-t)a + tb$, we have

$$\int_0^1 g(t)^{n-1} f_i(x(t))\,dt > 0$$

for $i = 1, 2$. Let $J_i = \{t \in [0,1] : \ x(t) \in S_i\}$ for $i = 1, 2$. Then we have

$$\int_0^1 g(t)^{n-1} h(x(t))\,dt < \frac{\delta}{5d\sqrt{n}} \min_{i=1,2} \int_{J_i} g(t)^{n-1} \ell(x(t))\,dt. \tag{6}$$

We abuse the notation by writing $h(t)$ and $\ell(t)$ for $h(x(t))$ and $\ell(x(t))$.

For $i = 1, 2$, define

$$S_i' = \left\{x \in S_i : h(x) < \frac{1}{9}\ell(x)\right\}.$$

Let $J_i' = \{t \in [0,1] : x(t) \in S_i'\}$ for $i = 1, 2$. Let

$$B = [0,1] \setminus (J_1' \cup J_2').$$

It is easy to see that $S_i'$, $J_i'$ and $B$ are all measurable sets.

Define a measure $\mu$ on $[0,1]$ by

$$\mu(T) = \int_T g(t)^{n-1}\ell(t)\,dt$$

for any measurable subset $T$. Since each point $x(t) \in B$ contributes at least $(1/9)g(t)^{n-1}\ell(t)$ to $\int_0^1 g(t)^{n-1}h(t)$, we get by (6) that

$$\mu(B) < \frac{9}{5}\frac{\delta}{d\sqrt{n}}\min\{\mu(J_1), \mu(J_2)\}$$

whence

$$\mu(B) < \frac{2\delta}{d\sqrt{n}}\min\{\mu(J_1'), \mu(J_2')\}. \tag{7}$$

First we consider the crucial case when $J_1' = [a,x)$, $B = [x,y]$ and $J_2' = (y,b]$ are intervals. If $y - x \geq \delta/\sqrt{n}$, then (7) contradicts Theorem 2.6 of Lovász–Simonovits (1993), applied to the 1-dimensional body $[a,b]$ and the log-concave function $g^{n-1}\ell$. So suppose that $y - x < \delta/\sqrt{n}$. Choose $u < x$ and $v > y$ such that $v - u < \frac{\delta}{\sqrt{n}}$. Let e.g. $\ell(u) \geq \ell(v)$. By Lemma 3.6,

$$h(u) + h(v) \geq \frac{1}{e+1}\ell(v),$$

On the other hand, we have by the choice of $u$ and $v$ that

$$h(u) < \frac{1}{9}\ell(u), \qquad h(v) < \frac{1}{9}\ell(v),$$

whence

$$\ell(v) < \frac{e+1}{9-(e+1)}\ell(u).$$

Letting $u$ tend to $x$ and $v$ to $y$, we get by the continuity of the function $\ell$ that $\ell(y) \leq \frac{e+1}{9-(e+1)}\ell(x)$ (or the other way around). Then we have $\frac{\ell(x)-\ell(y)}{\ell(x)} \geq \frac{1}{4}$. We apply Lemma 3.8 to get

$$\mu(B) \geq \frac{1}{16\sqrt{n}}\min\{\mu(J_1'), \mu(J_2')\},$$

which contradicts (7).

Now we turn to the general case. Let $[x_i, y_i]$ be all maximal intervals contained in $B$. By the special case settled above, we have

$$\mu([x_i, y_i]) \geq \frac{2\delta}{d\sqrt{n}}\mu(T_i),$$

where $T_i$ is either $[a, x_i]$ or $[y_i, b]$. Summing over all $i$, we get

$$\mu(B) \geq \frac{2\delta}{d\sqrt{n}}\mu(\cup_i T_i).$$

To finish, it suffices to notice that either $J_1'$ or $J_2'$ is contained in $\cup_i T_i$. Suppose not. Then, for any $u \in J_1' \setminus \cup T_i$ and $v \in J_2' \setminus \cup T_i$ there is an interval $[x_i, y_i]$ separating them, and then either $u$ or $v$ is contained in $T_i$. $\square$

# 4. Sampling

## 4.1. Speedy Walks

Let $K \subseteq dB \subseteq \mathbb{R}^n$ be a convex body, let $\delta > 0$ and set $B' = \delta B$. We consider a random walk in $K$ defined as follows: we select $v_0$ from some initial distribution $Q_0$. Given $v_k$, we flip a fair coin and if it is heads, we let $v_{k+1} = v_k$. Else, we generate a vector $u$ from the uniform distribution on $B'$, and consider $v_k + u$. If $v_k + u \in K$, we let $v_{k+1} = v_k + u$. (We then call this move a "proper" step.) Else, we let $v_{k+1} = v_k$. We call this walk the *lazy random walk in $K$ with $\delta$-steps*. It is straightforward to see that this is a time-reversible Markov chain and the uniform distribution $Q$ on $K$ is stationary.

We also consider a variation of the above random walk called the **speedy walk in $K$ with $\delta$-steps**: we start from a point $w_0$ drawn from some initial distribution $\tilde{Q}_0$. Given $w_k$, we let $w_{k+1} = w_k$ with probability 1/2 again. Else, we choose $w_{k+1}$ from the uniform distribution on $(w_k + B') \cap K$. (One could "implement" the speedy walk by doing a lazy walk, but selecting only those points which are either different from the previous point or correspond to flipping a "head".)

This defines a time-reversible Markov chain with stationary distribution $\hat{Q}$ given by

$$\hat{Q}(A) = \int_A \mathrm{vol}((x + B') \cap K)\, dx \Big/ \int_K \mathrm{vol}((x + B') \cap K)\, dx$$

We call $\hat{Q}$ the *speedy distribution* (on $K$, for step size $\delta$). In terms of the local conductance $\ell(x)$, defined in Section 3, the density function of $\hat{Q}$, with respect to the uniform distribution $Q$, is

$$\frac{d\hat{Q}}{dQ}(x) = \frac{\ell(x)}{\lambda}.$$

**Theorem 4.1.** *Let $K \subseteq dB$ be a convex body, and let $0 < \delta < d/32$. Let $Q_0$ be any probability distribution on $K$ with $M(Q_0, \hat{Q}) < \infty$. Let $(w_0, w_1, \ldots)$ be a speedy random walk in $K$ with $\delta$-steps, with $w_0$ drawn from $Q_0$. Let $Q_t$ be the distribution of $w_t$. Then we have*

$$M(Q_t, \hat{Q}) \leq M(Q_0, \hat{Q}) \exp\left(-\frac{t\delta^2}{800 d^2 n}\right).$$

*Further, for*

$$\tau = [M(Q_0, \hat{Q}) + 1] \exp\left(-\frac{t\delta^2}{800d^2n}\right),$$

*the random points $w_0$ and $w_t$ are $\tau$-independent.*

**Remark.** This means that $M(Q_t, \hat{Q})$ drops by a factor of $\frac{1}{e}$ in

$$800n\left(\frac{d}{\delta}\right)^2$$

steps. Hence, if we need $T$ steps to achieve a precision $\varepsilon$, then $\varepsilon/n^{10}$ can be achieved in almost the same number of steps: we loose only a factor of $O(\ln n)$. In other words, the precision is not a crucial issue in sampling.

**Proof.** Theorem 3.1 says that the conductance of the speedy random walk is at least $\frac{1}{2}\frac{\delta}{10d\sqrt{n}}$ (the 1/2 comes from the laziness) because the steady state probability densities are proportional to $\ell$. Then Theorem 2.6 directly yields the first result.

To see the $\tau$-independence, we argue as follows. Let $A, B$ be any two measurable sets. Let

$$\begin{aligned} f(A,B) &= |\mathsf{P}(w_0 \in A, w_t \in B) - \mathsf{P}(w_0 \in A)\mathsf{P}(w_t \in B)| \\ &= Q_0(A)|\mathsf{P}(w_t \in B \mid w_0 \in A) - Q_t(B)| \end{aligned} \tag{8}$$

Let $Q'_0$ be the distribution of $w_0$ conditioned on it being in $A$, i.e., $Q'_0(S) = Q_0(S \cap A)/Q_0(A)$ for any measurable $S$. Then, $\mathsf{P}(w_t \in B | w_0 \in A)$ is the distribution of $w_t$ if we start with $w_0$ drawn from $Q'_0$. So, applying the first part of the theorem, we get

$$|\mathsf{P}(w_t \in B \mid w_0 \in A) - \hat{Q}(B)| \le \sqrt{\hat{Q}(B)} M(Q'_0, \hat{Q}) \exp\left(-\frac{t\delta^2}{800d^2n}\right). \tag{9}$$

For any measurable $S$,

$$\hat{Q}(S) = \frac{1}{Q_0(A)}\hat{Q}(S \cap A) - \left(\frac{1}{Q_0(A)} - 1\right)\hat{Q}(S \cap A) + \hat{Q}(S \setminus A),$$

and hence

$$\begin{aligned} |Q'_0(S) - \hat{Q}(S)| \le\ & \frac{1}{Q_0(A)}|Q_0(S \cap A) - \hat{Q}(S \cap A)| \\ & + \left(\frac{1}{Q_0(A)} - 1\right)\hat{Q}(S \cap A) + \hat{Q}(S \setminus A). \end{aligned}$$

So, we have $M(Q'_0, \hat{Q}) \le \frac{1}{Q_0(A)}[M(Q_0, \hat{Q}) + 1]$ and now using (8) and (9), and the first part of this theorem, we get the claimed $\tau$-independence. $\square$

The case when we start from a given point $u$ needs a little additional care since $M$, as defined in the theorem, is infinite. But we can apply the theorem after making the first step. This will put us uniformly in the set $S = (u + \delta B) \cap K$.

**Corollary 4.2.** *Let $K \subseteq dB$ be a convex body and $d > 32\delta$. Let us start a speedy random walk from an arbitrary point of $K$. We have*

$$M(Q_{t+1}, \hat{Q}) \le \left(\frac{d}{\delta}\right)^n \exp\left(-\frac{t\delta^2}{800d^2n}\right).$$

**Proof.** Let $S = (u + \delta B) \cap K$. It is easy to see that $M(Q_1, \hat{Q}) \le 1/\sqrt{\hat{Q}(S)}$. Now, for any $x$, $\ell(x)$ satisfies $\ell(x)\mathrm{vol}(\delta B)(d/\delta)^n \ge \mathrm{vol}(K)$, since blowing up $(x + \delta B) \cap K$ from $x$ by a factor of $d/\delta$ covers $K$. So, for any two $x, y$, we have

$$\frac{\ell(x)}{\ell(y)} \ge \left(\frac{\delta}{d}\right)^n \frac{\mathrm{vol}(K)}{\mathrm{vol}(\delta B)\ell(y)} \ge \left(\frac{\delta}{d}\right)^n.$$

Now,

$$\hat{Q}(S) = \frac{\int_S \ell(x)\ dx}{\int_K \ell(x)\ dx} \ge \frac{\mathrm{vol}(S)}{\mathrm{vol}(K)} \left(\frac{\delta}{d}\right)^n \ge \left(\frac{\delta}{d}\right)^{2n},$$

the last inequality following again from the above argument about blowing up. This completes the proof. □

## 4.2. Local conductance

If we are to implement a speedy random walk, and the current point $v$ has small local conductance $\ell(v)$, then we have to carry out about $1/\ell(v)$ membership tests before a step in the speedy walk can be generated. Hence the bound on the number of steps of the speedy walk inferred from Theorem 4.1 does not reflect the full time-complexity of the algorithm. In fact, note that choosing $\delta = d$, the speedy walk yields a uniformly distributed point of $K$ in a single step!

Thus it is important to have good bounds on the average local conductance. Our bounds are based on the following lemma (for later use, we formulate it more generally than needed right now).

**Lemma 4.3.** *Let $L$ be a measurable subset of the surface of a convex set $K$ in $\mathbb{R}^n$ and let $S$ be the set of pairs $(x, y)$ with $x \in K$, $y \notin K$, $\|x - y\| \le \delta$, and such that the line segment $xy$ intersects $L$. Then the $(2n)$-dimensional measure of $S$ is at most*

$$\delta \mathrm{vol}_{n-1}(L) \frac{\pi_{n-1}}{(n+1)\pi_n} \mathrm{vol}(\delta B).$$

**Proof.** It suffices to prove the assertion for the case when $L$ is "infinitesimally small". In this case, the measure of $S$ is maximized when the surface of $K$ is a hyperplane in a large neighborhood of $L$. Then the measure of $S$ is independent of $K$ and, by a direct computation, is in fact equal to the upper bound given. □

**Corollary 4.4.** *Let $K$ and $L$ be as in Lemma 4.3. Choose $x$ uniformly from $K$ and choose $u$ uniformly from $\delta B$. The probability that $[x, x+u]$ intersects $L$ is at most*

$$\frac{\delta \mathrm{vol}_{n-1}(L)}{2\sqrt{n}\mathrm{vol}(K)}.$$

This bound gives us a lower bound on the average local conductance in terms of the surface/volume ratio of $K$.

**Corollary 4.5.** *The average local conductance $\lambda$ with respect to $\delta$-steps satisfies*

$$\lambda \geq 1 - \frac{\delta}{2\sqrt{n}} \frac{\mathrm{vol}_{n-1}(\partial K)}{\mathrm{vol}(K)}.$$

While the surface/volume ratio of an implicitly given convex body may be difficult to estimate, the following, slightly weaker bound is easier to apply.

**Corollary 4.6.** *If $K$ contains a ball of radius $r$, then*

$$\lambda \geq 1 - \frac{\delta\sqrt{n}}{2r}.$$

**Proof.** In this case,

$$\mathrm{vol}(K) \geq \frac{r}{n} \mathrm{vol}_{n-1}(\partial K).$$ □

### 4.3. Sampling by random walk

Now we turn to the problem of sampling. We describe two algorithms to sample from the distribution $\hat{Q}$; then we show a simple trick to use this to get a sample from the uniform distribution $Q$.

We assume that we already have a "reasonably good" starting point in the sense that the distribution is near the stationary. We will also assume that the average local conductance is at least 0.95. Soon we will see an important situation where both of these assumptions are valid.

**Algorithm 4.7.** [Sampling from a random starting point] Let $K \subseteq dB$ be a convex body and let $\delta, \varepsilon > 0$. Start a lazy random walk with $\delta$-steps from a distribution $Q_0$. Output the point we have immediately after

$$t = \left\lceil 801n \ln \frac{5}{\varepsilon} \left(\frac{d}{\delta}\right)^2 \right\rceil$$

*proper* steps.

To formulate the theorem analyzing this algorithm, we need the following definition. (We are quite sure that the theorem remains valid with "ordinary" expectation, but cannot prove it at the moment.)

**Definition 4.8.** Given a random variable $Z$, we say that it has expectation at most $E$ with exception $\varepsilon$ if there is an event $A$ with probability at least $1-\varepsilon$ such that the expectation of $Z$ conditional on $A$ is at most $E$.

**Remark 4.9.** For this and other algorithms in the paper, we will prove that their expected running time is at most a certain $T_0$ with exception $s$. The probability that we take more than $2T_0$ steps is clearly at most $(1+s)/2$.

We may change the algorithm description to say : if a run of the algorithm takes more than $2T_0$ steps, abandon it and start a new run with independent coin tosses; if we take more than $k$ runs, declare the algorithm a failure and stop. It is easy to see that if $\frac{1}{2} - s = \Omega(1)$, then the above algorithm will succeed with probability at least $1 - 2^{-\Omega(k)}$; also, the algorithm will now *always* take only $O(kT_0)$ steps.

**Theorem 4.10.** *Assume that* $0 < \delta < d/32$ *and that the average local conductance of* $K$ *with respect to* $\delta$*-steps is* $\lambda \geq .95$.

(a) *Assume that the starting distribution* $Q_0$ *satisfies* $M(Q_0, \hat{Q}) \leq 2$. *Then the distribution* $Q_f$ *of the point returned by Algorithm 4.7 satisfies*

$$M(Q_f, \hat{Q}) \leq \varepsilon.$$

*The starting point and the point returned by the algorithm are* $\varepsilon$*-independent.*

(b) *Assume that* $M_0 = M(Q_0, \hat{Q}) < \frac{1}{\sqrt{2}}$. *Then the expected number of proper and improper steps with exception* $2M_0^2$ *is at most*

$$\frac{2t}{\lambda(1-2M_0^2)}.$$

**Proof.** (a) The bound on $M(Q_f, \hat{Q})$ and the $\varepsilon$-independence follow from Theorem 4.1.

(b) The statement on the number of calls is a bit trickier. Let $V \subseteq K$ be the set of points $x$ where $dQ_0(x)/d\hat{Q}(x) > 2$. Then, $V$ is measurable and $Q_0(V) - \hat{Q}(V) \le M_0\sqrt{\hat{Q}(V)}$. Also, $Q_0(V) \ge 2\hat{Q}(V)$. From these two, it follows that $Q_0(V) \le 2M_0^2$. Let $A$ be the event that the starting point does not belong to $V$. Let $Z$ be the number of steps of the algorithm. Then, $\mathsf{P}(A) > 1 - 2M_0^2$. So it suffices to show that $E(Z|A)$ is at most $2t/(\lambda(1-2M_0^2))$. To this end, let $Q_i'$ be the distribution and $w_i$ the current point after the $i$th proper step conditioned on the event $A$. Then, for all measurable $S$,

$$Q_0'(S) = \frac{Q_0(S \setminus V)}{Q_0(K \setminus V)} \le \frac{2\hat{Q}(S \setminus V)}{1 - 2M_0^2}.$$

Then we see by induction on $i$ that

$$Q_i'(S) \le \frac{2}{1 - 2M_0^2}\hat{Q}(S).$$

If $w_i = x$, the expected number of (proper and improper) steps before the next proper move is $1/\ell(x)$. So the expected number of steps between the $i$th and $(i+1)$st proper move conditioned on $A$ is

$$\int_K \frac{1}{\ell(x)} dQ_i'(x) \le \frac{2}{1 - 2M_0^2}\int_K \frac{1}{\ell(x)} d\hat{Q}(x) = \frac{2}{\lambda(1 - 2M_0^2)}.$$

So, by linearity of expectation, it follows that $E(Z|A) \le 2t/(\lambda(1-2M_0^2))$. □

Now we address the question of how to obtain a "reasonably good" starting point. Note that the condition on the starting distribution was needed not only because we applied Theorem 4.1, but also because we had to bound the expected number of "wasted" steps. So we could not simply sacrifice a factor of (about) $n \ln n$ and just invoke Corollary 4.2. Instead, we use the "chain of bodies" trick as in all previous work. This construction will also be fundamental in the volume algorithm.

Let $K$ be a convex body with $B \subseteq K \subseteq dB$. Define $m = \lceil n \log d \rceil$ and $K_i = K \cap 2^{i/n}B$ $(i = 1, \ldots, m)$. Clearly

$$B = K_0 \subseteq K_1 \subseteq \ldots \subseteq K_m = K.$$

An important feature of these bodies is that

$$\mathrm{vol}(K_i) \le 2\mathrm{vol}(K_{i-1}).$$

We denote by $\ell_i(x)$ the local conductance of $K_i$ (with respect to $\delta$-steps), by $Q_i$, the uniform distribution on $K_i$, and by $\hat{Q}_i$, the speedy distribution on $K_i$. We will use this notation throughout.

**Algorithm 4.11.** [Sampling from a fixed starting point] Let $K$ be a convex body such that $B \subseteq K \subseteq dB$. Let $0 < \varepsilon < 1/(4m)$ be given.

Define

$$\delta = \frac{1}{10\sqrt{n \log(m/\varepsilon)}}.$$

Choose a point $w_0$ according to $\hat{Q}_0$ from $B$. For $i = 1, 2, \ldots m$, execute Algorithm 4.7 in $K_i$ starting at $w_{i-1}$ to get a sample point $w_i \in K_i$. Output the sequence $(w_i : \ i = 0, 1, 2, \ldots m)$.

**Remark 4.12.** Choosing $w_0$ from $\hat{Q}_0$ is easy to do : first pick $w_0$ according to the uniform distribution on $B$. Then pick a point $u$ uniformly from $\delta B$. If $w_0 + u \notin B$, then reject $w_0$ and repeat. Otherwise choose $w_0$. It is easy to see that we do expect to reject only $O(1)$ times.

**Remark 4.13.** The choice of $\delta$ will become clear in the next algorithm.

**Theorem 4.14.** *Let $P_i$ be the distribution of $w_i$ returned by Algorithm 4.11. Then $M(P_i, \hat{Q}_i) \leq \varepsilon$ for each $i$. The $w_i$ are pairwise $\varepsilon$-independent.*

*Further, the expectation of the number of calls on the oracle, with exception $2m\varepsilon^2$, is at most*

$$3mt = O\left(n^3 d^2 \left(\ln \frac{m}{\varepsilon}\right)^2 \ln d\right).$$

**Proof.** It follows from Corollary 4.6 that the average local conductance of each $K_i$ is at least 0.95. We prove by induction on $i$ that for all $i = 1, 2, \ldots, m$,

$$M(P_i, \hat{Q}_i) \leq \varepsilon. \tag{10}$$

This is obvious for $i = 0$. Let $S$ be any measurable subset of $K_i$. Then, by definition,

$$\hat{Q}_i(S) = \int_S \ell_i(x)\, dx \Big/ \int_{K_i} \ell_i(x)\, dx.$$

Further, if $S \subseteq K_{i-1}$, then

$$\hat{Q}_{i-1}(S) = \frac{\int_S \ell_{i-1}}{\int_{K_{i-1}} \ell_{i-1}} \leq \frac{\int_S \ell_i}{\int_{K_{i-1}} \ell_{i-1}} = \frac{\int_{K_i} \ell_i}{\int_{K_{i-1}} \ell_{i-1}} \hat{Q}_i(S).$$

Here

$$\int_{K_i} \ell_i \leq \mathrm{vol}(K_i),$$

and

$$\int_{K_{i-1}} \ell_{i-1} \geq \frac{19}{20}\mathrm{vol}(K_{i-1}) \geq \frac{19}{40}\mathrm{vol}(K_i).$$

Thus

$$\hat{Q}_{i-1}(S) \le \frac{40}{19}\hat{Q}_i(S). \tag{11}$$

Now let $S \subseteq K_i$ and let $S_1 = S \cap K_{i-1}$, $S_2 = S \setminus K_{i-1}$. We have

$$|P_{i-1}(S) - \hat{Q}_i(S)| \le |P_{i-1}(S_1) - \hat{Q}_{i-1}(S_1)| + |\hat{Q}_{i-1}(S_1) - \hat{Q}_i(S_1)| + \hat{Q}_i(S_2)$$

Here the first term is at most $\varepsilon\sqrt{\hat{Q}_{i-1}(S_1)}$ by the induction hypothesis. The second term is at most $(21/19)\hat{Q}_i(S_1)$ by (11). Combining these and using (11) again, we get

$$|P_{i-1}(S) - \hat{Q}_i(S)| \le \varepsilon\sqrt{\hat{Q}_{i-1}(S_1)} + \frac{21}{19}\hat{Q}_i(S_1) + \hat{Q}_i(S_2) < 2\sqrt{\hat{Q}_i(S)}.$$

From this it follows that $M(P_{i-1}, \hat{Q}_i) \le 2$ and now applying Theorem 4.10, we get that $M(P_i, \hat{Q}_i) \le \varepsilon$, as promised.

The pairwise $\varepsilon$-independence follows from Theorem 4.10.

The statement on the expected number of steps follows directly from Theorem 4.10(b), noting that from the previous proof we can take the $M_0$ there to be $\varepsilon$. □

Assume now that we want to get a sample from an approximately uniform distribution $Q$ on $K$, using any Algorithm $\mathcal{A}$ (for example, Algorithm 4.11) that generates independent samples from some distribution $P$ on $K$.

**Algorithm 4.15.** [Uniform sampling] Use Algorithm $\mathcal{A}$ to generate independent points $u_1, u_2, \dots$ from $P$ until a point $u_i$ is obtained with $(2n/(2n-1))u_i \in K$. Return $v = (2n/(2n-1))u_i$.

**Theorem 4.16.** *Assume that $|P - \hat{Q}|_{tv} \le \varepsilon$, $B \subseteq K$, and*

$$\delta \le \frac{1}{\sqrt{8n\ln(n/\varepsilon)}}.$$

*Then the distribution $P'$ of the point $v$ returned by Algorithm 4.15 satisfies*

$$|P' - Q|_{tv} < 10\varepsilon.$$

*The expected number of calls on Algorithm $\mathcal{A}$ is at most 2.*

**Proof.** Put $c = 1 - 1/(2n)$. We start by estimating the average local conductance on $cK$. We have

$$\mathrm{vol}(B')\int_{cK}(1 - \ell(x))dx = \mathrm{vol}_{2n}(\{(x,y) \in \mathbb{R}^{2n} : x \in cK, y \notin K, \|x - y\| \le \delta\}).$$

Let $t > 0$. For each $y \notin (1+t)K$, we have that there is a tangent hyperplane to $(1+t)K$ separating $y$ and the parallel tangent hyperplane to $cK$ is at least (in perpendicular distance) $1-c+t$ away (since $B \subseteq K$). So we get, by Lemma 2.9, that

$$\mathrm{vol}((y+B') \cap (cK)) \leq e^{-n(1-c+t)^2/(2\delta^2)} \mathrm{vol}(B') \leq e^{-1/(8n\delta^2)} \cdot e^{-t/(2\delta^2)} \mathrm{vol}(B').$$

The volume of $(1+t+dt)K \setminus (1+t)K$ is $n(1+t)^{n-1}\mathrm{vol}(K)\,dt$ which is at most $ne^{nt}\mathrm{vol}(K)\,dt$, and so by integrating with respect to $t$, we get

$$\int_{cK} (1-\ell(x))\,dx \leq \mathrm{vol}(K)\frac{1}{2}e^{-1/(8n\delta^2)} < \varepsilon \mathrm{vol}(cK).$$

Hence it follows that

$$\hat{Q}(cK) > \frac{1}{2}.$$

Now we return to the proof. Clearly, the distribution of $u_i$, the first point satisfying $(1/c)u_i \in K$, is proportional to the restriction of $P$ to $cK$. Therefore for every $S \subseteq K$,

$$P'(S) - Q(S) = \frac{P(cS)}{P(cK)} - Q(S) \leq \frac{\hat{Q}(cS) + \varepsilon}{\hat{Q}(cK) - \varepsilon} - Q(S)$$

Here

$$\hat{Q}(cS) \leq \frac{\mathrm{vol}(cS)}{\int_K \ell},$$

and

$$\hat{Q}(cK) = \frac{\mathrm{vol}(cK) - \int_{cK}(1-\ell)}{\int_K \ell} \geq \frac{\mathrm{vol}(cK) - \varepsilon\mathrm{vol}(cK)}{\int_K \ell}.$$

Hence

$$P'(S) - Q(S) \leq \frac{\mathrm{vol}(cS) + \varepsilon \int_K \ell}{\mathrm{vol}(cK) - \varepsilon\mathrm{vol}(cK) - \varepsilon \int_K \ell} - \frac{\mathrm{vol}(cS)}{\mathrm{vol}(cK)} < 10\varepsilon$$

because we may assume that $\varepsilon < 1/10$.

The fact that the expected number of trials is less than 2 follows from $\hat{Q}(cK) > 1/2$. □

**Proof of Theorem 2.2.** Suppose we wish to draw $N$ samples from $K$ satisfying $B \subseteq K \subseteq dK$ each nearly uniformly distributed and pairwise nearly independent. We can first apply Algorithm 4.11 to produce a point $w$ from a distribution $Q_1$ within $M$ distance 2 of $\hat{Q}$ in time $O^*(n^3d^2)$. Then we may apply Algorithm 4.7 starting from $w$ and run it $N$ times (with $\delta$ as in Algorithm 4.11) to get $3N$ points (in total time $O^*(n^2d^2N)$) which is an $\varepsilon$-good sample for a distribution $P$ satisfying the hypothesis of Theorem 4.16. We then apply Algorithm 4.15 to get (with high probability) the $N$ samples we need. The overall time for this process is $O^*(n^3d^2 + n^2d^2N)$ and as remarked in the Introduction, this is an improvement of $O^*(n)$ over previous algorithms. □

## 5. Transforming into isotropic position

It will be convenient to introduce the following notation: given a convex body $K \subseteq \mathbb{R}^n$ and a function $f:\ K \to \mathbb{R}^m$, we denote by $E_K(f)$ the "average of $f$ over $K$", i.e.,

$$E_K(f) = \frac{1}{\mathrm{vol}(K)} \int_K f(x)\, dx.$$

We denote by $b = b(K) = E_K(x)$ the center of gravity (baricenter) of $K$, and by $A(K)$ the $n \times n$ matrix

$$E_K((x-b)(x-b)^{\mathsf{T}}).$$

The trace of $A(K)$ is the average square distance of points of $K$ from the center of gravity, which we also call the *second moment* of $K$.

We recall from the introduction the definition of the isotropic position. The body $K \subseteq \mathbb{R}^n$ is in isotropic position if and only if $b(K) = 0$ and $A(K) = I$, the identity matrix. In this case we have $E_K(x_i) = 0$, $E_K(x_i^2) = 1$, $E_K(x_i x_j) = 0$. The second moment of $K$ is $n$, and therefore all but a fraction of $\varepsilon$ of its volume lies inside the ball $\sqrt{\frac{n}{\varepsilon}}B$.

We also need the following fact about isotropic bodies:

**Lemma 5.1.** *If $K$ is in isotropic position, then*

$$\sqrt{\frac{n+2}{n}}B \subseteq K \subseteq \sqrt{n(n+2)}B.$$

It will be enough to use the weaker but more convenient relations

$$B \subseteq K \subseteq (n+1)B.$$

This lemma is, in a sense, folklore. For centrally symmetric bodies, the corresponding result (in which case the bounds are somewhat sharper, but only by absolute constants) was proved by Milman and Pajor (1987). For the non-symmetric case, the containment, up to absolute constants, was proved by Sonnevend (1989). For a detailed proof see Kannan, Lovász and Simonovits (1995). Since we deal occasionally with nearly isotropic bodies, we need also the following version:

**Corollary 5.2.** *Let $\vartheta < 1/2$. If $K$ is in $\vartheta$-near isotropic position, then*

$$(1-2\vartheta)B \subseteq K \subseteq (1+2\vartheta)(n+1)B.$$

**Corollary 5.3.** *Let $K$ be a convex body and $f$, a linear function on $K$. Then*

$$E_K(f^2) \leq E_K(f)^2 + \left(\max_K f - E_K(f)\right)^2.$$

Indeed, when $K$ is in isotropic position, $E_K(f) = 0$, and $E_K(f^2) = 1$, then the assertion is just the second inequality in Lemma 5.1. The general case follows on noticing that the assertion is invariant under applying any affine transformation to $K$ as well as adding a constant to $f$ and scaling $f$.

As a special case we obtain that if the center of gravity of a convex body $K$ is the origin and $K$ is contained in the half-space $x_1 \leq c$, then

$$E_K(x_1^2) \leq c^2.$$

We shall need the following, slightly more general assertion.

**Lemma 5.4.** *If the center of gravity of a convex body $K$ is contained in $-\alpha K$, for $\alpha > 0$, and $K$ is contained in the half-space $x_1 \leq c$, then*

$$E_K(x_1^2) \leq (1 + 2\alpha + 2\alpha^2)c^2.$$

**Proof.** Since $b \in (-\alpha K)$, we have

$$E_K(x_1) = b_1 \geq -\alpha c.$$

We also have by assumption

$$\max_k x_1 \leq c.$$

So, by Corollary 5.3, we have

$$E_K((x_1)^2) \leq c^2 + (c + \alpha c)^2 = (1 + 2\alpha + 2\alpha^2)c^2.$$

□

## 5.1. Finding the center of gravity

The center of gravity of a convex body can be found in the obvious way, by drawing a sufficiently large sample and computing the center of gravity of this sample. The only issue is to estimate the error.

One cause of error is that we will have to use random points from $\hat{Q}$ rather than from $Q$. The following lemma estimates the error in the center of gravity and in the moment of the body that this causes.

**Lemma 5.5.** *Let $K$ be a convex body in $\mathbb{R}^n$ in isotropic position. Let $Q$ be the uniform distribution on $K$, and let $P$ be any other distribution on $K$, such that $P \le (1+\alpha)Q$ for some $0 < \alpha < 1/10$. Then for any unit vector $v$, we have,*

$$\left| \int_K v \cdot x \, dP(x) \right| \le 5\alpha \ln \frac{1}{\alpha}$$

*and*

$$1 - 26\alpha \ln^2 \frac{1}{\alpha} \le \int_K (v \cdot x)^2 \, dP(x) \le 1 + \alpha.$$

**Proof.** Without loss of generality, we may assume $v \cdot x = x_1$. The definition of $\alpha$ implies that we have a probability measure $R$ on $K$ such that

$$P = (1+\alpha)Q - \alpha R.$$

Let $F(t)$ denote the $(n-1)$-dimensional volume of the intersection of $K$ with the hyperplane $x_1 = t$, divided by $\mathrm{vol}(K)$. Then $F(t)$ is log-concave by the Brunn–Minkowski theorem, and the isotropic position of $K$ implies that

$$\int_{-\infty}^{\infty} t^2 F(t) \, dt = 1. \tag{12}$$

Let

$$G(t) = \int_t^{\infty} F(u) \, du.$$

$G$ is also log-concave. By a theorem of Grünbaum (1960), a hyperplane through the center of gravity of a convex body $K$ has at least a fraction of $1/e$ of the volume on each side. Hence $\frac{1}{e} \le G(0) \le 1 - \frac{1}{e}$. Let $\beta$ be chosen so that

$$G(\beta) = \frac{\alpha}{1+\alpha}. \tag{13}$$

Clearly, $\beta > 0$. Fix $A, B$ so that $Ae^{-Bt} = G(t)$ for $t = 0$ and $t = \beta$; thus $A = G(0)$ and $B = L/\beta$ for $L = \ln((1+\alpha)A/\alpha) > 1$.

First we show that $\beta$ cannot be too large. Integrating by parts, we have that

$$\begin{aligned} 1 &\ge \int_0^{\infty} t^2 F(t) \, dt = 2 \int_0^{\infty} tG(t) \, dt \ge 2 \int_0^{\beta} tG(t) \, dt \\ &\ge 2A \int_0^{\beta} te^{-Bt} \, dt = 2A \int_0^{\beta} te^{-Lt/\beta} \, dt = 2A \frac{\beta^2}{L^2} \int_0^L se^{-s} \, ds \\ &\ge 2A \frac{\beta^2}{L^2} \int_0^1 se^{-s} \, ds = 2A \frac{\beta^2}{L^2} \left(1 - \frac{2}{e}\right). \end{aligned}$$

Hence

$$\beta \leq \sqrt{\frac{1}{2A(1-(2/e))}}\, L. \tag{14}$$

Let $c_1 = \int_K x_1\, dR(x)$. We claim that

$$c_1 \leq \frac{1+\alpha}{\alpha} \int_\beta^\infty tF(t)\, dt. \tag{15}$$

In fact, the density of $R$ is at most $(1+\alpha)/\alpha$, and hence is maximized if $R$ is concentrated on the fraction of $K$ of size $\alpha/(\alpha+1)$ where $x_1$ is largest. This proves (15). To estimate $c_1$, we use

$$\begin{aligned}
&\int_\beta^\infty tF(t)\, dt = [-tG(t)]_\beta^\infty + \int_\beta^\infty G(t)\, dt \\
\leq\ & \frac{\alpha\beta}{1+\alpha} + A\int_\beta^\infty e^{-Bt}\, dt = \frac{\alpha\beta}{1+\alpha} - \frac{A}{B}\left[e^{-Bt}\right]_\beta^\infty = \frac{\alpha\beta}{1+\alpha} + \frac{A\beta}{L} e^{-L} \\
\leq\ & \frac{5}{2}\frac{\alpha}{1+\alpha} L + \frac{5}{2}\frac{\alpha}{1+\alpha} < 5\frac{\alpha}{1+\alpha} L.
\end{aligned}$$

Hence $c_1 \leq 5L$. Consequently,

$$\int_K x_1\, dP = -\alpha \int_K x_1\, dR \geq -5\alpha L \geq -5\alpha \ln\left(\frac{1}{\alpha}\right).$$

This proves the first inequality, because we can apply the same argument with $-v\cdot x$ instead of $v \cdot x$ as well.

To prove the second inequality we start with

$$\int_K x_1^2\, dQ = (1+\alpha)\int_K x_1^2\, dP - \alpha\int_K x_1^2\, dR = 1+\alpha-\alpha\int_K x_1^2\, dR.$$

This implies the upper bound immediately.

Let $\gamma$ be chosen so that

$$\int_{|t|\geq\gamma} F(t) = \frac{\alpha}{1+\alpha}.$$

Let

$$\alpha_1 = \int_{t\geq\gamma} F(t)\, dt \qquad \text{and} \qquad \alpha_2 = \int_{t\leq-\gamma} F(t)\, dt.$$

Define $L_i = \ln((1+\alpha_i)A/\alpha_i)$. Clearly $\max\{\alpha_1, \alpha_2\} \geq \alpha/2$ and hence $\min\{L_1, L_2\} \leq \ln(2A(1+\alpha)/\alpha)$. To estimate $\gamma$ from above, we apply (14) to get

$$\gamma \leq \sqrt{\frac{1}{2A(1-(2/e))}}\, L_1,$$

and similarly,

$$\gamma \le \sqrt{\frac{1}{2A(1-(2/e))}}\, L_2.$$

Hence (by $A > \frac{1}{e}$)

$$\gamma \le \sqrt{\frac{1}{2A(1-(2/e))}}\ \ln\frac{2A(1+\alpha)}{\alpha} < 2.27 \ln\frac{1}{\alpha}.$$

Similarly as before

$$\int_K x_1^2\, dR(x) \le \frac{1+\alpha}{\alpha}\int_{|t|>\gamma} t^2 F(t)\, dt. \tag{16}$$

Here

$$\begin{aligned}
\int_\gamma^\infty t^2 F(t)\, dt &= \left[-t^2 G(t)\right]_\gamma^\infty + 2\int_\gamma^\infty tG(t)\, dt \\
&\le \alpha_1\gamma^2 + 2A\int te^{-L_1 t/\gamma}\, dt = \alpha_1\gamma^2 + 2Ae^{-L_1}\left(\frac{\gamma^2}{L_1} + \frac{\gamma^2}{L_1^2}\right) \\
&= \alpha_1\gamma^2\left(1 + \frac{2}{L_1} + \frac{2}{L_1^2}\right) \le 5\alpha_1\gamma^2
\end{aligned}$$

(since $L_1 \ge L \ge 1$). Similarly,

$$\int_{-\infty}^{-\gamma} t^2 F(t)\, dt \le 5\alpha_2\gamma^2.$$

Thus

$$\int_{|t|\ge\gamma} t^2 F(t)\, dt \le 5\frac{\alpha}{1+\alpha}\gamma^2 \le 26\frac{\alpha}{1+\alpha}\ln^2\frac{1}{\alpha}.$$

Hence the lemma follows. $\square$.

**Corollary 5.6.** *Let $K$ be a convex body in $\mathbb{R}^n$ in isotropic position. Let $Q$ be the uniform distribution on $K$, and let $P$ be any other distribution on $K$, such that $P \le (1+\alpha)Q$ for some $0 < \alpha < 1/10$. Let $a = \int_K x\, dP$. Then $\|a\| \le 5\alpha \ln\frac{1}{\alpha}$.*

**Proof.**

$$\|a\| = \int_K a\cdot x dP$$

and so the previous lemma gives us the Corollary. $\square$

We can also formulate an affine invariant consequence:

**Corollary 5.7.** *Let $K$ be a convex body in $\mathbb{R}^n$ with $b(K) = 0$. Let $Q$ be the uniform distribution on $K$, and let $P$ be any other distribution on $K$, such that $P \le (1+\alpha)Q$ for some $0 < \alpha < 1/10$. Let $a = \int_K x\, dP$. Then $a \in 5\alpha\ln\frac{1}{\alpha}K$ and $a \in -\alpha K$.*

**Proof.** It is easy to see that the statement is invariant under linear transformations. So, we may assume that $K$ is in isotropic position. Then the first assertion of the Corollary follows from the previous one. The second is easily seen using the decomposition $P = (1+\alpha)Q - \alpha R$ from the proof of lemma 5.5:

$$a = \int_K x\,dP = (1+\alpha)\int_K x\,dQ - \alpha\int_K x\,dR = -\alpha\int_K x\,dR \in -\alpha K.$$

$\square$

Based on this, we can prove the correctness of the following algorithm.

**Algorithm 5.8.** [Approximating the center of gravity] Let $K$ be a convex body in $\mathbb{R}^n$ and $0 < \varphi, \eta < 1$ be given. Compute

$$m = \left\lceil \frac{8n}{\varphi\eta} \right\rceil \quad \text{and} \quad \varepsilon = \frac{\varphi\eta}{20(n+1)^2}.$$

Draw an $\varepsilon$-good (recall Definition 2.8) sample of $m$ random points from some distribution $P$ on $K$ and compute their center of gravity $g$.

**Theorem 5.9.** (a) *Assume that $P \le (1 + (\varphi/2))Q$. Then with probability at least $1-\eta$, $g - b(K) \in -\varphi(K - b(K))$.*

(b) *Assume that $P \le (1+\alpha)Q$ where $\alpha < 1/10$ and $10\alpha\ln(1/\alpha) \le \varphi$. Also assume that $K$ is in isotropic position. Then with probability at least $1-\eta$, $\|g - b(K)\| \le \varphi$.*

**Proof.** We describe the proof of (b); the proof of (a) is similar and, in fact, simpler.

Let $c = \int_K x\,dP(x)$. By Corollary 5.6, we have $\|c\| \le 5\alpha\ln\frac{2}{\alpha} \le \varphi/2$. Thus it suffices to show that with probability at least $1-\eta$, $\|g - c\| \le \varphi/2$. Let $(z_1, \ldots, z_m)$ be the sample, and define $y_k = z_k - c$. Let us compute the expectation of $\|g - c\|^2$.

$$\mathsf{E}(\|g-c\|^2) = \frac{1}{m^2}\left(\sum_{i=1}^m \mathsf{E}(\|y_i\|^2) + \sum_{i\ne j}\mathsf{E}(y_i^{\mathsf{T}} y_j)\right). \tag{17}$$

To estimate the first sum, fix an $i$ and consider a random point $z$ in $K$ from distribution $P$ that agrees with $z_i$ with probability at least $1-\varepsilon$. Let $y = z - c$. Then we have

$$\mathsf{E}(\|y_i\|^2) = \mathsf{E}(\|y\|^2) + \mathsf{E}(\|y_i\|^2 - \|y\|^2).$$

The first term is

$$\int_K \|z-c\|^2\,dP(x) = \int_K \|z\|^2\,dP(z) - \|c\|^2 \le (1+\alpha)\int_K \|x\|^2\,dQ(x) = (1+\alpha)n.$$

The second term is at most $\varepsilon(n+1)^2 < 1$ since $\|y_i\| \leq n+1$ by Lemma 5.1. Hence $\mathsf{E}(\|y_i\|^2) \leq 2n$. Also note that $\|\mathsf{E}(y_i)\| \leq \|\mathsf{E}(y)\| + \|\mathsf{E}(y_i - y)\| = 0 + \|\mathsf{E}(z_i - z)\| \leq 2\varepsilon(n+1)$.

To estimate the second sum in (17), consider a typical term and use Lemma 2.7:

$$\begin{aligned}
\mathsf{E}(y_i^{\mathsf{T}} y_j) &= \sum_k \mathsf{E}(y_{ik}y_{jk}) \leq \sum_k \mathsf{E}(y_{ik})\mathsf{E}(y_{jk}) + 4\varepsilon(n+1)^2 \\
&= \mathsf{E}(y_i)^{\mathsf{T}}\mathsf{E}(y_j) + 4\varepsilon(n+1)^2 \leq 4\varepsilon^2(n+1)^2 + 4\varepsilon(n+1)^2 \\
&\leq 5\varepsilon(n+1)^2.
\end{aligned}$$

Hence the second sum in (17) is at most $5m(m-1)\varepsilon(n+1)^2$. Summing up, we get

$$\mathsf{E}(\|g-c\|^2) \leq \frac{2n}{m} + 5\varepsilon(n+1)^2 \leq \frac{\varphi\eta}{2}.$$

by the choice of $\varepsilon$ and $m$. Hence by Markov's inequality, the probability that $\|g-c\| > \varphi/2$ is at most $\eta$. $\square$

## 5.2. Isotropy through sampling

Let $K$ be a convex body; our goal is to describe an algorithm that brings $K$ into a $\vartheta$-nearly isotropic position with probability at least $1-\eta$. In this section we describe an algorithm that achieves this provided that we have a subroutine to generate almost uniformly distributed random points in $K$. This will be combined with a sampling algorithm in the next sections.

The idea is to generate a set of $m = O^*(n^2)$ random points in $K$, and bring those into near-isotropic position. We prove that bringing these sample points from such a distribution into isotropic position actually brings $K$ into near isotropic position.

**Algorithm 5.10.** [Isotropy transformation using sampling] Let $K \subseteq \mathbb{R}^n$ be a convex body, and let $0 < \vartheta, \eta < 1/4$.

(1) Compute

$$\varepsilon_1 = \frac{\eta^2\vartheta^2}{32(n+1)^4} \qquad m = \left\lceil \frac{80n^2}{\vartheta^2\eta^2} \right\rceil .$$

Draw an $\varepsilon_1$-good sample, from some distribution $P$, of $m$ random points $y_1, \ldots, y_m$ from $K$. Compute the vector

$$\overline{y} = \frac{1}{m}\sum_i y_i$$

and the matrix

$$Y = \frac{1}{m}\sum_{i=1}^m (y_i - \overline{y})(y_i - \overline{y})^{\mathsf{T}} .$$

If $Y^{1/2}$ is not invertible, declare the attempt a failure and repeat. Otherwise, output $K' = Y^{-1/2}(K - \overline{y})$.

**Remark.** It is easy to see that since $K$ has nonzero volume and $m \geq n$, $Y$ is invertible with probability 1.

**Theorem 5.11.** *Assume that $P$ satisfies $P \leq (1+\alpha)Q$, where $40\alpha \ln^2 \frac{1}{\alpha} = \vartheta$. Then with probability at least $1-\eta$, the body $K'$ produced by Algorithm 5.10 is in $\vartheta$-nearly isotropic position.*

**Proof.** It is easy to check that the assertion is invariant under affine transformation of $K$, so we may assume that $K$ is in isotropic position.

We start with proving the second condition of the $\vartheta$-isotropy. We want to prove that with probability at least $1-\eta$, every vector $w \in \mathbb{R}^n$ satisfies

$$(1-\vartheta)\|w\|^2 \leq \frac{1}{\mathrm{vol}(K')} \int_{K'-b(K')} (w^\mathsf{T} y)^2 \, dy \leq (1+\vartheta)\|w\|^2. \tag{18}$$

By a change of variables, inequality (18) can be written as

$$(1-\vartheta)v^\mathsf{T} Y v \leq \frac{1}{\mathrm{vol}(K)} \int_K (v^\mathsf{T} y)^2 \, dy \leq (1+\vartheta) v^\mathsf{T} Y v.$$

Here we may assume that $\|v\| = 1$, then the middle term is 1, and so we have to prove that

$$\frac{1}{1+\vartheta} \leq v^\mathsf{T} Y v \leq \frac{1}{1-\vartheta}. \tag{19}$$

We have

$$Y = Z - \overline{y}\overline{y}^\mathsf{T},$$

where

$$Z = \frac{1}{m} \sum_{i=1}^m y_i y_i^\mathsf{T}.$$

Hence it suffices to show that with probability at least $1-\eta$, for every $v \in \mathbb{R}^n$, $\|v\| = 1$ we have

$$\frac{1}{1+\vartheta} + (v^\mathsf{T}\overline{y})^2 \leq v^\mathsf{T} Z v \leq \frac{1}{1-\vartheta} + (v^\mathsf{T}\overline{y})^2. \tag{20}$$

To prove this, we need some auxiliary inequalities of similar nature. Let

$$A = \int_K x x^\mathsf{T} \, dP(x).$$

By Lemma 5.5, for every unit vector $v$,

$$1 - 26\alpha \ln^2 \frac{2}{\alpha} \le v^{\mathsf{T}} A v < 1 + \alpha. \tag{21}$$

Now we prove that for all unit vectors $v$,

$$|v^{\mathsf{T}}(\mathsf{E}Z - A)v| \le \frac{\vartheta}{4}. \tag{22}$$

Indeed, let $P_i$ be the distribution of $y_i$. Then

$$v^T \mathsf{E} Z v = \frac{1}{m}\sum_{i=1}^{m} v^T \mathsf{E}(y_i y_i^{\mathsf{T}}) v = \frac{1}{m}\sum_{i=1}^{m} \int_K (v^{\mathsf{T}} x)^2 \, dP_i(x) = \int_K (v^{\mathsf{T}} x)^2 \, dP'(x),$$

where $P' = (1/m)\sum_i P_i$. Hence

$$|v^{\mathsf{T}}(\mathsf{E}Z - A)v| = \left| \int_K (v^{\mathsf{T}} x)^2 \left(dP'(x) - dP(x)\right) \right| \le \varepsilon_1 (n+1)^2,$$

since $|P - P'|_{tv} \le \varepsilon_1$ and $(v^{\mathsf{T}} x)^2 \le (n+1)^2$. This proves (22).

Next we prove that with probability at least $1 - \eta/2$, we have for all unit vectors $v$

$$|v^{\mathsf{T}}(Z - \mathsf{E}Z)v| \le \frac{\vartheta}{4}. \tag{23}$$

To prove (23), we use that

$$\|Z - \mathsf{E}Z\|^2 \le \mathrm{Tr}((Z - \mathsf{E}Z)^2)$$

Let us compute the expectation of this trace. We have

$$\begin{aligned} m^2 (Z - \mathsf{E}Z)^2 &= \left( \sum_{i=1}^{m} (y_i y_i^{\mathsf{T}} - \mathsf{E}(y_i y_i^{\mathsf{T}})) \right)^2 \qquad (24) \\ &= \sum_{i=1}^{m} \Big( y_i y_i^{\mathsf{T}} - \mathsf{E}(y_i y_i^{\mathsf{T}}) \Big)^2 + \sum_{i \ne j} \Big( y_i y_i^{\mathsf{T}} - \mathsf{E}(y_i y_i^{\mathsf{T}}) \Big) \Big( y_j y_j^{\mathsf{T}} - \mathsf{E}(y_j y_j^{\mathsf{T}}) \Big) \end{aligned}$$

The first term is handled as follows : fix any $i$, then

$$\mathsf{E}\Big( y_i y_i^{\mathsf{T}} - \mathsf{E}(y_i y_i^{\mathsf{T}}) \Big)^2 = \mathsf{E}((y_i y_i^{\mathsf{T}})^2) - (\mathsf{E}(y_i y_i^{\mathsf{T}}))^2$$

and hence

$$\mathsf{E}\mathrm{Tr}\Big( \Big( y_i y_i^{\mathsf{T}} - \mathsf{E}(y_i y_i^{\mathsf{T}}) \Big)^2 \Big) \le \mathsf{E}\mathrm{Tr}\Big( (y_i y_i^{\mathsf{T}})^2 \Big) = \mathsf{E}(\|y_i\|^4).$$

Since $|P_i - P|_{tv} < \varepsilon_1$, we can construct a random variable $z_i$ distributed according to $P$ such that $z_i = y_i$ with probability $1 - \varepsilon_1$. Then,

$$\mathsf{E}(\|y_i\|^4) = \mathsf{E}(\|z_i\|^4) + \mathsf{E}(\|y_i\|^4 - \|z_i\|^4).$$

The first term is bounded above by $(1+\alpha)\int_K \|x\|^4 dx/\mathrm{vol}(K)$ which is at most $8(1+\alpha)(\int_K \|x\|^2/\mathrm{vol}(K))^2 = 8(1+\alpha)n^2$, by the assumption that $P \le (1+\alpha)Q$ and the fact that $\mathsf{E}_Q(\|x\|^4) \le 8(\mathsf{E}_Q(\|x\|^2))^2$ by a Theorem of Gromov and Milman (1984). The second term is at most $\varepsilon_1(n+1)^4$. Hence

$$\sum_{i=1}^m \mathsf{E}\mathrm{Tr}\Big(\Big(y_i y_i^{\mathsf{T}} - \mathsf{E}(y_i y_i^{\mathsf{T}})\Big)^2\Big) \le 8(1+\alpha)mn^2 + m\varepsilon_1(n+1)^4.$$

The second term in (24) can be estimated using the "almost independence" of the sample points. We can write its trace as

$$\begin{aligned}&\sum_{i\ne j} \mathrm{Tr}\Big(y_i y_i^{\mathsf{T}} - \mathsf{E}(y_i y_i^{\mathsf{T}})\Big)\Big(y_j y_j^{\mathsf{T}} - \mathsf{E}(y_j y_j^{\mathsf{T}})\Big)\\ &= \sum_{i\ne j}\sum_k\sum_r \Big(y_{ik}y_{ir} - \mathsf{E}(y_{ik}y_{ir})\Big)\Big(y_{jk}y_{jr} - \mathsf{E}(y_{jk}y_{jr})\Big).\end{aligned}$$

Invoking Lemma 2.7, we get that the expectation of each term here is bounded by $4\varepsilon_1(n+1)^4$, and hence the expectation of the whole sum is bounded by $4\varepsilon_1 m(m-1)(n+1)^4$.

Summing up, we get that

$$\mathsf{E}(\|Z-\mathsf{E}Z\|)^2 \le \mathsf{E}(\|Z-\mathsf{E}Z\|^2) \le 8(1+\alpha)\frac{1}{m}n^2 + \frac{1}{m}\varepsilon_1(n+1)^4 + 4\varepsilon_1(n+1)^4 < \frac{\eta^2\vartheta^2}{4},$$

and hence with probability at least $1-\eta$, $\|Z-\mathsf{E}Z\| \le \vartheta/4$. This proves (23).

Finally we remark that by Theorem 5.9, with probability at least $1-\eta/2$,

$$\|\overline{y}\| \le \frac{\vartheta}{4}, \tag{25}$$

and hence $(v^{\mathsf{T}}y)^2 < \vartheta^2/16$.

Combining these inequalities, we have

$$\begin{aligned} v^{\mathsf{T}}Zv &= v^{\mathsf{T}}Av + v^{\mathsf{T}}(\mathsf{E}Z-A)v + v^{\mathsf{T}}(Z-\mathsf{E}Z)v \le \left(1+\frac{\vartheta}{4}\right) + \frac{\vartheta}{4} + \frac{\vartheta}{4}\\ &< \frac{1}{1-\vartheta} + (v^{\mathsf{T}}y)^2,\end{aligned}$$

and similarly,

$$\begin{aligned} v^{\mathsf{T}}Zv &= v^{\mathsf{T}}Av + v^{\mathsf{T}}(\mathsf{E}Z-A)v + v^{\mathsf{T}}(Z-\mathsf{E}Z)v \ge \left(1-\frac{\vartheta}{4}\right) - \frac{\vartheta}{4} - \frac{\vartheta}{4}\\ &> \frac{1}{1+\vartheta} + (v^{\mathsf{T}}y)^2.\end{aligned}$$

This proves (19).

To complete the proof, it suffices to remark that $b(K') = -Y^{1/2}\overline{y}$ and hence if (19) and (25) hold, then

$$\|b(K')\| = \sqrt{\overline{y}^{\mathsf{T}} Y \overline{y}} \le \sqrt{\frac{\vartheta^2}{16}\frac{1}{1-\vartheta}} < \vartheta. \qquad \square$$

### 5.3. Improving local conductance

In this section we describe an algorithm to find an affine transformation that increases the local conductance while keeping the second moment of the body bounded.

First we describe a simple linear transformation that brings a convex body "closer" to its isotropic position. Throughout this section, let $K$ be a convex body in $\mathbb{R}^n$ for which

(i) the center of gravity $b(K) \in (-\frac{1}{10})K$ and

(ii) $E_K(\|x\|^2) \le n$.

**Lemma 5.12.** *Let $h$ be a unit vector for which*

$$h^{\mathsf{T}} x < \frac{1}{2}$$

*for every $x \in K$. Define the linear transformation*

$$U_h = \left(1 - \frac{1}{2n}\right)(I + hh^{\mathsf{T}}).$$

*Then*

(a) *if $K$ satisfies (i) and (ii), then so does $U_h K$;*

(b) *the volume of $U_h K$ is at least $9/8$ times the volume of $K$.*

The geometric meaning of $U_h$ is the following: if the tangent plane is too near to the origin, then we stretch $K$ by a factor of about 2 in the direction of the normal vector of the tangent plane, and then shrink in all directions by a factor of $\left(1 - \frac{1}{2n}\right)$.

**Proof.** The center of gravity is affine invariant, so (i) is trivially preserved. Lemma 5.4 implies that

$$E_K((h^{\mathsf{T}} x)^2) < \frac{1}{3},$$

and hence

$$\begin{aligned} E_{U_h K}(\|x\|^2) &= E_K(\|U_h x\|^2) = \left(1 - \frac{1}{2n}\right)^2 E_K(\|x + (x^{\mathsf{T}} h)h\|^2) \\ &\le \left(1 - \frac{1}{2n}\right)^2 E_K(\|x\|^2 + 3(x^{\mathsf{T}} h)^2) = \left(1 - \frac{1}{2n}\right)^2 (n+1) < n. \end{aligned}$$

This concludes the proof of (a).

Since

$$\det U_h = 2\left(1 - \frac{1}{2n}\right)^n \geq \frac{9}{8},$$

we have (b). $\square$

From now on, we assume that, in addition to (i) and (ii) above,

(iii) $K$ contains a ball with radius $1/\sqrt{n}$ centered at the origin.

**Definition 5.13.** [Flat steps] Let $v \in K$ but $u \notin K$. Find, using binary search, a point $u'$ on the segment $[u, v]$ such that $u' \notin K$ but $u' \in 2^{1/n}K$. The separation oracle called for $u'$ returns a separating hyperplane $H$. If $H$ is closer to the origin than $1/2$, we call the pair $(u, v)$ a *flat step.*

**Remark.** We shall apply the above definitions to random walks where $\|v - u\| < \delta < \frac{1}{\sqrt{n}}$. If (iii) holds, then $2\log n$ oracle calls will find $u'$.

**Lemma 5.14.** *Let $K$ be a convex body containing the origin and let $v$ be a uniformly distributed random point in $K$. Make one step of a lazy random walk starting from $v$. Then the probability that this step is a non-flat improper step is at most $4\delta\sqrt{n}$.*

**Proof.** Put $K_1 = \operatorname{conv}(K \cup \frac{1}{2}B)$. Assume that the attempted step $[v, u]$ is non-flat improper. Then trivially $u \notin K_1$. We prove that the $(2n)$-dimensional measure of the set $S$ of pairs $[v, u]$ with $v \in K$, $u \in \mathbb{R}^n \setminus K_1$ and $\|u - v\| \leq \delta$ is at most $4\delta\sqrt{n}\mathrm{vol}(K)\mathrm{vol}(B')$; this will prove the lemma.

Let $q'$ be the point of intersection of the segment $[v, u]$ and $\partial K_1$. Then clearly $q'$ belongs to

$$F' = \partial K_1 \cap (2^{1/n}K).$$

Applying Lemma 4.3 to $K_1$, we get that

$$\mathrm{vol}_{2n}(S) \leq \delta \mathrm{vol}_{n-1}(F') \frac{\pi_{n-1}}{(n+1)\pi_n} \mathrm{vol}(B') < \frac{\delta}{\sqrt{n}} \mathrm{vol}_{n-1}(F')\mathrm{vol}(B').$$

The hyperplane supporting $K_1$ at any point of $F'$ has distance at least $1/2$ from the origin. Hence the union $U$ of segments connecting 0 to $F'$ has volume at least $\mathrm{vol}_{n-1}(F')/(2n)$. On the other hand, clearly $U \subseteq 2^{1/n}K$. This implies that

$$\mathrm{vol}_{n-1}(F') \leq 4n\mathrm{vol}(K),$$

and so

$$\mathrm{vol}_{2n}(S) < 4\delta\sqrt{n}\mathrm{vol}(K)\mathrm{vol}(\delta B). \qquad \square$$

**Lemma 5.15.** *Let $K \subseteq dB$ ($d \geq 1$) be a convex body with average local conductance $\lambda$ with respect to $\delta$-moves where $0 < \delta < d/32$ . Let $u \in K$. Starting from $u$, do a lazy random walk in $K$ with step size $\delta$ until at least*

$$T = \left\lceil 1600n^2 \left(\frac{d}{\delta}\right)^2 \ln\frac{d}{\delta}\right\rceil$$

*proper steps were made. Then the probability that no flat steps were attempted is at most $\lambda + 6\delta\sqrt{n}$.*

**Proof.** We may assume that $\delta < 1/(6\sqrt{n})$. Consider a random walk in the body $K_1 = \mathrm{conv}(K \cup \frac{1}{2}B)$, starting at $u$. Until this walk hits $K_1 \setminus K$, it can be considered a random walk in $K$. Conversely, a random walk in $K$ can be considered a random walk in $K_1$ until the first flat step is attempted because until then, any time we attempt to step out of $K$, we are actually stepping out of $K_1$. Hence the probability that a random walk of length $T$ in $K$ attempts a flat step is at least as large as the probability that a random walk in $K_1$ of length $T$ hits $K_1 \setminus K$.

Now $(1/2)B \subseteq K_1 \subseteq dB$. Corollary 4.6 implies that the average local conductance of $K_1$ is at least $1 - \delta\sqrt{n}$. Let $\hat{Q}_1$ be the speedy distribution on $K_1$, then Corollary 4.2 implies that for the distribution $Q_j$ of the point $w_j$ at the $(T-1)$th proper step we have $M(Q_j, \hat{Q}_1) \leq (\delta/d)^n \leq \delta\sqrt{n}$ and therefore, if $w$ is the last point before the $T$th proper step, then

$$P(w \in A) - \frac{\mathrm{vol}(A)}{\mathrm{vol}(K_1)} \leq \hat{Q}_1(A) + \delta\sqrt{n} - \hat{Q}_1(A)(1 - \delta\sqrt{n}) \leq 2\delta\sqrt{n}$$

for every $A \subseteq K_1$ and so the variational distance between the distribution of $w$ and the uniform distribution on $K_1$ is at most $2\delta\sqrt{n}$. For convenience, consider a uniformly distributed random point $v$ of $K$ that agrees with $w$ with probability $1 - 2\delta\sqrt{n}$, and make a step from $v$. If the walk in $K$ attempted no flat step during the first $T$ steps, then either $v \neq w$, or $v \in K$ and the step made from $v$ was not flat. By Lemma 5.14, this latter probability is at most $4\delta\sqrt{n} + \lambda$; thus the probability that no flat step was attempted is at most $\lambda + 6\delta\sqrt{n}$. $\square$

Next we describe the key algorithm to improve the local conductance.

**Algorithm 5.16.** Let $K$ be a convex body in $\mathbb{R}^n$, and let $0 < \vartheta, \eta < 1$ be given.

(0) Let

$$\delta = \frac{\min\{\vartheta, \eta\}}{24\sqrt{n}}, \qquad d = \sqrt{\frac{2n}{\vartheta}},$$

$$M = \left\lceil \frac{32}{\vartheta} n \log n \right\rceil, \qquad T = \left\lceil 1600 n^2 \left(\frac{d}{\delta}\right)^2 \ln \frac{d}{\delta} \right\rceil.$$

Select a random integer $N$ uniformly from $\{0, \ldots, M-1\}$.

(1) Generate a point $u$ in $K$ whose distribution is closer than $\eta/6$ to the uniform (in total variation distance).

(2) Let $K_0 = K$. For $i = 0, 1, \ldots, N-1$, do the following. Starting from $u$, do a lazy random walk in $K_i' = K_i \cap dB$ until either $T$ proper steps were made, or a flat step was made, whichever comes first.

If we end with $T$ proper steps, we go to the next $i$. If we end with a flat step (with respect to $K_i$), then let $H$ be the separating hyperplane whose distance from the origin is at most 1/2. Let $h$ be the normal of $H$ of unit length, directed away from the origin. Apply the linear transformation $U_h$ of Lemma 5.12 to $K_i$ to get $K_{i+1} = U_h K_i$.

(3) Output the body $K_N$.

**Remarks.** 1. The random choice of $N$ is certainly an artifact of the proof below. $M = N$ should do.

2. The random choice of $u$ may also be unnecessary; any point sufficiently far away from the corners (in particular, the origin) should be just as good.

Note that we use uniform distribution for $u$ instead of the speedy distribution $\hat{Q}$. One reason of this is that the body $K$ keeps changing during the algorithm and the uniform distribution is invariant while $\hat{Q}$ is not under linear transformations.

3. Intersecting $K_i$ with $dB$ is probably another artifact. We could walk inside $K_i$ if we could prove a version of the isoperimetric inequality (Theorem 3.1), with the diameter $d$ replaced by the square root of the second moment. Theorem 5.1 of Kannan, Lovász and Simonovits (1995) is of this nature; however, it would not be directly applicable here.

4. We assume that we have a subroutine to generate the "almost uniform" point $u$. How this is actually done will be discussed when the main algorithm is used.

**Theorem 5.17.** *Assume that $K$ satisfies (i), (ii) and (iii). Then algorithm 5.16 produces a convex body $K_N$ satisfying (i) and (ii). The expectation of the average local conductance of* *the output body is at least $1-\vartheta$. With probability at least $1-\eta$, the number of oracle calls it uses is at most*

$$3MT = O\left(\frac{n^5}{\eta^2\vartheta^3}\ln n \ln \frac{n}{\eta\vartheta}\right) = O^*(n^5).$$

**Proof.** The first assertion is clear by Lemma 5.12. This also implies that the volume of $K$ never exceeds the volume of the isotropic ball, which is $(n+2)^{n/2}\pi_n$. By (iii), the volume of the original body is at least $n^{-n/2}\pi_n$. Thus $\mathrm{vol}(K)$ can increase by at most a factor $(n(n+2))^{n/2} < (n+1)^n$. Since the volume of $K$ increases by a factor of at least $9/8$ (by Lemma 5.12) after the linear transformation is carried out at the end of such a walk, at most $8n\log n$ flat step transformations can occur.

Consider the algorithm going on for $M$, rather than $N$, iterations. Let $L_i$ be the average local conductance of $K_i$ and $L_i'$, the average local conductance of $K_i'$ (these are random variables!). Since $E_{K_i}(\|x\|^2) \le n$ is preserved, and $K_i' \subseteq dB$, we have (by Markov inequality),

$$\frac{\mathrm{vol}(K_i')}{\mathrm{vol}(K_i)} \ge 1 - \frac{\vartheta}{2}.$$

So,

$$L_i \ge L_i' \frac{\mathrm{vol}(K_i')}{\mathrm{vol}(K_i)} \ge L_i'\left(1-\frac{\vartheta}{2}\right).$$

Let $\lambda_i$ be the expectation of $L_i$.

Let $X_i$ be the indicator variable of the event that the $i$-th random walk ended with a flat step. Then $\sum_i X_i$ is the number of such walks, and hence,

$$\sum_i X_i \le 8n\log n. \tag{26}$$

On the other hand, from Lemma 5.15 we get that

$$\mathsf{P}(X_{i+1}=1 \mid \text{previous events}) \ge 1 - 6\delta\sqrt{n} - L_i' \ge 1 - \frac{\vartheta}{4} - \frac{1}{1-\vartheta/2}L_i.$$

Therefore

$$\sum_{i=0}^{M-1} E(X_i) \ge \sum_{i=0}^{M-1}\left(1-\frac{\vartheta}{4} - \frac{1}{1-\vartheta/2}\lambda_i\right)$$

and hence by (26),

$$\frac{1}{M}\sum_{i=0}^{M-1}\lambda_i \ge \left(1-\frac{\vartheta}{2}\right)\left(1-\frac{8n\log n}{M} - \frac{\vartheta}{4}\right) \ge 1-\vartheta.$$

Recall that $N$ is a random element of $\{0,\dots,M-1\}$. Thus

$$E(L_N) = E(\lambda_N) \ge 1-\vartheta.$$

To estimate the number of oracle calls, we count different kinds of steps. The number of proper steps is at most $MT$; the number of flat improper steps is at most $8n \log n \leq MT$. We show that the probability that the number of non-flat improper steps is larger than $MT$ is less than $\eta$.

For simplicity, imagine that the last walk goes on, if necessary, until a total of at least $3MT$ steps are made. If the number of non-flat improper steps during the algorithm is larger than $MT$ then their number among the first $3MT$ steps is larger than $MT$. Since $u$, and therefore every given point in the sequence, has a distribution that is closer to uniform than $\eta/6$ (in total variation distance), the probability that a given step is non-flat improper is at most $\eta/6 + 4\delta\sqrt{n} < \eta/3$ by Lemma 5.14. Thus the expected number of non-flat improper steps is at most $\eta MT$. By Markov's inequality, the probability that there are more than $MT$ such steps is at most $\eta$. □

### 5.4 Isotropy for rounded bodies

Assume that we are given a convex body $K$ and a $\vartheta, \eta > 0$ (later $\vartheta$ and $\eta$ will be fixed constants); we wish to bring $K$ into $\vartheta$-nearly isotropic position, with probability at least $1 - \eta$. In this section we treat the special case when $K$ already satisfies the conditions

$$B \subseteq K \subseteq 10nB$$

(the general case will be treated in the next section). The basic tool is Algorithm 5.10, but to implement it, we have to describe how to generate a sample of size $80n^2/(\vartheta\eta)$. This can be done using $O^*(n^5)$ oracle calls, by combining our previous results.

**Step 1**. Let $Q$ be the uniform distribution on $K$. We apply Algorithm 4.15 with Algorithm 4.11 as $\mathcal{A}$ to get a single random point in $K$ from a distribution $Q^*$ satisfying $|Q^* - \hat{Q}|_{tv} < \eta/30$.

By Theorems 4.16 and 4.14 this takes an expected

$$O\left(n^5 \left(\ln \frac{n}{\eta}\right)^2 \ln n\right)$$

oracle calls.

**Step 2.** Next, we can scale down $K$ so as to satisfy

$$\frac{1}{\sqrt{n}} B \subseteq K \subseteq 10\sqrt{n} B.$$

and then apply Algorithm 5.16. (As the result of Step 1, we have an $\eta/30$-good sample point to start from.) We get a convex body satisfying (i) and (ii) and having average local conductance at least .999. By Theorem 5.17, with probability at least $1-\eta/5$, we use

$$O\left(\frac{n^5}{\vartheta^3\eta^2}\ln\frac{1}{\vartheta\eta}\ln^2 n\right)=O^*(n^5).$$

oracle calls.

**Step 3**. This allows us to use Algorithm 4.7 with $\delta$ as in Step 2 repeatedly (starting with a point generated by Step 1) to generate an $\varepsilon_1$-good sample of $m=\lceil 80n^2/(\vartheta\eta)^2\rceil$ points where $\varepsilon_1=\eta^2\vartheta^2/(32(n+1)^4)$. This takes

$$O\left(\frac{n^5}{\vartheta^6}\cdot\ln\frac{1}{\vartheta\eta}\ln n\right)=O^*(n^5)$$

calls to the oracle.

**Step 4.** We use this $\varepsilon_1$-good sample of $m$ points in Algorithm 5.10 to bring $K$ into near isotropic position. This does not use the oracle; it uses only matrix arithmetic; finding the square root of a matrix $Y$.

To summarize, in this section, we have shown using the above plus Remark 4.9 :

**Lemma 5.18** *Let $1>\vartheta,\eta>0$. If $K$ is a convex body satisfying $B\subseteq K\subseteq 10nB$, then we have a randomized algorithm which brings K into $\vartheta$-nearly isotropic position with probability at least $1-\eta$. The algorithm never makes more than*

$$O\left(\frac{n^5}{\eta^2\vartheta^6}(\ln n)^3(\ln(1/\eta\vartheta))^2\right)$$

*oracle calls.*

### 5.5. Isotropy for general bodies

Now consider a general convex body, satisfying only

$$B\subseteq K\subseteq dB$$

for some $d$. As remarked in the introduction, we may achieve $d=n^{3/2}$ by a standard application of the ellipsoid algorithm; but the order of magnitude of the time complexity of our algorithm remains the same for any $d<n^{\text{const}}$.

**Algorithm 5.19.** [Isotropy transformation for general convex body] Let a convex body $B \subseteq K \subseteq dB$ and $0 < \vartheta, \eta < 1$ be given. Compute $p = \lceil \log d \rceil$. Let $K_0 = K$. For $i = 0, 1, \ldots, p$, do the following: Let $K_i' = K_i \cap 10nB$. Find an affine transformation $\alpha_i$ that maps $K_i'$ to a $\vartheta$-nearly isotropic position with probability at least $1 - \eta/(\log d)$. Let $K_{i+1}$ be the image of $K_i$ under this map. Output $K_p$.

**Theorem 5.20.** *The convex body produced by Algorithm 5.19 is in $\vartheta$-isotropic position with probability at least $1 - \eta$.*

**Remark.** Algorithm 5.19 could use any algorithm $\mathcal{A}$ bringing a sandwiched $K$ into $\vartheta$-near isotropic position. If $\mathcal{A}$ uses $T$ steps (when called with error-bound $\eta/\ln d$, then Algorithm 5.19 uses at most $(\log d)T$ steps.

**Proof of Theorem 5.20.** Assume that $K \subseteq dB$ and define for $i = 1, \ldots, p$

$$d_i = \max\left\{\frac{d}{2^i}, 10n\right\}.$$

It suffices to prove (by induction on $i$) that if all the iterations were successful (which happens with probability at least $1 - \eta$), then

$$K_i \subseteq d_i B. \tag{27}$$

So (by $d_{p-1} = 10n$) $K_{p-1} \subseteq 10nB$ and therefore $K_p$ is already in $\vartheta$-nearly isotropic position.

The case $i = 0$ is trivial. Let $i > 0$. Let $v \in K_i$, and let $v$ be the image of $u \in K_{i-1}$ under $\alpha_{i-1}$. If $u \in K_{i-1}'$ then $v$ lies in a $\vartheta$-isotropic body, and hence by Corollary 5.2 $\|v\| \leq (1 + 2\vartheta)(n + 1) < 2n$. So suppose that $u \in K_{i-1} \setminus K_{i-1}'$; let $q$ be the point where the segment $[0, u]$ intersects the boundary of $K_{i-1}'$. Let $z = \alpha_{i-1}(0)$ and $s = \alpha_{i-1}(q)$. Since $0, q \in K_{i-1}'$, we have $\|z\| < 2n$ and $\|s\| < 2n$.

Now $u = \tau q + (1 - \tau)0$, where $\tau = \|u\|/(10n) > 1$. Since $\alpha$ is affine, it follows that $v = \tau s + (1 - \tau)z$ and hence

$$\|v\| \leq \tau\|s\| + |1 - \tau|\|z\| < 4n\tau < \frac{1}{2}\|u\| \leq \frac{d}{2^i}.$$

This proves (27). The bound on the expected number of oracle calls follows from Lemma 5.18. $\square$

## 6. Estimating the volume

Now we describe how the previous sampling and "rounding" algorithms can be used to estimate the volume of a convex set $K$. In Section 5, we have seen that $K$ can be brought into near isotropic position, using $O^*(n^5)$ oracle calls. After this, we intersect $K$ with a ball with radius $2\sqrt{2n}\log(1/\varepsilon)$; clearly, we loose at most a fraction of $\varepsilon$ of its volume.

To conclude, it suffices to describe an algorithm that computes (approximately) the volume of a convex body $K$ satisfying $B \subseteq K \subseteq O^*(\sqrt{n})B$, in time $O^*(n^5)$. More generally, we show how to compute the volume of a convex body $K$ satisfying $B \subseteq K \subseteq dB$ in time $O^*(n^4 d^2)$.

Define, as in Section 4, $K_i = K \cap 2^{i/n} B$ $(i = 0, \dots, m = \lceil n \log n \rceil)$. We will apply Algorithm 4.11. $\delta$ is as defined there. $K_0 = B$ and $K_m = K$. Moreover,

$$1 \geq \frac{\mathrm{vol}(K_{i-1})}{\mathrm{vol}(K_i)} \geq \frac{1}{2}.$$

We denote by $\ell_i(x)$ the local conductance of $K_i$ at point $x$; so $\ell(x) = \ell_m(x)$. We define $\ell_i(x) = 0$ for $x \notin K_i$. We also define $\bar{\ell}_i = \int_{K_i} \ell_i(x)\,dx$. Clearly $\ell_i(x) \leq \ell_{i+1}(x)$. It follows from Corollary 4.6 that

$$.95\mathrm{vol}(K_i) \leq \bar{\ell}_i \leq \mathrm{vol}(K_i).$$

Now we describe the volume algorithm. Roughly speaking, we follow the method of Dyer, Frieze and Kannan (1989): we generate $p = 400\varepsilon^{-2} n \log n = O^*(n)$ random points in each $K_i$ and count how many of them fall in $K_{i-1}$; this gives an estimate of the ratio $\mathrm{vol}(K_{i-1})/\mathrm{vol}(K_i)$. However, our methods are more efficient in generating random points from the distribution $\hat{Q}_i$, and so we approximate the ratios $\bar{\ell}_{i-1}/\bar{\ell}_i$ instead. The ratio $\bar{\ell}_m/\mathrm{vol}(K)$ is determined by the same sort of method, but this time we use uniformly distributed points in $K$. The value $\bar{\ell}_0$ is computed easily.

**Algorithm 6.1.** Let $K$ be a convex body, and let $0 < \varepsilon < 1$ be given. Compute

$$p = \left\lceil \frac{400m}{\varepsilon^2} \right\rceil, \qquad \varepsilon_0 = \frac{\varepsilon^2}{12000m^2}.$$

Execute $p$ independent runs of Algorithm 4.11, to produce $mp$ points $w_{ir}$, $1 \leq i \leq m$, $1 \leq r \leq p$. For each $i$, the points $w_{i1}, \dots w_{ip}$ are totally independent points in $K_i$, from a distribution $P_i$ such that $|P_i - \hat{Q}_i|_{tv} < \varepsilon_0$, and $w_{ir}$ and $w_{js}$ are $\varepsilon_0$-independent for all $i, j, r$ and $s$. For each $i$ and $r$, make a step of the speedy walk from $w_{ir}$ , to get a point $w'_{ir} \in K_i$. Set

$$a_{ir} = \begin{cases} 1, & \text{if } w_{ir} \in K_{i-1} \text{ and } w'_{ir} \in K_{i-1}, \\ 0, & \text{otherwise.} \end{cases}$$

Compute $b_i = a_{i1} + \ldots + a_{ip}$.

Execute $p$ further independent runs of Algorithm 4.11 (with the same parameters), followed by Algorithm 4.15, to get $p$ independent points $w_1, \ldots, w_p$ in $K$ from a distribution $P$ such that $|P - Q|_{tv} < 5\varepsilon_0$. For each $1 \le r \le p$, generate a uniformly distributed random point $w'_r$ in the ball $w_r + B'$. Set

$$a_{m+1,r} = \begin{cases} 1, & \text{if } w'_r \in K, \\ 0, & \text{otherwise.} \end{cases}$$

Compute $b_{m+1} = a_{m+1,1} + \ldots a_{m+1,p}$ and return

$$\zeta = \frac{p^{m+1}\bar{\ell}_0}{b_1 \ldots b_{m+1}}$$

as the estimate of the volume of $K$.

**Theorem 6.2.** *Assume that $K$ satisfies $B \subseteq K \subseteq dB$. Then the probability that the value $\zeta$ returned by the algorithm is between $(1-\varepsilon)\mathrm{vol}(K)$ and $(1+\varepsilon)\mathrm{vol}(K)$ is at least $3/4$. With probability at least $9/10$, the total number of oracle calls is*

$$O\left(\frac{n^4 d^2}{\varepsilon^2} \ln n \ln d \Big(\ln \frac{n}{\varepsilon}\Big)^2\right).$$

**Proof.** Let $\rho_i = \bar{\ell}_{i-1}/\bar{\ell}_i$ $(1 \le i \le m)$, and $\rho_{m+1} = \bar{\ell}_m/\mathrm{vol}(K)$. It is easy to check that $2/5 \le \rho_i \le 1$. The key fact we will use is that for each $1 \le i \le m$,

$$\begin{aligned} \mathsf{E}(a_{ir}) &= \mathsf{P}(w_{ir} \in K_{i-1}, w'_{ir} \in K_{i-1}) \\ &= \int_{K_{i-1}} \mathsf{P}(w'_{ir} \in K_{i-1} \mid w_{ir} = x)\, dP_i(x) \\ &= \int_{K_{i-1}} \frac{\ell_{i-1}(x)}{\ell_i(x)}\, dP_i(x) \\ &= \int_{K_{i-1}} \frac{\ell_{i-1}(x)}{\ell_i(x)}\, d\hat{Q}_i(x) + \varepsilon_i, \end{aligned}$$

where $|\varepsilon_i| \le \|P_i - \hat{Q}_i\|_1 < \varepsilon_0$. Here

$$\int_{K_{i-1}} \frac{\ell_{i-1}(x)}{\ell_i(x)}\, d\hat{Q}_i(x) = \frac{\int_{K_{i-1}} \ell_{i-1}(x)\, dx}{\int_{K_i} \ell_i(x)\, dx} = \rho_i.$$

Hence

$$|\mathsf{E}(a_{ir}) - \rho_i| < \varepsilon_0. \tag{28}$$

for $1 \le i \le m$. A similar argument shows that this inequality also holds for $i = m+1$. It follows that $\mathsf{E}(a_{ir}) > 1/3$ for all $i$.

Informally, we see that $b_i \approx p\bar{\ell}_{i-1}/\bar{\ell}_i = p\rho_i$ for $1 \le i \le m$, and similarly, $b_{m+1} \approx p\bar{\ell}_m/\mathrm{vol}(K) = p\rho_{m+1}$, whence $\zeta \approx \mathrm{vol}(K)$. To analyze the error of the product of the $b_i$, let $\beta_i = \mathsf{E}(b_i) = p\mathsf{E}(a_{ir})$, and consider the random variable

$$X = \sum_{i=1}^{m+1} \left( \ln \frac{b_i}{\beta_i} \right)$$

We claim that with probability at least .75,

$$|X| \le \frac{\varepsilon}{2}. \tag{29}$$

First we remark that $b_i = \sum_r a_{ir}$, where $0 \le a_{ir} \le 1$ and, for fixed $i$, the $a_{ir}$ are independent. Hence we may apply the inequality of Chernoff–Hoeffding, and get that with probability at least .99,

$$b_i \ge \frac{\beta_i}{\sqrt{2}} \qquad \text{for all } i. \tag{30}$$

Set

$$A = \sum_i \frac{b_i - \beta_i}{\beta_i}, \qquad C = \sum_i \left( \frac{b_i - \beta_i}{\beta_i} \right)^2,$$

and

$$D = \sum_{i<j} \frac{(b_i - \beta_i)(b_j - \beta_j)}{\beta_i \beta_j}$$

Using the formula for the variance of the binomial distribution, we get

$$\mathsf{E}(C) = \sum_i \frac{1}{\beta_i} \left( 1 - \frac{\beta_i}{p} \right) \le \frac{2m}{p} < \frac{\varepsilon^2}{200}.$$

Next we estimate the expectation of $D$. The expectation of a typical summand can be estimated using $\varepsilon_0$-independence and Lemma 2.7:

$$\frac{1}{\beta_i \beta_j} \mathsf{E} \left[ \sum_{r,s} (a_{ir} - \mathsf{E}(a_{ir}))(a_{js} - \mathsf{E}(a_{js})) \right] \le \frac{4}{\beta_i \beta_j} p^2 \varepsilon_0 < 36\varepsilon_0,$$

and hence

$$\mathsf{E}(D) < 18m^2 \varepsilon_0 < \frac{\varepsilon^2}{640}.$$

We claim that whenever (30) holds, $C < \varepsilon^2/30$ and $D < \varepsilon^2/64$, then we have $|X| < \varepsilon/2$. Since $A^2 = C + 2D$, we get that in this case $|A| < \varepsilon/\sqrt{15}$. If $X \ge 0$, then using the inequality $\ln x \le x - 1$, we get that $X \le A < \varepsilon/2$. If $X < 0$, then (using (30)) we can apply the inequality $\ln x \ge x - 1 - (x-1)^2$ (for $x \ge 1/\sqrt{2}$), to get $X \ge A - C \ge -\varepsilon/2$. In both cases, the claim follows. By Markov's inequality, the probability that either $C > \varepsilon^2/30$ or $D > \varepsilon^2/64$ is at most $.15 + .1 = .25$. Thus with probability at least 3/4, we have $|X| \le \varepsilon/2$.

To prove the first assertion of the theorem, we use that

$$\zeta = \frac{p^{m+1}}{b_1 \dots b_{m+1}} \bar{\ell_0}$$

and

$$\mathrm{vol}(K) = \frac{\bar{\ell_0}}{\rho_1 \dots \rho_{m+1}}$$

whence

$$\frac{\mathrm{vol}(K)}{\zeta} = \frac{b_1 \dots b_{m+1}}{\beta_1 \dots \beta_{m+1}} \frac{\beta_1 \dots \beta_{m+1}}{(p\rho_1) \dots (p\rho_{m+1})}.$$

Inequality (28) implies that

$$\left| \frac{\beta_i}{p\rho_i} - 1 \right| < \frac{\varepsilon_0}{\rho_i} \le \frac{5}{2}\varepsilon_0,$$

and hence

$$\left| \ln \prod_{i=1}^{m+1} \frac{\beta_i}{p\rho_i} \right| \le \sum_{i=1}^{m+1} \left| \ln \frac{\beta_i}{p\rho_i} \right| \le 2 \sum_{i=1}^{m+1} \left| \frac{\beta_i}{p\rho_i} - 1 \right| < 6m\varepsilon_0 < \frac{\varepsilon}{10}.$$

Hence whenever (29) holds, we have

$$\left| \ln \frac{\zeta}{\mathrm{vol}(K)} \right| \le \left| \ln \frac{b_1 \dots b_{m+1}}{\beta_1 \dots \beta_{m+1}} \right| + \left| \ln \frac{\beta_1 \dots \beta_{m+1}}{(p\rho_1) \dots (p\rho_{m+1})} \right| < \frac{3}{5}\varepsilon.$$

This proves the first assertion. The bound on the number of oracle calls follows easily. □

**Proof of Theorem 2.1** : Using Algorithm 5.19, we can bring the given convex body $K$ into 1/10-nearly isotropic position with probability at least 9/10 in time $O(n^5 (\ln n)^4)$ (see Theorem 5.20). Then we may apply the volume algorithm 6.1 of this section with $d = 2\sqrt{2n}/\log(1/\varepsilon)$ and with probability at least 8/10, get a volume estimate within relative error $\varepsilon/4$ where also, with probability at least 9/10 that the algorithm takes time

$$O\left( \frac{n^5}{\varepsilon^2} (\ln n)^3 (\ln(1/\varepsilon))^3 \right).$$

Applying Remark 4.9, we may again ensure that the algorithm definitely takes time

$$O\left( \frac{n^5}{\varepsilon^2} (\ln n)^3 (\ln(1/\varepsilon))^3 \ln(1/\eta) \right)$$

and fails with probability at most $(2/10) + \eta$ for any $\eta > 0$. Repeating the whole process $O(\ln(1/\eta))$ times and taking the median (as discussed in the Introduction) gives us an estimate of volume within relative error $\varepsilon$ and probability of success at least $1 - \eta$ proving Theorem 2.1.

**Acknowledgements**: We thank Alan Frieze, Vitaly Milman, György Sonnevend and Peter Winkler for many helpful discussions and correspondence.

**References**

D. Applegate and R. Kannan (1990): Sampling and integration of near log-concave functions, *Proc. 23th ACM STOC*, 156–163.

I. Bárány and Z. Füredi (1986): Computing the volume is difficult, *Proc. of the 18th Annual ACM Symposium on Theory of Computing*, pp.442–447.

C. Borell (1975): The Brunn–Minkowski inequality in Gauss spaces, *Invent. Math.* **30** 207-216

M. Dyer and A. Frieze (1988): On the complexity of computing the volume of a polytope. *SIAM J. Comp.* **17**, 967–974.

M. Dyer and A. Frieze (1992): Computing the volume of convex bodies: a case where randomness provably helps, in: *Probabilistic Combinatorics and Its Applications* (ed. Béla Bollobás), Proceedings of Symposia in Applied Mathematics, Vol. 44, 123–170.

M. Dyer, A. Frieze and R. Kannan (1989): A Random Polynomial Time Algorithm for Approximating the Volume of Convex Bodies, *Proc. of the 21st Annual ACM Symposium on Theory of Computing*, 375-381.

G. Elekes (1986): A geometric inequality and the complexity of computing volume, *Discrete and Computational Geometry* **1**, 289-292.

J. A. Fill (1991): Eigenvalue bounds on convergence to stationarity for nonreversible Markov Chains, with an application to the Exclusion Process, The Annals of Applied Probability, Vol. 1, pp 62-87.

M. Gromov and V. D. Milman (1984): Brunn theorem and a concentration of volume of convex bodies, GAFA Seminar Notes, Tel Aviv University.

M. Grötschel, L. Lovász and A. Schrijver (1988): *Geometric Algorithms and Combinatorial Optimization*, Springer-Verlag.

P. M. Gruber and J. M. Wills (1993) (eds): Handbook of Convex Geometry, Elsevier Science Publ.

B. Grünbaum (1960): Partitions of mass-distributions and of convex bodies by hyperplanes, *Pacific J. Math.* **10**, 1257–1261.

M. R. Jerrum, L. G. Valiant and V. V. Vazirani (1986): Random generation of combinatorial structures from a uniform distribution, *Theoretical Computer Science* **43**, 169–188.

R. Kannan, L. Lovász and M. Simonovits (1995): Isoperimetric problems for convex bodies and a localization lemma, *J. Discr. Comput. Geom.* **13** 541–559.

L. G. Khachiyan (1988): On the complexity of computing the volume of a polytope. *Izvestia Akad. Nauk SSSR, Engineering Cybernetics* **3**, 216–217.

L. G. Khachiyan (1989): The problem of computing the volume of polytopes is NP-hard. *Uspekhi Mat. Nauk 44*, 199-200.

L. G. Khachiyan (1993): Complexity of polytope volume computation, in: *New Trends in Discrete and Computational Geometry* (ed. J. Pach), 1993, Springer, 91–101.

J. Lawrence (1989): Polytope volume computation, Preprint NISTIR 89-4123, U.S.Dept. of Commerce, Natl. Inst. of Stand. Tech., Galthersburg.

L. Lovász (1986): An Algorithmic Theory of Numbers, Graphs and Convexity, CBMS-NSF Regional Conference series in Applied Mathematics **50**, SIAM Philadelphia.

L. Lovász (1992): How to compute the volume? *Jber. d. Dt. Math.-Verein, Jubiläumstagung 1990*, B. G. Teubner, Stuttgart, 138–151.

L. Lovász and M. Simonovits (1990): Mixing rate of Markov chains, an isoperimetric inequality, and computing the volume. *Proc. 31st Annual Symp. on Found. of Computer Science,* IEEE Computer Soc., 346–355.

L. Lovász and M. Simonovits (1992): On the randomized complexity of volume and diameter, *Proc. 33rd IEEE FOCS*, 482–491.

L. Lovász and M. Simonovits (1993): Random walks in a convex body and an improved volume algorithm, *Random Structures and Algorithms* **4**, 359–412.

V. D. Milman and A. Pajor (1987): Isotropic position and inertia ellipsoids and zonoids of the unit ball of a normed $n$-dimensional space, GAFA Seminar Notes, Tel Aviv University, Lecture Notes 1376, Springer 64-104.

A. Sinclair and M. Jerrum (1988): Conductance and the rapid mixing property for Markov chains: the approximation of the permanent resolved, *Proc. 20th ACM STOC*, pp. 235–244.

G. Sonnevend (1989): Applications of analytic centers for the numerical solution of semiinfinite, convex programs arising in control theory, DFG report Nr. 170/1989, Univ. Würzburg,

Inst.f. angew. Mathematik.