Flavors of Geometry
MSRI Publications
Volume **31**, 1997

# Volume Estimates and Rapid Mixing

BÉLA BOLLOBÁS

Contents



## Lecture 1. Introduction

Computing, or at least estimating, the volume of a body is one of the oldest questions in mathematics, studied already in Egypt and continued by Euclid and Archimedes. Here we are mostly concerned with the *computational complexity* of estimating volumes of convex bodies in $\mathbb{R}^n$, with $n$ large. The problems and results to be discussed are all rather recent, most of them less than ten years old, and although there was a breakthrough a few years ago and by now there are several substantial and exciting results, there is much to be done.

In vague terms, we would like to find a fast algorithm that computes, for each convex body $K$ in $\mathbb{R}^n$, positive numbers $\underline{\mathrm{vol}}\, K$ and $\overline{\mathrm{vol}}\, K$ such that

$$\underline{\mathrm{vol}}\, K \le \mathrm{vol}\, K \le \overline{\mathrm{vol}}\, K,$$

and $\overline{\mathrm{vol}}\, K / \underline{\mathrm{vol}}\, K$ is as small as possible.

This formulation has several flaws. It is not clear what our algorithm is allowed to do and how its speed is measured; we also have to decide how our convex body is given and how small $\overline{\mathrm{vol}}\, K / \underline{\mathrm{vol}}\, K$ we wish to make. Our first aim is then to make this problem precise.

We write $|x|$ for the standard Euclidean norm of a vector $x \in \mathbb{R}^n$, and $\langle x, y\rangle$ for the standard inner product. $B^n = B_2^n$ denotes the Euclidean ball of radius 1 in $\mathbb{R}^n$, and $B^n(\varepsilon)$ the ball of radius $\varepsilon$. If there is no danger of confusion, we write

$\operatorname{vol} K$ for the volume of a body $K$, in whatever the appropriate dimension is; if we want to draw attention to the dimension we write $\operatorname{vol}_k L$ for the $k$-dimensional volume of a $k$-dimensional body $L$.

By a *convex body* in $\mathbb{R}^n$ we mean a compact convex subset of $\mathbb{R}^n$, with nonempty interior. Following [Grötschel, Lovász, and Schrijver 1988], we shall assume that our convex body $K \subset \mathbb{R}^n$ is given by a certain oracle; this will enable us to obtain results that are valid for a large class of specific algorithms. An *oracle* is a "black box" that answers various questions put to it. When we talk about a convex body $K$ given by an oracle, the questions tend to be simple, for example: "What about the point $x \in \mathbb{R}^n$?" A *strong membership oracle* answers this question in one of two ways: "$x \in K$" or "$x \notin K$." A *strong separation oracle* answers either "$x \in K$" or, in addition to replying that "$x \notin K$," it gives a linear functional separating $x$ from $K$. In other words, when the answer is negative, the oracle justifies its assertion by displaying a vector $c \in \mathbb{R}^n$ such that $\langle c, x\rangle > \langle c, y\rangle$ for every $y \in K$. We can assume that $\|c\|_\infty = 1$, where $\|\cdot\|_\infty$ denotes the sup norm.

In a *weak membership oracle* the question is slightly different: "What about the point $x \in \mathbb{R}^n$ and the positive number $\varepsilon$?" Let

$$K_\varepsilon = K + B^n(\varepsilon) = \{y \in \mathbb{R}^n : |y - z| \le \varepsilon \text{ for some } z \in K\}$$

and

$$K_{-\varepsilon} = \mathbb{R}^n \setminus (\mathbb{R}^n \setminus K)_\varepsilon = \{y \in \mathbb{R}^n : |y - z| > \varepsilon \text{ for every } z \notin K\};$$

the weak membership oracle answers either "$x \in K_\varepsilon$" or "$x \notin K_{-\varepsilon}$". (For $x \in K_\varepsilon \setminus K_{-\varepsilon}$ it may return either answer.)

Similarly, a *weak separation oracle* replies to the same question in one of the following two ways: either

$$\text{“}x \in K_\varepsilon\text{”}$$

or "$x \notin K_{-\varepsilon}$ and here is a functional $c \in \mathbb{R}^n$, $\|c\|_\infty = 1$, proving it:

$$\langle c, y\rangle < \langle c, x\rangle + \varepsilon \quad \text{for all } y \in K_{-\varepsilon}.\text{”}$$

An algorithm then is a sequence of questions to the oracle, each question depending on the answers to the previous questions. The *complexity* or *running time* of an algorithm is the number of questions asked before the bounds are produced.

A moment's thought tells us that with the oracles just described no algorithm can produce bounds other than $\underline{\operatorname{vol}} K = 0$ and $\overline{\operatorname{vol}} K = \infty$, since we can do no better than this if we imagine that $K$ is "at infinity" in the direction of the $x_1$-axis, say; if to the question: "What about $x = (x_i)_1^n$?" the oracle replies that

$$K \subset \{y = (y_i)_1^n : y_1 > x_1 + 1\},$$

then after some questions all we know is that $\min\{y_1 : (y_i) \in K\}$ is large, but we know nothing about the volume of the body. Indeed, we cannot even find a single point of $K$ in a finite time, even if $K$ is known to have large volume. In order to give the algorithm a chance, we need some guarantees about $K$, namely that it is not "at infinity" and it is not too small. The standard way to do this is to assume that

$$rB^n \subset K \subset RB^n \tag{1.1}$$

for some positive numbers $r$ and $R$. If (1.1) holds we say that the algorithm is *well guaranteed*, with guarantees $r$ and $R$.

We should also specify how the data are measured: the *size* of the input $(r, R, n)$ for a convex body $K \subset \mathbb{R}^n$ satisfying (1.1) is

$$\langle K\rangle = n + \langle r\rangle + \langle R\rangle,$$

where $\langle x\rangle$ is the number of binary digits of a dyadic rational $x$.

In what follows, it will make very little difference which of the above oracles we shall use: the difficulty is not in going from one oracle to another but in finding suitable algorithms. For example, Grötschel, Lovász, and Schrijver [1988] showed that a weak separation oracle can be obtained from a weak membership oracle in polynomial time.

We are only interested in algorithms that are fairly fast, namely those whose complexity is polynomial and of not too high a degree. In view of this, our problem could be restated as follows. Given a polynomial $f(x) = x^a + b$, find a function $g(x)$ such that, if the oracle describing our convex body $K \subset \mathbb{R}^n$ has guarantees $2^{-l_1}$ and $2^{l_2}$, we can compute, after no more than $f(n + l_1 + l_2)$ appeals to the oracle, numbers $\underline{\mathrm{vol}}\, K$ and $\overline{\mathrm{vol}}\, K$ such that

$$\underline{\mathrm{vol}}\, K \le \mathrm{vol}\, K \le \overline{\mathrm{vol}}\, K$$

and $\overline{\mathrm{vol}}\, K / \underline{\mathrm{vol}}\, K \le g(n)$. Moreover, $g(x)$ should grow as slowly as possible.

As we shall see in the next section, the solution to this problem is rather disappointing: no matter what polynomial $f$ we choose, the approximation $g(n)$ cannot be guaranteed to be better than polynomial. However, if we do not insist that our approximations $\underline{\mathrm{vol}}\, K$ and $\overline{\mathrm{vol}}\, K$ be valid every time, we can do much better. In 1989, Dyer, Frieze, and Kannan [1991] devised a randomized algorithm that approximates with high probability the volume of a convex body as closely as desired in polynomial time. After this breakthrough, faster and faster algorithms have been devised, but it is unlikely that we are near to a best possible algorithm. In Section 4 we describe one of the most elegant (although not quite the fastest) algorithms found so far. This algorithm and every other are based on rapidly mixing random walks: we shall present the relevant results in Section 3.

## Lecture 2. Volumes of Convex Hulls and Deterministic Bounds

Our aim in this section is to show that there is no fast deterministic algorithm for approximating the volume of a convex body:

THEOREM 2.1. *For every polynomial-time algorithm for computing the volume of a convex body in* $\mathbb{R}^n$ *given by a well-guaranteed separation oracle*, *there is a constant* $c > 0$ *such that*

$$\frac{\overline{\mathrm{vol}}\,K}{\underline{\mathrm{vol}}\,K} \le \left(\frac{cn}{\log n}\right)^n$$

*cannot be guaranteed for* $n \ge 2$.

Nevertheless, we start with a positive result, claiming that in polynomial time we *can* achieve some kind of approximation. To be precise, Lovász [1986] proved that for every convex body $K$ given by a well-guaranteed oracle there is an affine transformation $\varphi : x \to Ax + b$ computable in polynomial time and such that

$$B \subset \varphi(K) \subset n\sqrt{n+1}\,B.$$

In particular, this algorithm produces estimates $\underline{\mathrm{vol}}\,K$ and $\overline{\mathrm{vol}}\,K$ with

$$\overline{\mathrm{vol}}\,K/\underline{\mathrm{vol}}\,K \le n^n(n+1)^{n/2}.$$

Furthermore, Lovász showed that if $K$ is centrally symmetric—say, if it is the unit ball of a norm on $\mathbb{R}^n$—there is a polynomial-time algorithm that produces estimates $\underline{\mathrm{vol}}\,K$ and $\overline{\mathrm{vol}}\,K$ with $\overline{\mathrm{vol}}\,K/\underline{\mathrm{vol}}\,K \le n^n$.

Elekes [1986] was the first to realize that this bound is not as outrageous as it looks at first sight. Indeed, he showed that, for $0 < \varepsilon < 2$, there is no polynomial-time algorithm that returns $\overline{\mathrm{vol}}\,K$ and $\underline{\mathrm{vol}}\,K$ with

$$\overline{\mathrm{vol}}\,K/\underline{\mathrm{vol}}\,K \le (2-\varepsilon)^n.$$

What Elekes noticed was that the convex hull of polynomially many points in $B^n$ is only a small fraction of $B^n$, and that this fact implies that every polynomial algorithm is bound to give a poor result. The theorem of Elekes was soon improved by Bárány and Füredi [1988] to an essentially best possible result: this is the main result we shall present.

Let's start with the problem of approximating the unit ball $B^n \subset \mathbb{R}^n$ by the convex hull of $m$ points of $B^n$. As we shall see later, a similar (and essentially equivalent) problem is that of approximating $B^n$ by the intersection of $m$ slabs, each containing $B^n$. A question of this type was considered in Section 2 of Keith Ball's lectures in this volume [Ball 1997], where the measure of approximation was the Banach–Mazur distance. Here our aim is rather different: we wish to approximate $B^n$ by polytopes contained by $B^n$ (or containing $B^n$) that have relatively few vertices, and the measure of our approximation is the difference of volumes.

The following beautiful and simple result of Elekes [1986] shows that we cannot hope for a good approximation unless we take exponentially many points. This result is reminiscent of [Ball 1997, Theorem 2.1], but the proof is considerably simpler.

THEOREM 2.2. *Let* $v_1, \dots, v_m \in \mathbb{R}^n$ *and* $K = \operatorname{conv}\{v_1, \dots, v_m\}$. *Then* $K \subset \bigcup_{i=1}^m B_i$, *where* $B_i = B(v_i/2, |v_i|/2)$ *is the ball of centre* $v_i/2$ *and radius* $|v_i|/2$. *In particular*, *if each* $v_i$ *is in the unit ball* $B^n$ *of* $\mathbb{R}^n$ *then*

$$\operatorname{vol}_n K / \operatorname{vol}_n B^n \le m/2^n.$$

PROOF. Suppose $x \notin B_i$, so that $\langle x - v_i/2,\, x - v_i/2\rangle > \|v_i\|^2/4$. Then

$$|x|^2 > \langle x, v_i\rangle,$$

that is, $v_i$ is in the open half-space $H(x) = \{y \in \mathbb{R}^n : \langle x, y\rangle < \|x\|^2\}$. Hence if $x \notin \bigcup_{i=1}^m B_i$ then $v_i \in H(x)$ for every $i$, and so $K = \operatorname{conv}\{v_1, \dots, v_m\} \subset H(x)$. But $x \notin H(x)$, so $x \notin K$. Hence $K \subset \bigcup_{i=1}^m B_i$, as claimed. The last statement of the theorem follows immediately. □

To give a more geometric argument for why $K \subset \bigcup_{i=1}^m B_i$, note that $B_i$ is the set of points $x$ for which the angle $v_i\, x\, v_0$ is at least $\pi/2$, where $v_0 = 0 \in \mathbb{R}^n$. All we have to notice is that if $v = \lambda v_i + (1-\lambda) v_j$ for some $\lambda$ with $0 < \lambda < 1$, then $x \notin B_i \cup B_j$ implies that the angle $v\, x\, v_0$ is less than $\pi/2$, as the angles $v_i\, x\, v_0$ and $v_j\, x\, v_0$ are less than $\pi/2$. But then if $x \notin \bigcup_{i=1}^m B_i$ and $v \in K$ then the angle $v\, x\, v_0$ is less than $\pi/2$. Hence $\bigcup_{i=1}^m B_i$ contains the set

$$\alpha(K) = \{x \in \mathbb{R}^n : \text{angle } v\, x\, v_0 \ge \pi/2 \text{ for some } v \in K\},$$

which certainly contains $K$.

If $m$ is not too large the inequality of Theorem 2.2 is fairly good, but if $m$ is exponential, let alone at least $2^n$, it is very weak. Our next aim is to present a result of Bárány and Füredi [1988] (Theorem 2.5) that gives an essentially best possible bound.

Let's write $V(n, m)$ for the maximal volume of the convex hull of $m$ points in $B^n$:

$$V(n, m) = \max\left\{\operatorname{vol}_n K : K = \operatorname{conv}\{v_1, \dots, v_m\} \subset B^n\right\},$$

and let

$$W(n, m) = \frac{V(n, m)}{\operatorname{vol}_n B^n}$$

be the proportion of the volume. In order to get our upper bound for $V(n, m)$, we need an extension of a result from [Fejes Tóth 1964] closely related to Fritz John's theorem [1948]. Theorem 3.1 in [Ball 1997] is a sharper version of Fritz John's theorem: here we shall need only the original result.

THEOREM 2.3. *For every convex body $K \subset \mathbb{R}^n$, there is a unique ellipsoid $E$ of maximal volume contained in $K$. If $E$ has centre* 0 *then*

$$E \subset K \subset nE. \tag{2.1}$$

PROOF. By a simple compactness argument, there is at least one ellipsoid of maximal volume. To prove uniqueness, suppose that there are two ellipsoids of maximal volume, say $E$ and $E'$. Then the ellipsoid $\frac{1}{2}(E+E')$ is contained in the convex hull of $E \cup E'$. By the Brunn–Minkowski Theorem (or by the AM/GM inequality), $\mathrm{vol}_n\bigl(\frac{1}{2}(E+E')\bigr)$ is greater than $\mathrm{vol}_n E = \mathrm{vol}_n E'$, unless $E'$ is a translate of $E$. If $E'$ is a translate of $E$ we can assume without loss of generality that each is a unit ball, and then $\mathrm{conv}(E \cup E')$ is easily seen to contain an ellipsoid $E''$ with $\mathrm{vol}_n E'' > \mathrm{vol}_n E$. Hence $E$ is indeed unique.

To see that $K \subset nE$, all we have to check is that if $B^n$ is the unit ball in $\mathbb{R}^n$ and $|v| > n$ then $\mathrm{conv}(B^n \cup \{v\})$ contains another ellipsoid of volume $\mathrm{vol}_n B^n$. This can be done by simple calculations. $\square$

Let $S_0$ be a regular simplex with inscribed ball $B_0 = B^n$ and so with circumscribed ball $nB_0$. By the uniqueness of the ellipsoid of maximal volume in a convex body, $B_0$ is *the* ellipsoid of maximal volume in $S_0$. Also, it is simple to see that $S_0$ is a simplex of maximal volume contained in $nB_0$. Since the volume ratio is affine invariant, it follows that if $S$ is any simplex in $\mathbb{R}^n$ and $E_1 \subset S \subset E_2$ for some ellipsoids $E_1$ and $E_2$, then $\mathrm{vol}\, E_1/\mathrm{vol}\, E_2 \le n^{-n}$. In particular, if $E \subset S \subset \lambda E$ for some ellipsoid $E$ and positive real $\lambda$ then $\lambda \ge n$. *A fortiori*, if $K$ is a simplex, we cannot replace $n$ in Theorem 2.3 by a smaller constant.

Also, if a simplex $S \subset \mathbb{R}^n$ contains a ball of radius $r_1$ and is contained in a ball of radius $r_2$, then $r_2 \ge nr_1$. This last assertion is the result from [Fejes Tóth 1964] that we shall need and extend.

In order to state and prove this extension, we introduce some notation. First, given a set $S \subset \mathbb{R}^n$, let $U = \mathrm{span}\, S = \mathrm{lin}\{x-y : x, y \in S\}$ be the subspace of $\mathbb{R}^n$ defined by $S$, and for $\rho > 0$ set

$$S^\rho = S + \bigl(U^\perp \cap \rho B^n\bigr),$$

where $U^\perp$ is the orthogonal complement of the subspace $U$. Thus $S^\rho$ is the set of $x \in \mathbb{R}^n$ for which $n(x,S)$ is attained at some $y \in S$ with $n(x,y) \le \rho$ and $(x-y) \perp U$. If $S$ is convex and $\dim U = k$, then clearly

$$\mathrm{vol}_n S^\rho = (\mathrm{vol}_k S)\bigl(\mathrm{vol}_{n-k}\, B^{n-k}\bigr)\, \rho^{n-k}. \tag{2.2}$$

Secondly, for $1 \le k \le n$ define

$$\rho(n,k) = \begin{cases} 1 & \text{if } k = 0, \\ \sqrt{(n-k)/(nk)} & \text{if } 1 \le k \le n-2, \\ 1/n & \text{if } k = n-1. \end{cases}$$

LEMMA 2.4. *Let* $x \in S = \operatorname{conv}\{v_0, v_1, \dots, v_n\} \subset B^n$. *Then, for every* $k$ *such that* $0 \le k \le n-1$, *the simplex* $S$ *has a* $k$*-dimensional face* $S_k = \operatorname{conv}\{v_{i_0}, v_{i_1}, \dots, v_{i_k}\}$ *and a point* $x_k$ *in the interior of the face* $S_k$ (*in the sense that* $x_k = \sum_{j=0}^{k} \lambda_j v_{i_j}$ *with* $\sum_{j=0}^{k} \lambda_j = 1$ *and* $\lambda_j > 0$ *for every* $j$), *such that* $(x - x_k) \perp \operatorname{span} S_k$ *and* $\|x - x_k\| \le \rho(n,k)$. *In particular*, $x \in S_k^{\rho(n,k)}$.

PROOF. We know the result for $k = n-1$, and from it we shall deduce the result for $1 \le k \le n-2$. Set $x_n = x$, $S_n = S$, and let $S_{n-1}$ be a $(n-1)$-dimensional face of $S_n$ containing a point $x_{n-1}$ such that $(x_n - x_{n-1}) \perp \operatorname{span} S_{n-1}$ and $|x_n - x_{n-1}| \le 1/n$. Next, let $S_{n-2}$ be a $(n-2)$-dimensional face of $S_{n-1}$ containing a point $x_{n-2}$ such that $(x_{n-1} - x_{n-2}) \perp \operatorname{span} S_{n-2}$ and $|x_{n-1} - x_{n-2}| \le 1/(n-1)$. Proceed in this way up to $x_k$ in $S_k$. Then the vectors $x_n - x_{n-1}$, $x_{n-1} - x_{n-2}$, $\dots$, $x_{k+1} - x_k$ are orthogonal, with $|x_l - x_{l-1}| \le 1/l$. Hence

$$x - x_k = x_n - x_k = (x_n - x_{n-1}) + (x_{n-1} - x_{n-2}) + \cdots + (x_{k+1} - x_k)$$

is orthogonal to $\operatorname{span} S_k$ and

$$\begin{aligned}
|x - x_k|^2 &\le \frac{1}{n^2} + \frac{1}{(n-1)^2} + \cdots + \frac{1}{(k+1)^2} \\
&\le \frac{1}{n(n-1)} + \frac{1}{(n-1)(n-2)} + \cdots + \frac{1}{k(k+1)} \\
&= \frac{1}{k} - \frac{1}{k+1} + \frac{1}{k+1} - \frac{1}{k+2} + \cdots + \frac{1}{n-1} - \frac{1}{n} = \frac{1}{k} - \frac{1}{n} = \frac{n-k}{nk},
\end{aligned}$$

as required.

Finally, let's consider the case $k = 0$. Suppose that $\|x - v_i\| > 1$ for every $i$, and so

$$\{v_0, v_1, \dots, v_n\} \subset B^n \setminus B(x,1).$$

But then every point of $S = \operatorname{conv}\{v_0, v_1, \dots, v_n\}$ is closer to 0 than to $x$ and so $x \notin S$. Hence $S \subset \bigcup_{i=0}^{n} B(v_i, 1)$, as claimed. $\square$

The alert reader must have noticed that this simple proof does not give a tight bound except in the cases $k = n$ and $k = 1$. This is because the later simplices $S_{n-1}, S_{n-2}, \dots, S_{k+1}$ are likely to be contained in balls of radii less than 1. However, the loss is surprisingly little. Trivially, we cannot do better than in the case of a regular simplex inscribed in $B^n$: to cover the origin by neighbourhoods of the $k$-dimensional faces we must have $\rho(n,k)$ at least $\sqrt{(n-k)/(n(k+1))}$ rather than $\sqrt{(n-k)/(nk)}$. It is very likely that, in fact, $\rho(n,k) = \sqrt{(n-k)/(n(k+1))}$.

The Bárány–Füredi upper bound for $V(n,m)$ follows easily from Lemma 2.4:

THEOREM 2.5. *There is a constant* $c > 0$ *such that for* $m = m(n) \ge 1$ *we have*

$$W(n,m) \le \left(\frac{c(\log(m/n) + 1)}{n}\right)^{n/2}$$

*and so*

$$V(n,m) \le \left(\frac{\gamma(\log(m/n)+1)^{1/2}}{n}\right)^n,$$

*where* $\gamma = (2\pi ec)^{1/2}$. *Furthermore, if* $\varepsilon > 0$ *is fixed and we take* $m/n \to \infty$ *and* $n/\log(m/n) \to \infty$, *then*

$$V(n,m) \le \left(\frac{(2e+\varepsilon)\log(m/n)}{n^2}\right)^{n/2}.$$

Proof. Let $K = \operatorname{conv}\{v_1, v_2, \dots, v_m\} \subset B^n$. By Carathéodory's theorem ([Carathéodory 1907]; see also [Eckhoff 1993]) $K$ is the union of its $n$-dimensional simplices:

$$K = \bigcup_{i_0 < \dots < i_n} \operatorname{conv}\{v_{i_0}, v_{i_1}, \dots, v_{i_n}\}.$$

Hence, by Lemma 2.4, for all $k$ such that $1 \le k \le n-1$ we have

$$K \subset \bigcup_{i_0 < \dots < i_k} \bigl\{S^{\rho(n,k)} : S = \operatorname{conv}\{v_{i_0}, \dots, v_{i_k}\}\bigr\},$$

and so

$$\operatorname{vol}_n K \le \binom{m}{k+1} \max\bigl\{\operatorname{vol}_n S^{\rho(n,k)} : S = \operatorname{conv}\{x_0, x_1, \dots, x_k\} \subset B^n\bigr\}.$$

By identity (2.2), for a simplex $S$ as above,

$$\operatorname{vol}_n S^{\rho(n,k)} = (\operatorname{vol}_k S)\left(\operatorname{vol}_{n-k} B^{n-k}\right)\rho(n,k)^{n-k}.$$

Furthermore, easy computations show that the maximal volume of an $n$-simplex in $B^n$ is $(n+1)^{(n+1)/2}/n^{n/2}n!$ and

$$\operatorname{vol}_n B^n = \frac{\pi^{n/2}}{\Gamma(n/2+1)} \le (2\pi e/n)^n.$$

Putting together the last four relations and the definition of $\rho(n,k)$ we get

$$\operatorname{vol}_n K \le \binom{m}{k+1} \frac{(k+1)^{(k+1)/2}}{k^{k/2}k!} \frac{\pi^{(n-k)/2}}{\Gamma((n-k+2)/2)} \left(\frac{n-k}{nk}\right)^{(n-k)/2}.$$

Therefore

$$\operatorname{vol}_n K \le \left(\frac{em}{k+1}\right)^{k+1} \left(\frac{e}{k}\right)^k \left(\frac{2e\pi}{nk}\right)^{(n-k)/2}$$

and so

$$\frac{\operatorname{vol}_n K}{\operatorname{vol}_n B^n} \le \left(\frac{em}{k+1}\right)^{k+1} n^{k/2} k^{-(n+k)/2}. \tag{2.3}$$

All that remains is to find a value of $k$ for which the right-hand side is small.

Let's do this under the assumptions $m/n \to \infty$ and $n/\log(m/n) \to \infty$; the existence of $c$ can be shown similarly. We claim that $k = \lceil n/2\log(m/n) \rceil$ is a suitable choice. To avoid too much clutter, we shall just take $k = n/2\log(m/n)$. Note that $k \to \infty$, $k = o(n)$ and

$$\left(\frac{m}{k+1}\right)^k \le \exp(k\log(m/k)) = \exp\left(k(\log(m/n) + \log(n/k))\right)$$
$$= \exp\left(\frac{n}{2} + n\frac{\log(2\log(m/n))}{2\log(m/n)}\right) = e^{(1+o(1))n/2}. \tag{2.4}$$

Also,

$$(n/k)^k = (2\log(m/n))^k = \exp\left(\frac{n}{2\log(m/n)}\log(2\log(m/n))\right) = e^{o(n)},$$

since $m/n \to \infty$. Hence

$$n^{k/2}k^{-(n+k)/2} = (n/k)^{(n+k)/2}n^{-n/2} = e^{o(n)}(2\log(m/n))^{n/2}n^{-n/2}.$$

Together with (2.3) and (2.4), this implies that

$$\frac{\mathrm{vol}_n K}{\mathrm{vol}_n B^n} \le e^{o(n)}\left(2e\log(m/n)/n\right)^{n/2},$$

completing the proof. □

Theorem 2.5 is essentially best possible in a large range of $m$, except for the constant $2e$. In fact, the theorem can be read out of some earlier results of Carl [1985], and over the years it has been discovered many times, having been published in [Bárány and Füredi 1988; Carl and Pajor 1988; Gluskin 1988; Bourgain, Lindenstrauss, and Milman 1989]. Numerous related results can be found in [Vaaler 1979; Figiel and Johnson 1980; Bárány and Füredi 1986; 1987; Ball and Pajor 1990; Gordon, Reisner, and Schütt ≥ 1997], and elsewhere.

We see from Theorem 2.5 that a polytope $K$ contained in $B^n$ with $\mathrm{vol}_n\, K \ge \frac{1}{2}\mathrm{vol}_n\, B^n$, say, has exponentially many vertices. In fact, if $\mathrm{vol}_n K$ can be close to $\mathrm{vol}_n B^n$ then $1 - \mathrm{vol}_n K/\mathrm{vol}_n B^n$ is a more significant measure of the volume approximation. Gordon, Reisner and Schütt [≥ 1997] proved that in order to get $1-\varepsilon$ proportion of the volume, we need about $n^{n/2}$ points: there are positive constants $\varepsilon_0$ and $\varepsilon_1$ such that $1 - \mathrm{vol}_n K/\mathrm{vol}_n B^n \ge \varepsilon_0$ whenever $K$ is a polytope in $B^n$ with $m \le (\varepsilon_1 n)^{n/2}$ vertices—in other words,

$$W(n,m) \le 1 - \varepsilon_0$$

if $m \le (\varepsilon_1 n)^{n/2}$. Even more, there are positive constants $\delta_0$ and $\delta_1$ such that

$$1 - W(n,m) \ge \delta_0 n m^{-2/(n-1)}$$

whenever $n \ge 2$ and $m \ge (\delta_1 n)^{n/2}$.

Many of the papers fleetingly mentioned above concern the volume of the intersection of slabs, rather than the volume of the convex hull of points. In order to prove the main result of this section, namely that computing the volume is difficult, we need one of these results as well. Given natural numbers $n$ and

$m$, let $S(n,m)$ be the infimum of the volumes of intersections of $m$ slabs in $\mathbb{R}^n$, each of the form

$$\{x : |\langle x, v\rangle| \le 1\},$$

where $v \in \mathbb{R}^n$ is a vector of length at most 1. The intersection of $m$ such slabs is precisely a centrally symmetric polytope $K$ containing $B^n$, with at most $2m$ facets [Ball 1997, Theorem 2.1]. The following lower bound for $S(n,m)$ is given in [Carl and Pajor 1988; Gluskin 1988].

THEOREM 2.6. *There is a constant $\delta > 0$ such that if $1 \le n \le m$ then*

$$S(n,m) \ge \left(\frac{\delta}{(\log(m/n)+1)^{1/2}}\right)^n.$$

Rather than proving this directly (which would not be difficult), we shall deduce it from Theorem 2.5 and a beautiful and important result, the reverse Santaló inequality of Bourgain and Milman. For a convex body $K$ in $\mathbb{R}^n$, the *polar* of $K$ is

$$K^\circ = \{x \in \mathbb{R}^n : \langle x, y\rangle \le 1 \text{ for all } y \in K\}.$$

If $K$ is a *ball*, that is the unit ball $B(X)$ of a normed space $X = (\mathbb{R}^n, \|\cdot\|)$, then $K^\circ$ is precisely the unit ball of the dual: $K^\circ = B(X^*)$. For us this is precisely the most important case.

What can one say about the product $\operatorname{vol} K \operatorname{vol} K^\circ$ for a ball $K \subset \mathbb{R}^n$? Santaló [1949] proved that it is at most $(\operatorname{vol} B^n)^2$, so that the maximum is attained for $K = B^n$. Thus

$$\operatorname{vol} K \operatorname{vol} K^\circ \le \pi^n/\Gamma(n/2+1)^2 \sim \left(\frac{2\pi e}{n}\right)^n \Big/ \pi n = (2e)^n \pi^{n-1}/n^{n+1}.$$

Taking $X = l_1^n$, so that $X^* = l_\infty^n$, we see that $\operatorname{vol} K \operatorname{vol} K^\circ$ can be as small as $\operatorname{vol} B(l_1^n) \operatorname{vol} B(l_\infty^n) = 4^n/n! \sim (4e/n)^n/\sqrt{2\pi n}$. Mahler conjectured that this value is, in fact, the minimum of the product. Although this long-standing conjecture is still open, Bourgain and Milman [1985] proved the following reverse Santaló inequality, which is only a little weaker than Mahler's conjecture.

THEOREM 2.7. *There is a constant $c_0 > 0$ such that if $K$ is any ball in $\mathbb{R}^n$ then*

$$\operatorname{vol} K \operatorname{vol} K^\circ \ge (c_0/n)^n. \qquad \square$$

It is frequently convenient to state both inequalities together as follows: if $K$ is a ball in $\mathbb{R}^n$ then

$$c_0 \le \left(\frac{(\operatorname{vol} K)(\operatorname{vol} K^\circ)}{(\operatorname{vol} B^n)^2}\right)^{1/n} \le 1$$

for some constant $c_0 > 0$.

Theorem 2.6 is an easy consequence of Theorems 2.5 and 2.7. Indeed, let

$$L = \{x \in \mathbb{R}^n : |\langle x, u_i\rangle| \le 1,\ i = 1, \dots, m\},$$

where $u_1, \ldots, u_m \in B^n$, be the intersection of $m$ slabs containing $B^n$. Then $L = K^\circ$, where $K = \operatorname{conv}\{u_1, \ldots, u_m\}$. Hence, by Theorems 4 and 5,

$$\begin{aligned}\operatorname{vol} L &\geq (c_0/n)^n / \operatorname{vol} K \geq (c_0/n)^n (n/(\gamma \log(m/n) + 1)^{1/2})^n \\ &= \left(c_0^2 \gamma (\log(m/n) + 1)\right)^{n/2}.\end{aligned}$$

Proof of Theorem 2.1. It suffices to prove the theorem for large $n$. Playing the role of the oracle, we shall give away much more than we have to. First of all, we specify that

$$B\left(l_1^n\right) \subset K \subset B\left(l_\infty^n\right).$$

Thus $r \leq 1/\sqrt{n}$ and $R \geq \sqrt{n}$ will do, so we can have input size at most $2n$. For $x \in \mathbb{R}^n$, $x \neq 0$, define $x^\circ = x/\|x\|$, $H^+(x^\circ) = \{z \in \mathbb{R}^n : \langle z, x^\circ \rangle \leq 1\}$, and $H^-(x^\circ) = \{z \in \mathbb{R}^n : \langle z, -x^\circ \rangle \leq 1\}$. Here $H^+(x_\circ)$ and $H^-(x_\circ)$ are half-spaces containing $B^n$, and their intersection is a slab.

To the question posed by the algorithm "And what about $x$?", the oracle replies very generously that $x^\circ \in K$, $-x^\circ \in K$, and $K$ is contained in the slab $H^+(x^\circ) \cap H^-(x^\circ)$. This is, of course, consistent with $K = B^n$.

Now let's run the algorithm until $m \leq d^a/2 - n$ questions have been asked for some $a \geq 2$, say $x_1, x_2, \ldots, x_m$. Setting $C = \operatorname{conv}\{\pm e_1, \pm e_2, \ldots, \pm e_n, \pm x_1^\circ, \pm x_2^\circ, \ldots, \pm x_m^\circ\}$, we see that the answers are consistent with $K = C$ and $K = C^\circ$ as well. Consequently,

$$\overline{\operatorname{vol}}\, C = \overline{\operatorname{vol}}\, C^\circ \geq \operatorname{vol} C^\circ$$

and

$$\underline{\operatorname{vol}}\, C \leq \operatorname{vol} C.$$

Therefore

$$\frac{\overline{\operatorname{vol}}\, C}{\underline{\operatorname{vol}}\, C} \geq \frac{\operatorname{vol} C^\circ}{\operatorname{vol} C} \geq \frac{S(n, n^a)}{V(n, n^a)}.$$

By Theorems 2.5 and 2.6,

$$\frac{\overline{\operatorname{vol}}\, C}{\underline{\operatorname{vol}}\, C} \geq \frac{\left(\delta/(a \log n)^{1/2}\right)^n}{\left(\gamma (a \log n)^{1/2}/n\right)^n} = \left(\frac{n}{\gamma \delta a \log n}\right)^n,$$

proving the assertion. □

Numerous related results concerning the hardness of approximations can be found in [Khachiyan 1988; 1989; 1993; Lawrence 1991; Lovász and Simonovits 1992].

To conclude this section, let's say a few words about the volumes of intersections of slabs. For $u_1, \ldots, u_m \in \mathbb{R}^n$, set

$$S(u_1, \ldots, u_m) = \{x \in \mathbb{R}^n : |\langle x, u_i \rangle| \leq 1,\ i = 1, \ldots, m\}.$$

Then Theorem 2.6 claims that if $\max |u_i| \leq 1$ then

$$\operatorname{vol} S(u_1, \ldots, u_m) \geq \{\delta/(\log(m/n) + 1)\}^n.$$

The earliest significant slab-intersection theorem is in [Vaaler 1979], and says that if $\sum_{i=1}^m |u_i|^2 \le n$ then

$$\operatorname{vol} S(u_1,\dots,u_m) \ge 2^n. \tag{2.5}$$

An attractive reformulation of this result is the following: for $1 \le k \le n$, any central section of the unit cube $[-1/2,-1/2]^n$ by a $k$-dimensional subspace has volume at least 1. For $k = n-1$ this was first proved by Hensley [1979], who also showed that such an $(n-1)$-dimensional intersection has volume at most 5. Subsequently, Ball [1986] improved the upper bound to the following surprising and beautiful best possible result: any section of the unit cube $[-1/2,1/2]^n$ by an $(n-1)$-dimensional affine subspace has volume at most $\sqrt{2}$.

Clearly, Theorem 2.6 and Vaaler's inequality (2.5) are the $p=\infty$ and $p=2$ members of a family of inequalities parameterized by $p$. The general case was proved in [Ball and Pajor 1990]: if $1 \le p < \infty$, $m \ge n$, and $u_1,\dots,u_m \in \mathbb{R}^n$ are such that $\sum_{i=1}^m |u_i|^p \le r^p n$, then

$$\operatorname{vol} S(u_1,\dots,u_m) \ge \begin{cases} (2\sqrt{2}/\sqrt{p}\,r)^n & \text{if } p \ge 2,\\ r^{-n} & \text{if } 1 \le p \le 2.\end{cases}$$

Vaaler's theorem is the case $p=2$ of this result. It is interesting to note that Vaaler's theorem was used by Bombieri and Vaaler [Bombieri and Vaaler 1983] to sharpen an important result in the geometry of numbers, namely Siegel's lemma. In turn, Ball and Pajor made use of their extension above to prove a generalization of Siegel's lemma.

## Lecture 3. Rapidly Mixing Random Walks

We saw in the preceding lecture that there is a polynomial-time algorithm that, for every convex body $k \subset \mathbb{R}^n$, produces volume estimates $\underline{\operatorname{vol}}\,K$ and $\overline{\operatorname{vol}}\,K$ satisfying $\overline{\operatorname{vol}}\,K/\underline{\operatorname{vol}}\,K \le n^n$, and this is the best one can do, except for a factor $(c\log n)^{-n}$. The exciting part of the story is that if we are willing to replace *certainty* by *high probability*—that is, if we are willing to consider *randomized algorithms* that fail with a small probability—then we can do much better. Estimating the volume of a convex body $K$ is akin to sampling at random from the uniform distribution on $K$. In order to find a random point of $K$, one runs a random walk on $K$ (to be precise, a discrete version of $K$) till the distribution of the last point is close to the stationary distribution, which is the uniform distribution. The problem is then to decide when we can stop so that we are likely to be close to the stationary distribution. This leads us to the question of *mixing time*, the time it takes to get close to the stationary distribution, and to criteria for *rapid mixing*, that is getting close to the stationary distribution in unexpectedly few steps. The aim of this section is to give a beautiful and simple condition for rapid mixing in terms of the conductance of the random walk.

Alon and Milman [Alon and Milman 1985; Alon 1986] were the first to connect combinatorial properties—especially expansion properties—of a graph with the second eigenvalue of its Laplace operator. Loosely speaking, a graph $G$ with $n$ vertices *expands well* if, for every set $U$ of at most $n/2$ vertices, there are relatively many edges with precisely one endvertex in $U$. The *Laplacian* of a simple graph $G = (V, E)$ (where $V$ is the set of vertices, $E$ is the set of edges, and *simple* means there is at most one edge joining two vertices and no loops from a vertex to itself) is the linear operator map $Q : L^2(V) \to L^2(V)$ given by the matrix

$$\operatorname{diag}(d(v))_{v\in V} - A,$$

where $d(v)$ is the *degree* of the vertex $v \in V$ (the number of edges incident on $v$) and $A$ is the *adjacency matrix* of $G$ (the matrix whose rows and columns are indexed by $V$ and where each entry is 1 or 0, depending on whether or not there is an edge in $E$ connecting the two vertices in question). (See [Bollobás 1979] for details and other standard terminology.)

Alon and Milman also proved a discrete version of Cheeger's inequality [1970] related to isoperimetric inequalities on manifolds. Connecting expansion with mixing time, Aldous [1987] showed that random walks on graphs with good expansion properties of low degree mix rapidly. Building on these ideas, Jerrum and Sinclair [1989; Sinclair and Jerrum 1989] defined the conductance of a random walk, and showed that large conductance implies fast mixing rate.

Our aim here is to present the connection between conductance and mixing rate. Rather than consider general Markov chains, we shall take essentially the simplest case, that of simple random walks on regular graphs. As so often, it takes no effort to step up from here to a more general setting. We shall follow the simple and elegant approach of [Mihail 1989]; for a more substantial review of random walks, conductances and eigenvalues, see [Vazirani 1991]; for the related spectral properties of graphs, see [Chung 1996].

Let $G = (V, E)$ be a connected $d$-regular simple graph ($d$-regular means that every vertex has degree $d$). We write $V = \{1, \dots, n\}$ for notational simplicity. For the purposes of these lectures, a *simple random walk on $G$ with initial state $X_0$* is a sequence of random variables

$$\tilde{X} = (X_0, X_1, \dots),$$

taking values in $V$, such that for $i, j \in V$ and $t \geq 0$ we have

$$\mathbb{P}(X_{t+1} = j \mid X_t = i) = \begin{cases} \frac{1}{2} & \text{if } i = j, \\ 1/2d & \text{if } ij \in E, \\ 0 & \text{otherwise.} \end{cases}$$

Intuitively, if $X_t$ represents the probability distribution of the random walker's position at time $t$; that is, $\mathbb{P}(X_t = i)$ is the probability that she will be at vertex $i$ at time $t$. The display above says that from time $t$ to time $t + 1$ the random walker has a 50% chance of staying put, and equal chances of moving away from the current vertex along any of the edges incident on it. Since the transition probabilities are independent of $t$, the sequence $(X_0, X_1, \ldots)$ is a Markov chain.

The elementary theory of Markov processes guarantees that, with these probabilities, $\mathbb{P}(X_t = i) \to 1/n$ for all $i$ as $t \to \infty$, no matter what the initial state is (see [Kemeny and Snell 1976], for example). The question is, how fast? Set $p_i^{(t)} = \mathbb{P}(X_t = i)$ and let $e_{i,t} = p_i^{(t)} - 1/n$ be the *excess probability* at $i$. The excess probabilities satisfy

$$\begin{aligned} e_{i,\,t+1} = p_i^{(t+1)} - \frac{1}{n} &= \left(\tfrac{1}{2} p_i^{(t)} + \frac{1}{2d} \sum_{j \in \Gamma(i)} p_j^{(t)}\right) - \frac{1}{n} \\ &= \tfrac{1}{2}\bigl(p_i^{(t)} - 1/n\bigr) + \frac{1}{2d} \sum_{j \in \Gamma(i)} \bigl(p_j^{(t)} - 1/n\bigr) \\ &= \tfrac{1}{2} e_{i,t} + \frac{1}{2d} \sum_{j \in \Gamma(i)} e_{j,t} = \frac{1}{2d} \sum_{j \in \Gamma(i)} (e_{i,t} + e_{j,t})\,. \end{aligned} \tag{3.1}$$

Define

$$d_1(t) = d_1(\tilde{X}, t) = \sum_i |e_{i,t}|$$

and

$$d_2(t) = d_2(\tilde{X}, t) = \sum_i e_{i,t}^2.$$

A simple random walk $\tilde{X}$ on $G$ is *rapidly mixing* if there is a polynomial $f$ such that if $0 < \varepsilon < \frac{1}{3}$ and $t \ge f(\log n) \log(1/\varepsilon)$ then $d_1(t) \le \varepsilon$.

Strictly speaking, this definition does not make much sense since if we have only one graph $n$ itself is really a constant. For a proper definition, we need a *sequence* $(G_i)_{i=1}^{\infty}$ of regular graphs, where each $G_i$ has $n_i$ vertices and $n_i \to \infty$. We say that the simple random walks on $G_1, G_2, \ldots$ are *rapidly mixing* if there is a polynomial $f$, depending only on the sequence $(G_i)$, such that if $0 < \varepsilon < \frac{1}{3}$ and $t \ge f(\log n_i) \log(1/\varepsilon)$ then $d_1(\tilde{X}_i, t) \le \varepsilon$ whenever $\tilde{X}_i$ is a simple random walk on $G_i$.

Let's define the *conductance* of $G$ or the *conductance* of a simple random walk on $G$ as follows. For $U \subset V$ set $\bar{U} = V - U$ and

$$\Phi_G(U) = \frac{e(U, \bar{U})}{d\,|U|}.$$

Note that $0 \le \Phi_G(U) \le 1$. Also, for $1 \le |U| \le n/2$, $\Phi_G(U)$ is small if there are relatively few $U - \bar{U}$ edges, that is, if there is a "bottleneck" when we try to go from $U$ to $\bar{U}$. The *conductance* of $G$ is then

$$\Phi_G = \min_{|U| \le n/2} \Phi_G(U).$$

The conductance is also called the *isoperimetric number* of the graph or its *Cheeger constant*. The quantity $d\,|U|$ is the "volume" of $U$, the sum of the degrees of its vertices. If $G = (V, E)$ is not necessarily regular then for $U \subset V$ the *volume* of $U$ is $\operatorname{vol} U = \sum_{u \in U} d(u)$, and the conductance of $G$ is

$$\min_{U \subset V} \frac{e(U, \bar{U})}{\min\{\operatorname{vol} U, \operatorname{vol} \bar{U}\}}.$$

With this definition, the results below are easily extended to general graphs and beyond; we shall state one of these results at the end of the lecture.

Clearly, we have $0 \le \Phi_G \le 1$, although the upper bound is somewhat unrealistic: if $\Phi_G = 1$ then $G$ is either the trivial graph consisting of one vertex, or a single edge, or a triangle. If $G$ has many vertices, the best we can hope is that $\Phi_G$ is not far from $\frac{1}{2}$. Concerning the lower bound, note that $\Phi_G = 0$ if and only if $G$ is disconnected.

Our main aim is to prove the following fundamental result, which clearly shows the importance of the conductance.

THEOREM 3.1. *Every simple random walk on $G$ satisfies*

$$d_2(t+1) \le \bigl(1 - \tfrac{1}{4}\Phi_G^2\bigr) d_2(t).$$

*In particular*, *as* $d_2(0) \le 2$,

$$d_2(t) \le \bigl(1 - \tfrac{1}{4}\Phi_G^2\bigr)^t d_2(0) \le 2\bigl(1 - \tfrac{1}{4}\Phi_G^2\bigr)^t.$$

We shall deduce this result from two lemmas that are of interest in their own right.

LEMMA 3.2. $d_2(t+1) \le d_2(t) - \dfrac{1}{2d} \displaystyle\sum_{ij \in E} (e_{i,t} - e_{j,t})^2.$

PROOF. By (3.1),

$$d_2(t+1) = \frac{1}{4d^2} \sum_{i=1}^{n} \Biggl( \sum_{j \in \Gamma(i)} (e_{i,t} + e_{j,t}) \Biggr)^2.$$

Applying the Cauchy–Schwarz inequality to the inner sum, we find that, as $|\Gamma(i)| = d$,

$$\begin{aligned} d_2(t+1) &\le \frac{1}{4d^2} \sum_{i=1}^{n} \Biggl( \sum_{j \in \Gamma(i)} (e_{i,t} + e_{j,t})^2 \Biggr) d \\ &= \frac{1}{2d} \sum_{ij \in E} (e_{i,t} + e_{j,t})^2 = \frac{1}{2d} \sum_{ij \in E} \bigl\{ 2\,(e_{i,t}^2 + e_{j,t}^2) - (e_{i,t} - e_{j,t})^2 \bigr\} \\ &= d_2(t) - \frac{1}{2d} \sum_{ij \in E} (e_{i,t} - e_{j,t})^2. \end{aligned}$$

□

The second lemma needs a little more work.

LEMMA 3.3. *Suppose weights $x_i$ are assigned to the elements of the vertex set $V = \{1, \ldots, n\}$, satisfying $\sum_{i=1}^n x_i = 0$. Then*

$$\sum_{ij\in E} (x_i - x_j)^2 \geq \frac{d}{2}\Phi_G^2 \sum_{i=1}^n x_i^2.$$

PROOF. Set $m = \lceil n/2 \rceil$. We shall prove that if $y_1 \geq y_2 \geq \ldots \geq y_n$, with $y_m = 0$, then

$$\sum_{ij\in E} (y_i - y_j)^2 \geq \frac{d}{2}\Phi_G^2 \sum_{i=1}^n y_i^2. \tag{3.2}$$

This is stronger than the desired inequality. Indeed, in the statement of the lemma we may assume that $x_1 \geq x_2 \geq \ldots \geq x_n$. Setting $y_i = x_i - x_m$, inequality (3.2) gives

$$\sum_{ij\in E} (x_i - x_j)^2 = \sum_{ij\in E} (y_i - y_j)^2 \geq \frac{d}{2}\Phi_G^2 \sum_{i=1}^n (x_i - x_m)^2 = \frac{d}{2}\Phi_G^2 \sum_{i=1}^n x_i^2 + \frac{nd}{2}\Phi_G^2 x_m^2,$$

since $\sum_{i=1}^n x_i = 0$.

In order to prove (3.2), set

$$u_i = \begin{cases} y_i & \text{if } i \leq m, \\ 0 & \text{if } i > m, \end{cases} \qquad v_i = \begin{cases} 0 & \text{if } i \leq m, \\ y_i & \text{if } i > m. \end{cases}$$

Thus $y_i = u_i + v_i$ for every $i$. Also, if $u_i \neq 0$ then $u_i > 0$ and $i < m$, and if $v_i \neq 0$ then $v_i < 0$ and $i > m$. Since $(y_i - y_j)^2 = (u_i - u_j + v_i - v_j)^2 \geq (u_i - u_j)^2 + (v_i - v_j)^2$ for every edge $ij$, it suffices to prove that

$$\sum_{ij\in E} (u_i - u_j)^2 \geq \frac{d}{2}\Phi_G^2 \sum_{i=1}^m u_i^2 \tag{3.3}$$

and

$$\sum_{ij\in E} (v_i - v_j)^2 \geq \frac{d}{2}\Phi_G^2 \sum_{i=m}^n v_i^2.$$

Furthermore, as $m \geq n - m$, it suffices to prove (3.3). We may assume that $u_1 > 0$. By the Cauchy–Schwarz inequality,

$$\begin{aligned}
\left( \sum_{ij\in E} \left(u_i^2 - u_j^2\right) \right)^2 &= \left( \sum_{ij\in E} (u_i - u_j)(u_i + u_j) \right)^2 \leq \sum_{ij\in E} (u_i - u_j)^2 \sum_{ij\in E} (u_i + u_j)^2 \\
&\leq \sum_{ij\in E} (u_i - u_j)^2 \sum_{ij\in E} 2\left(u_i^2 + u_j^2\right) \\
&= 2d \sum_{i=1}^n u_i^2 \sum_{ij\in E} (u_i - u_j)^2. \qquad (3.4)
\end{aligned}$$

We may assume that, in all the sums $\sum_{ij\in E}$ over the edges $ij$, we have $i<j$. Note that

$$\sum_{ij\in E}\left(u_i^2-u_j^2\right)=\sum_{ij\in E}\sum_{l=i}^{j-1}\left(u_l^2-u_{l+1}^2\right)=\sum_{l=1}^{n-1}\left(u_l^2-u_{l+1}^2\right)e(U_l,\,\bar{U}_l),$$

where $U_l=\{1,\ldots,l\}$ and $\bar{U}_l=\{l{+}1,\ldots,n\}$. Since $u_m=u_{m+1}=\ldots=u_n=0$, this gives

$$\begin{aligned}\sum_{ij\in E}\left(u_i^2-u_j^2\right)&=\sum_{l=1}^{m-1}\left(u_l^2-u_{l+1}^2\right)e(U_l,\,\bar{U}_l)\geq\sum_{l=1}^{m-1}\left(u_l^2-u_{l+1}^2\right)d\,\Phi_G l\\&=d\Phi_G\sum_{l=1}^{m-1}u_l^2=d\Phi_G\sum_{l=1}^{n}u_l^2.\end{aligned}\tag{3.5}$$

Inequalities (3.4) and (3.5) give

$$\sum_{ij\in E}(u_i-u_j)^2\geq\left(d\Phi_G\sum_{i=1}^n u_i^2\right)^2\bigg/\left(2d\sum_{i=1}^n u_i^2\right)=\frac{d}{2}\Phi_G^2\sum_{i=1}^n u_i^2,$$

as desired. □

Proof of Theorem 3.1. By Lemma 3.2,

$$d_2(t)-d_2(t+1)\geq\frac{1}{2d}\sum_{ij\in E}(e_{i,t}-e_{j,t})^2.$$

Applying Lemma 3.3 with $x_i=e_{i,t}$, we find that

$$d_2(t)-d_2(t+1)\geq\tfrac14\Phi_G^2\sum_{i=1}^n e_{i,t}^2=\tfrac14\Phi_G^2 d_2(t),$$

completing the proof. □

By the Cauchy–Schwarz inequality, $d_1(t)\leq(n\,d_2(t))^{1/2}$, so Theorem 3.1 has the following immediate consequence.

Corollary 3.4. *Every simple random walk on a graph $G$ of order $n$ with conductance $\Phi_G$ satisfies*

$$d_1(t)\leq(2n)^{1/2}\bigl(1-\tfrac14\Phi_G^2\bigr)^{t/2}.$$ □

Corollary 3.4 implies that if we assume that $G$ connected (so that $\Phi_G>0$), that $0<\varepsilon<1/3$, and that

$$t>8\Phi_G^{-2}\bigl(\log(1/\varepsilon)+\tfrac12\log(2n)\bigr)$$

then

$$d_1(t)\leq(2n)^{1/2}\bigl(1-\tfrac14\Phi_G^2\bigr)^{t/2}<\exp\bigl(\tfrac12\log(2n)-\tfrac18\Phi_G^2 t\bigr)<\varepsilon.$$

In particular, if $n\geq 3$ and $t\geq 8\Phi_G^{-2}\log n\log(1/\varepsilon)$ then $d_1(t)<\varepsilon$. This gives us the following sufficient condition for rapid mixing.

THEOREM 3.5. *Let $(G_i)_1^\infty$ be a sequence of regular graphs with $|G_i| = n_i \to \infty$. If there is a $k \in \mathbb{N}$ such that*

$$\Phi_{G_i} \geq (\log n_i)^{-k}$$

*for sufficiently large $i$, the simple random walks on $(G_i)_1^\infty$ are rapidly mixing.*

PROOF. We have just seen that $f(x) = 8x^{2k+1}$ will do if $n_i \geq 3$. $\square$

Let's see some families of regular graphs for which we can give a good lower bound for the conductance. As a trivial example, take the complete graph $K_n$. It is immediate that $\Phi_{K_n} > \frac{1}{2}$ for $n \geq 2$, so the simple random walks on $(K_n)$ are rapidly mixing. Of course, this can be derived very simply from first principles as well.

As a less trivial example, we take the cubes $Q_1, Q_2, \ldots$, defined as follows: the vertex set of $Q_d$ is $\{0,1\}^d$, the set of sequences $x = (x_i)_1^n$, $x_i = 0$ or $1$, and two sequences joined by an edge if they differ in only one term. $Q_d$ is obviously $d$-regular, and it is easy to prove that $\Phi_{Q_d} = 1/d$. The worst bottlenecks arise between the "top" and "bottom" of $Q_n$: for $U = \{(x_i) \in Q_d : x_1 = 1\}$ and $\bar{U} = \{(x_i) \in Q_d : x_1 = 0\}$, say. Clearly, $e(U, \bar{U}) = |U| = 2^{d-1}$ so that $\Phi_{Q_n}(U) = 1/d$.

Since $\Phi_{Q_d} = 1/d = 1/\log n$, where $n = 2^d = |Q_d|$, simple random walks on $(Q_d)_1^\infty$ are rapidly mixing.

The cube $Q_d$ is just $K_2^d = K_2 \times \ldots \times K_2$, that is, the product of $d$ paths of lengths 1. Taking the product of $d$ cycles, each of length $l$, we get the torus $T_l^d$. This graph has $l^d$ vertices and it is $2d$-regular. One can show that for $G = T_{2l}^d$ we have $\Phi_G = 2/(ld)$. (Note that $T_4^d$ is just the cube $Q_{2d}$.) Hence, for a fixed value of $l$, simple random walks on $(T_{2l}^d)_{d=1}^\infty$ are rapidly mixing.

It is straightforward to extend Theorem 3.1 to aperiodic random walks. To be precise, let $V$ be a finite set and let $X$ be a random walk on $V$ with transition probabilities $p(u,v)$ such that $p(u,u) \geq \frac{1}{2}$. Suppose that $X$ is *reversible*, that is, there is a (stationary) probability distribution $\lambda$ on $V$ with $\lambda(u)p(u,v) = \lambda(v)p(v,u)$. Here it is natural to view $\lambda(u)p(u,v)$ as the *flow* from $u$ to $v$: it is the same as the flow from $v$ to $u$. Such a random walk is a straightforward generalization of a simple random walk on a regular graph discussed above; in fact, it is hardly more than that. Denoting by $\lambda(U) = \sum_{u \in U} \lambda(u)$ the "volume" of a $U \subset V$, the *conductance* of $X$ is

$$\tilde{\Phi}_X = \min_{\lambda(U) \leq 1/2} \frac{\sum_{u \in U} \sum_{v \in V \setminus U} \lambda(u)p(u,v)}{\lambda(U)}.$$

Note that this definition makes the conductance half as large as before since if $X$ is the simple random walk on a $d$-regular graph then $p(u,v) = 1/(2d)$, so $\tilde{\Phi}_X = \frac{1}{2}\Phi_G$. The fact is that for random walks this is the natural definition, while for graphs $\Phi_G$ is natural.

Needless to say, we are interested in convergence to the stationary distribution $\lambda$. To measure the distance from $\lambda$, as before, we put

$$d_2(t) = \sum_{v \in V} \bigl(p_v^{(t)} - \lambda(v)\bigr)^2.$$

Let's state then the analogue of Theorem 3.1: the proof is unchanged.

THEOREM 3.6. *With the notation above*,

$$d_2(t+1) \le \bigl(1 - \tilde{\Phi}_X^2\bigr) d_2(t)$$

*and so*

$$d_2(t) \le 2\bigl(1 - \tilde{\Phi}_X^2\bigr)^t. \qquad \square$$

This is the theorem we shall need in the next section.

## Lecture 4. Randomized Volume Algorithms

We have seen that no polynomial-time algorithm can estimate the volume of a convex body substantially better than within a factor of $n^n$. Thus, if we want our algorithm to produce a lower bound and an upper bound that are guaranteed to be valid in *every instance* and be reasonably fast, we cannot demand that the ratio of the two bounds be substantially less than $n^n$. The situation is entirely different if we allow *randomization* and do not insist that the bounds of the algorithm be valid every time, only that they be valid with *high probability.*

Estimating the volume can be viewed as a game between *Hider*, trying to "hide" the volume of a convex body, and *Seeker*, the algorithm trying to pin down the volume. In the case of a deterministic algorithm, Hider is allowed to change his mind as the game progresses: to be precise, there is no way of telling whether he changes his mind or not, as all he has to make sure is that the answers he gives remain consistent with *some* convex body. On the other hand, a randomized algorithm is applicable only if Hider is required to play an honest game, that is if he has to fix a convex body once and for all at the beginning of the game. Then Seeker may keep tossing coins in order to decide his next appeal to the oracle, and so he may come up with a randomized algorithm that gets good results fast, with probability close to 1. Seeker trades certainty for speed and efficiency, with large probability. The probability of failure should be small and independent of the body Hider chooses.

Let's assume that our body $K \subset \mathbb{R}^n$, where $n \ge 2$, is given by a well-guaranteed strong membership oracle (although, as we mentioned earlier, it is unimportant which membership oracle we take). Let $\varepsilon$ and $\eta$ be small positive numbers, say less than $\frac{1}{3}$. An *$\varepsilon$-approximation* to $\operatorname{vol} K$ is a number $\widetilde{\operatorname{vol}}\, K$ such that

$$(1-\varepsilon)\widetilde{\operatorname{vol}}\, K < \operatorname{vol} K < (1+\varepsilon) \operatorname{vol} K.$$

If all goes well, we may hope to find a *fully polynomial approximation scheme* (FPRAS) for approximating the volume of a convex body: a randomized algorithm that runs in time polynomial in $\langle K\rangle$, $1/\varepsilon$ and $\log(1/\eta)$, and with probability at least $1-\eta$ produces an $\varepsilon$-approximation to $\operatorname{vol} K$.

In 1989, Dyer, Frieze, and Kannan [1991] found precisely such an algorithm. In describing the speed of an FPRAS, it is convenient to use the "soft-$O$" notation $O^*$, one that ignores powers of $\log n$ and polynomials of $1/\varepsilon$ and $\log(1/\eta)$. In this notation, Dyer, Frieze, and Kannan produced an FPRAS running in time $O^*(n^{23})$—to be precise, in time $O(n^{23}(\log n)^5\varepsilon^{-2}(\log 1\varepsilon)(\log 1/\eta))$. With this result, the floodgates opened: Lovász and Simonovits [1990] found a $O^*(n^{16})$ algorithm, Applegate and Kannan [1991] and Lovász and Simonovits [1992] reduced the complexity to $O^*(n^{10})$, then Dyer and Frieze [1991] to $O^*(n^8)$, Lovász and Simonovits [1993] to $O^*(n^7)$, and Kannan, Lovász, and Simonovits [$\geq$ 1997] to $O^*(n^5)$.

All algorithms are modelled on the original algorithm of Dyer, Frieze and Kannan, so they use a multiphase Monte Carlo algorithm to reduce volume computation to sampling, and use random walks to sample. In order to decide when we are likely to be close to the stationary distribution, conductance is used to bound the mixing time. Finally, isoperimetric inequalities are used to bound the conductance.

In this lecture, we shall sketch one of the most beautiful of these algorithms, given in [Dyer and Frieze 1988]. We do not give nearly all the details, to avoid making the presentation too technical.

Before stating the result, let's say a few words about an obvious naive approach to estimating the volume by a randomized algorithm, which goes as follows. Place a fine grid on $K$: for example, assuming $n^2B^n_\infty \subset K \subset n^4B^n_\infty$, we may take $\mathbb{Z}^n \cap n^4B^n_\infty$. Consider a random walk on this grid, whose stationary distribution is exactly the uniform distribution on the grid. Run such a random walk long enough so that it gets close to the stationary distribution. Stop it and check whether the point is in $K$ or not. Roughly, with probability $\operatorname{vol} K/(2n^4)^n$ we should get a point of $K$, so running this walk sufficiently many times, we should be able to estimate $\operatorname{vol} K$.

All this, of course, leads nowhere, since the probability of our random walk ending in $K$ is likely to be exponentially small: it need not even be more than $n^{-n}$. Thus, to estimate it, we would have to run $O(n^n)$ walks.

The moral of all this is that we should try to estimate only rather large *ratios* of volumes. The problem is easily reduced to this, but at the expense of not knowing the shape of the larger body either. Thus what we can have is two bodies $L \subset K$, given by oracles, with $\operatorname{vol} L > \frac{1}{2}\operatorname{vol} K$, say, and our task is to estimate $\operatorname{vol} L/\operatorname{vol} K$. Now it would be good enough to estimate this ratio by running a random walk on the part of a fine grid inside $K$, with the uniform distribution being the stationary distribution. But now the problem is that we would like to define a random walk on the grid inside a body $K$ we know almost nothing about, in a way that makes its stationary distribution uniform. This is a pretty tall order.

The following beautiful idea solves our difficulty. Define on a set *larger* than the grid inside $K$ (an entire grid graph, say) a random walk with the following properties: the stationary distribution is uniform on the points belonging to $K$; the stationary distribution gives a fairly large probability to the set of points in $K$; and the walk converges to the stationary distribution fast: it is rapidly mixing. At the first sight, all this seems to be a pie in the sky, but the beauty of it all is that Dyer, Frieze, and Kannan managed to define precisely such a random walk.

Needless to say, there are a good number of technical difficulties to overcome, the most important of which is that the random walk is rapidly mixing. This is proved with the aid of isoperimetric inequalities.

After this preamble, let's state the main result of this lecture, first proved in [Dyer, Frieze, and Kannan 1991], and sketch its proof.

Theorem 4.1. *There is a fully polynomial randomized approximation scheme for the volume of a convex body given by a well-guaranteed membership oracle.*

Proof. Sketch of proof We divide the proof into seven steps, saying rather little about each. Let $K \subset \mathbb{R}^n$ be a convex body given by the strong membership oracle, with guarantees $r \geq 1$ and $R$, so that the size of the input is $\langle K \rangle = n + \lceil \log_2 r \rceil + \lceil \log_2 R \rceil$. As always, we may assume that $n$ is large.

**1. Rounding.** Let's write $B_\infty^n = \{x \in \mathbb{R}^n : |x_i| \leq 1 \text{ for every } i\}$ for the cube of side length 2 centred at the origin. There is a polynomial algorithm that replaces $K$ by its affine image (also denoted by $K$) such that

$$2n^2 B_\infty^n \subset K \subset 2n^4 B_\infty^n,$$

say. The ratio $n^2$ of the radii of the balls is rather unimportant: $n^{20}$ would do just as well. In fact, one can do much better: making use of an idea of Lenstra [1983], Applegate and Kannan [1991] showed that we can achieve

$$B_\infty^n \subset K \subset 2(n+1) B_\infty^n$$

as well. To this end, we start with a right simplex $S$ in $K$ and gradually expand it. By rescaling everything so that $S$ becomes the standard simplex $\operatorname{conv}\{0, e_1, e_2, \ldots, e_n\}$, where $(e_i)_1^n$ is the standard basis of $\mathbb{R}^n$, one can check in polynomial time whether the region $\{x \in K : |x_i| \geq 1 + 1/n^2\}$ is empty. If it is not empty, we replace $S$ by a simplex $S' \subset K$ with $\operatorname{vol} S' \geq (1 + 1/n^2) \operatorname{vol} S$, and if it is empty we terminate the process.

**2. Subdivision.** Let's place a number of cubes $C_i = r_i B^n_\infty$ between the cube $C_0 = 2n^2 B^n_\infty$ and $C_l = 2n^4 B^n_\infty$, where $l = \lceil 2(n+1)\log_2 n \rceil$: we take $r_i = \lfloor 2^{i/(n+1)} n^2 \rfloor$ for $0 \le i < l$. Also, set $K_i = C_i \cap K$, so that $K_0 = C_0$ and $K_l = K$. With

$$\alpha_i = \mathrm{vol}\, K_{i-1}/\mathrm{vol}\, K_i,$$

we have

$$\mathrm{vol}\, K = \frac{\mathrm{vol}\, K_l}{\mathrm{vol}\, K_{l-1}} \cdot \frac{\mathrm{vol}\, K_{l-1}}{\mathrm{vol}\, K_{l-2}} \cdot \cdots \cdot \frac{\mathrm{vol}\, K_1}{\mathrm{vol}\, K_0} \cdot \mathrm{vol}\, K_0 = 2^n \Big/ \prod_{i=1}^{l} \alpha_i.$$

Hence it suffices to find an approximation of each $\alpha_i$.

At the first sight, this does not seem to be much of a progress. However, since

$$K_i \subset \frac{r_i}{r_{i-1}} K_{i-1}$$

and

$$r_i/r_{i-1} \le 2^{1/(n+1)} \frac{r_{i-1}+1}{r_{i-1}} < (1 + 1/2n^2) 2^{1/(n+1)} < 2^{1/n},$$

we have

$$\alpha_i = \frac{\mathrm{vol}\, K_{i-1}}{\mathrm{vol}\, K_i} > \frac{1}{2}.$$

Thus the gain is that it suffices to approximate the proportion of the volume of a convex body $L$ in a convex body $K$, when this proportion is rather large. In other words, we do not have to search for an exponentially small body inside another body, only for a body taking up at least half of the volume.

**3. Density.** With a slight abuse of notation, we set $K = K_i$ and $L = K_{i-1}$, where $1 \le i \le l$. Thus

$$K_0 = 2n^2 B^n_\infty \subset L \subset K \subset K_l = wn^4 B^n_\infty$$

and

$$L \subset K \subset 2^{1/n} L.$$

Let $V$ be the set of lattice points $\mathbb{Z}^n$ in $K_l$ and let $G = (V, E)$ be the subgraph of $\mathbb{Z}^n$ induced by $V$. Then $G$ is the *grid graph* $P^n_m$, with $m = 4n^4 + 1$, having $m^n$ vertices: the product of $n$ paths, each of length $4n^4$ and so with $m = 4n^4 + 1$ vertices.

We shall define a distribution on $V$ that will turn out to be the stationary distribution of a certain random walk on the graph $G$.

First, for $x \in \mathbb{R}^n$ set

$$\varphi_0(x) = \min\left\{ s \ge 0 : n^2 x \in \left(n^2 + s\right) K \right\}.$$

Clearly, $\varphi_0$ is a convex function that varies at most 1 on points at distance at most 1 in $\|\cdot\|_\infty$: if $\|x - y\|_\infty \le 1$ then $n^2(x-y) \in n^2 B^n_\infty \subset K$, so $n^2 x = n^2 y + n^2(x-y) \in (n^2 + \varphi_0(y))K + K = (n^2 + \varphi_0(y) + 1)K$, giving $\varphi_0(x) \le \varphi_0(y) + 1$.

For $x \in V = \mathbb{Z}^n \cap K_l$, set $\varphi(x) = \lceil \varphi_0(x) \rceil$, and let $\varphi_1(x)$ be the maximal convex function on $K_l$ dominated by $\varphi$. Then

$$\varphi_0(x) \le \varphi_1(x) \le \varphi(x) < \varphi_0(x) + 1$$

for $x \in V$. Finally, for $x \in K_l$, set

$$f(x) = 2^{-\varphi(x)} \qquad \text{and} \qquad f_1(x) = 2^{-\varphi_1(x)}.$$

Then

$$\tfrac{1}{2} f_1(x) < f(x) \le f_1(x)$$

and $f(x) = 1$ on $K$.

Our aim is to define a random walk on the grid graph $G$ whose stationary distribution is precisely $f$, suitably normalized.

**4. The random walk.** Let's define a random walk on the grid graph $G = (V, E)$ by giving the transition probabilities:

$$p(x,y) = \begin{cases} 1/(4n) & \text{if } xy \in E \text{ and } \varphi(y) \le \varphi(x), \\ 1/(8n) & \text{if } xy \in E \text{ and } \varphi(y) = \varphi(x) + 1, \\ 1 - \sum_{z \in \Gamma(x)} p(x,z) & \text{if } y = z, \\ 0 & \text{otherwise.} \end{cases}$$

Thus for every $x \in V$ our random walk stays put at $x$ with probability at least $\frac{1}{2}$; with probability $1/4n$ it goes to a neighbouring vertex $y$ of "norm" $\varphi(y) \le \varphi(x)$, and with *half as much* probability it goes to a neighbour of larger "norm" $\varphi(y)$.

This random walk is *reversible*, with stationary distribution $\lambda(x) = cf(x)$, where $c > 0$ is a normalizing constant. $\big($Thus $c \sum_{x \in V} f(x) = 1.\big)$Indeed, if $xy \in E$ and $\varphi(x) = \varphi(y)$ then

$$f(x)p(x,y) = 2^{-\varphi(x)} \frac{1}{4n} = 2^{-\varphi(y)} \frac{1}{4n} = f(y)p(y,x),$$

and if $\varphi(x) + 1 = \varphi(y)$ then

$$f(x)p(x,y) = 2^{-\varphi(x)} \frac{1}{8n} = 2^{-\varphi(y)} \frac{1}{4n} = f(y)p(y,x).$$

Another very important aspect of this random walk is that, although it has been tailored for $K$, it is very efficient to compute the transition probabilities at the points *where we need it.* (We certainly cannot afford to compute the transition probabilities at *all* the vertices! That would need exponentially many steps.) All we have to do is to *carry the value of* $\varphi$: as $\varphi$ is known to change by at most one at the next step, at most $4n$ appeals to the oracle give us the values of $\varphi$ at all the neighbours. Having got these values, we know all the transition probabilities from our point, so we can take the next step of our random walk. To keep things simple, we start from a point of $K$, say from $O \in n^2 B_\infty^n \subset K$; then $\varphi(O) = O$, and we are away.

**5. The error term.** For the probability distribution $\lambda$ on $V$, we have

$$\lambda(K) = \lambda(V \cap K) > \frac{1}{2}. \tag{4.1}$$

In other words, running our random walk long enough, the probability of ending in $K$ is more than $\frac{1}{2}$. Indeed,

$$f(K) = \sum_{x \in K \cap V} f(x) = |K \cap V| \sim \operatorname{vol} K.$$

Also, if $x \in \mathbb{Z}^n$ and for a positive integer $s$ we have

$$x \in \left\{ \Big(1 + \frac{s}{n^2}\Big) K - \Big(1 + \frac{s-1}{n^2}\Big) K \right\}$$

then $s - 1 < \varphi_0(x) \le \varphi(x) \le s$, so $f(x) = 2^{-s}$. The number of lattice points satisfying (4.1) is about the volume of the body on the right-hand side, so about

$$\left\{ \Big(1 + \frac{s}{n^2}\Big)^n - \Big(1 + \frac{s-1}{n^2}\Big)^n \right\} \operatorname{vol} K,$$

so it is certainly at most

$$\big(e^{s/n} - 1\big) f(K).$$

Hence,

$$f(\mathbb{Z}^n) = \sum_{x \in \mathbb{Z}^n} f(x) \le f(K) + \sum_{s=1}^{\infty} 2^{-s} \big(e^{s/n} - 1\big) f(K) < 2 f(K).$$

But $\lambda$ is just $cf$, so

$$\lambda(K) = \lambda(K)/\lambda(\mathbb{Z}^n) = f(K)/f(\mathbb{Z}^n) > \tfrac{1}{2},$$

proving the claim.

**6. A coin toss.** Set $\alpha = \operatorname{vol} L / \operatorname{vol} K$ and $\alpha' = \lambda(L \cap V)/\lambda(K \cap V) = \lambda(L)/\lambda(K)$. Then $\alpha'$ is sufficiently close to $\alpha$, so all we have to do is estimate $\alpha'$. This will be done by tossing a biased coin, with probability about $\alpha'$ for heads. Here is how one coin toss works.

We run our random walk long enough, till it is close enough to the stationary distribution. Say, we stop our random walk $X_0 = 0, X_1, \ldots$ at $X_t \in \mathbb{Z}^d$. Let $E_0$ be the event that $X_t \notin K$, $E_1$ the event that $X_t \in L$ and $E_2$ the event that $X_t \in K \setminus L$. Then $\mathbb{P}(E_0)$ is not too large: by (4.1), it is not much larger than $\frac{1}{2}$, so $\mathbb{P}(E_1 \cup E_2)$ is substantial, at least $\frac{1}{3}$. Since, on the lattice points of $K$, the stationary distribution $\lambda$ is uniform, we have

$$\mathbb{P}(E_2)/\mathbb{P}(E_1 \cup E_2) \sim \alpha',$$

with good enough approximation.

By repeating the coin toss sufficiently many times, our approximation will be good enough with high enough probability.

**7. The crunch.** What we have to show now is that it suffices to run our random walk for polynomial time to get close to its stationary distribution—in short, that our random walk is rapidly mixing. By Theorem 3.6, all we need is to show that our random walk has large enough conductance. That this is the case, and so Theorem 4.1 holds, follows from the following *isoperimetric inequality*, essentially due to Lovász and Simonovits[Lovász and Simonovits 1990].

THEOREM 4.2. *Let $M \subset \mathbb{R}^n$ be a convex body and let $\mathcal{B}(M)$ be the σ-field of Borel subsets of $M$. Let $F : \operatorname{Int} M \to \mathbb{R}^+$ be a log-concave function and let $\mu$ be the measure on $\mathcal{B}(M)$ with density $F$:*

$$\mu(A) = \int_A F\,dx$$

*for $A \in \mathcal{B}(M)$. Then, for $A_1, A_2 \in \mathcal{B}(M)$, we have*

$$\min\{\mu(A_1), \mu(A_2)\} \le \tfrac{1}{2}\,\frac{\operatorname{diam} M}{d(A_1, A_2)}\,\mu(M \setminus A_1 \cup A_2),$$

*where* $\operatorname{diam} M = \max\{|x-y| : x, y \in M\}$ *is the diameter of $M$ and $d(A_1, A_2) = \inf\{|x-y| : x \in A_1,\ y \in A_2\}$ is the distance between $A_1$ and $A_2$.* □

The proof uses repeated bisections and is somewhat similar to the method of [Payne and Weinberger 1960]. Lovász and Simonovits [1990] proved the inequality with a constant 1 instead of the best possible constant $\frac{1}{2}$, which was inserted in [Dyer and Frieze 1991].

Let's see then that Theorem 4.2 implies that the conductance of our random walk is not too small. Let $U \subset V$, $0 < \lambda(U) < \frac{1}{2}$, $\bar{U} = V \setminus U$, and let $\partial U$ be the boundary of $U$, that is, the set of vertices in $\bar{U}$ having at least one neighbour in $U$. Let $M$ be the union of unit cubes centred at the points of $V$, so that $M$ is a solid cube. Let $A_1$ be the union of unit cubes centred at the vertices of $U$, let $B$ be the union of cubes of volume 2 centred at the vertices of $\partial U$, and set $A_2 = M \setminus (A_1 \cup B)$.

Writing $c_1, c_2, \ldots$ for positive constants, we clearly have

$$d(A_1, A_2) \ge c_1/n$$

and

$$\sum_{u\in U}\sum_{v\in\bar{U}} \lambda(u)p(u,v) \ge \frac{c_2}{n}\lambda(B).$$

Hence we may assume that $\lambda(B)$ is small, say $\lambda(B) < 1/n$.

Define a measure $\mu$ on $\mathcal{B}(M)$, as in Theorem 4.2, with $F = f_1 = 2^{-\varphi_1}$, where $\varphi_1$ is the maximal convex function on $M$ dominated by $\varphi$. Note that, for every $u \in V$, $\lambda(u)$ is within a constant factor of the $\mu$-measure of the unit cube centred at $u$. Hence, by Theorem 4.2,

$$\frac{\sum_{u\in U}\sum_{v\in\bar{U}} \lambda(u)p(u,v)}{\lambda(U)} \ge \frac{c_3}{n}\,\frac{\mu(M \setminus A_1 \cup A_2)}{\min\{\mu(A_1), \mu(A_2)\}} \ge n^{-7},$$

since, rather crudely, $\operatorname{diam} M = O(n^5)$ and $d(A_1, A_2) \geq c_1/n$. This completes the sketch of a proof of Theorem 4.1.

We have made no attempt to get a really fast algorithm: in particular, one could use an isoperimetric inequality closer to the problem at hand than Theorem 4.2. Using a more careful analysis, Dyer and Frieze [1991] showed that the running time of the algorithm above is $O^*(n^8)$.

We conclude with a few words about the latest results concerning FRPAS for computing the volume. Improving the earlier results, Kannan, Lovász and Simonovits [1995] proved the following theorem.

THEOREM 4.3. *Let $c > 0$ be a constant and let $K \subset \mathbb{R}^n$ be a convex body given by a separation oracle, with guarantee*

$$B^n \subset K \subset n^c B^n.$$

*There is a fully randomized polynomial approximation scheme that, given $\varepsilon, \eta > 0$, returns positive numbers $\underline{\mathrm{vol}}\, K$ and $\overline{\mathrm{vol}}\, K$ such that $\underline{\mathrm{vol}}\, K \leq (1+\varepsilon)\overline{\mathrm{vol}}\, K$ and*

$$\underline{\mathrm{vol}}\, K \leq \mathrm{vol}\, K \leq \overline{\mathrm{vol}}\, K$$

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

$$O\Big(\frac{n^5}{\varepsilon^2}\Big(\log\frac{1}{\varepsilon}\Big)^3\Big(\log\frac{1}{\eta}\Big)(\log n)^5\Big) = O^*\Big(n^5\Big)$$

*calls to the oracle.*

The basis of their proof is, once again, a fast sampling algorithm, that is, a fast algorithm that generates $N$ random points $v_1, v_2, \ldots, v_N$ of $K$ with *almost uniform* and *almost independent* distributions. To be precise, if

$$B^n \subset K \subset dB^n \tag{4.2}$$

then we can achieve that

(a) the distribution of each $v_i$ is close to the uniform distribution in the total variation distance: for $U \in \mathcal{B}(K)$ we have

$$|\mathbb{P}(v_i \in U) - \mathrm{vol}\, U/\mathrm{vol}\, K| < \varepsilon;$$

(b) for $1 \leq i < j \leq N$ and $A, B \in \mathcal{B}(K)$ we have

$$|\mathbb{P}(v_i \in A,\, v_j \in B) - \mathbb{P}(V_i \in A)\mathbb{P}(V_j \in B)| < \varepsilon;$$

(c) the algorithm uses only $O^*(n^3 d^2 + N n^2 d^2)$ calls to the oracle.

The main innovation in finding such an algorithm is that instead of demanding (4.2), Kannan, Lovász and Simonovits [$\geq$ 1997] are satisfied with 'approximate sandwiching', that is, if $B^n$ is contained in $K$ and $d'B^n \cap K$ is *most of* $K$, provided $d'$ is small and $K$ can be 'turned' into such a position by a fast algorithm. Thus we want an efficient way of finding an affine transformation $T$ such that $B^n \subset TK$ and $d'B^n \cap TK$ is most of $TK$ for $d'$ fairly small.

It is easy to show [Pisier 1989; Ball 1997] that (4.2) cannot be guaranteed with $d < n$; and it is not known whether in polynomial time one can guarantee it with $d$ not much larger than $n$. However, if we demand only approximate sandwiching then we *can* find an affine transformation $T$ in $O^*(n^5)$ time and $d' = O(\sqrt{n}/\log(1/\varepsilon))$. After much work, this leads to a volume algorithm with $O^*(n)$ calls to the oracle.

Finally, having emphasized how surprising it is that there are fully randomized polynomial time algorithms to approximate the volume of a convex body, let us note that there seems to be no nontrivial *lower bound* on the speed of such an algorithm. For example, it is not impossible that there are algorithms running in time $O^*(n^2)$.

## References

[Aldous 1987] Aldous, D. (1987). On the markov chain simulation method for uniform combinatorial distributions and simulated annealing. *Probab. Engrg. Inform. Sci. 1*, 33–46.

[Alon 1986] Alon, N. (1986). Eigenvalues and expanders. *Combinatorica 6*, 86–96.

[Alon and Milman 1985] Alon, N. and V. D. Milman (1985). $\lambda_1$, isoperimetric inequalities for graphs and superconcentrators. *J. Combin. Theory Ser. B 38*, 73–88.

[Applegate and Kannan 1991] Applegate, D. L. and R. Kannan (1991). Sampling and integration of near log-concave functions. In *Proc. 23rd Annual ACM Symposium on Theory of Computing*, New York, pp. 156–163. ACM.

[Ball 1986] Ball, K. M. (1986). Cube slicing in $\setminus^n$. *Proc. Amer. Math. Soc. 97*, 465–473.

[Ball 1997] Ball, K. M. (1997). An elementary introduction to modern convex geometry. In S. Levy (Ed.), *Flavors of Geometry*, Number 31 in MSRI Publications, pp. 1–58. New York: Cambridge U. Press.

[Ball and Pajor 1990] Ball, K. M. and A. Pajor (1990). Convex bodies with few faces. *Proc. Amer. Math. Soc. 110*, 225–231.

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

[Bárány and Füredi 1987] Bárány, I. and Z. Füredi (1987). Computing the volume is difficult. *Discrete Comput. Geom. 2*, 319–326.

[Bárány and Füredi 1988] Bárány, I. and Z. Füredi (1988). Approximation of the sphere by polytopes having few vertices. *Proc. Amer. Math. Soc. 102*, 651–659.

[Bollobás 1979] Bollobás, B. (1979). *Graph theory: an introductory course*. Number 63 in Graduate Texts in Math. Heidelberg: Springer.

[Bombieri and Vaaler 1983] Bombieri, E. and J. D. Vaaler (1983). On siegel's lemma. *Invent. Math. 73*, 11–32.

[Bourgain et al. 1989] Bourgain, J., J. Lindenstrauss, and V. D. Milman (1989). Approximation of zonoids by zonotopes. *Acta Math. 162*, 73–141.

[Bourgain and Milman 1985] Bourgain, J. and V. D. Milman (1985). Sections euclidiennes et volume des corps symétriques convexes dans $\diagdown^n$. *C. R. Acad. Sci. Paris Sér. I 300*, 435–437.

[Carathéodory 1907] Carathéodory, C. (1907). Über den variabilitätsbereich der koeffizienten von potenzreihen, die gegebene werte nicht annehmen. *Math. Ann. 64*, 95–115.

[Carl 1985] Carl, B. (1985). Inequalities of Bernstein–Jackson-type and the degree of compactness of operators in Banach spaces. *Ann. Inst. Fourier* (*Grenoble*) *35*(3), 79–118.

[Carl and Pajor 1988] Carl, B. and A. Pajor (1988). Gel'fand numbers of operators with values in a Hilbert space. *Invent. Math. 94*, 479–504.

[Cheeger 1970] Cheeger, J. (1970). A lower bound for the smallest eigenvalue of the laplacian. In R. C. Gunning (Ed.), *Problems in analysis*, pp. 195–199. Princeton: Princeton U. Press.

[Chung 1996] Chung, F. R. K. (1996). *Spectral graph theory.* CBMS Lecture Notes. Providence: Amer. Math. Soc.

[Dyer and Frieze 1988] Dyer, M. E. and A. M. Frieze (1988). On the complexity of computing the volume of a polyhedron. *SIAM J. Comput. 17*, 967–974.

[Dyer and Frieze 1991] Dyer, M. E. and A. M. Frieze (1991). Computing the volume of convex bodies: a case where randomness provably helps. In B. Bollobás (Ed.), *Probabilistic combinatorics and its applications*, Number 44 in Proc. Symp. Applied Math., Providence, pp. 123–169. Amer. Math. Soc.

[Dyer et al. 1991] Dyer, M. E., A. M. Frieze, and R. Kannan (1991). A random polynomial time algorithm for approximating the volume of convex bodies. *J. ACM 38*, 1–17.

[Eckhoff 1993] Eckhoff, J. (1993). Helly, Radon, and Carathéodory type theorems. In P. M. Gruber and J. M. Wills (Eds.), *Handbook of convex geometry A*, pp. 389–448. Amsterdam: North-Holland.

[Elekes 1986] Elekes, G. (1986). A geometric inequality and the complexity of measuring the volume. *Discrete Comput. Geom. 1*, 289–292.

[Fejes Tóth 1964] Fejes Tóth, L. (1964). *Regular Figures.* Number 48 in Monographs in Pure and Applied Math. New York: MacMillan.

[Figiel and Johnson 1980] Figiel, T. and W. B. Johnson (1980). Large subspaces of $l_\infty^n$ and estimates of the gordon–lewis constants. *Israel J. Math. 37*, 92–112.

[Gluskin 1988] Gluskin, E. D. (1988). Extremal properties of rectangular parallelipipeds and their applications to the geometry of banach spaces. *Mat. Sb.* (*N. S.*) *136*, 85–95. In Russian.

[Gordon et al. ] Gordon, Y., S. Reisner, and C. Schütt. Umbrellas and polytopal approximation of the euclidean ball. To appear.

[Grötschel et al. 1988] Grötschel, M., L. Lovász, and A. Schrijver (1988). *Geometric algorithms and combinatorial optimization.* Number 2 in Algorithms and combinatorics. Berlin: Springer.

[Hensley 1979] Hensley, D. (1979). Slicing the cube in $\diagdown^n$ and probability (bounds for the measure of a central cube slice in $\diagdown^n$ by probability methods). *Proc. Amer. Math. Soc. 73*, 95–100.

[Jerrum and Sinclair 1989] Jerrum, M. R. and A. J. Sinclair (1989). Approximating the permanent. *SIAM J. Comput. 18*, 1149–1178.

[John 1948] John, F. (1948). Extremum problems with inequalities as subsidiary conditions. In *Studies and essays presented to R. Courant on his* 60*th birthday*, pp. 187–204. New York: Interscience.

[Kemeny and Snell 1976] Kemeny, J. G. and J. L. Snell (1976). *Finite Markov chains*. Undergraduate Texts in Mathematics. New York: Springer. Reprinting of the 1960 original.

[Khachiyan 1988] Khachiyan, L. G. (1988). On the complexity of computing the volume of a polytope. *Izvestia Akad. Nauk SSSR Tekhn. Kibernet. 3*, 216–217.

[Khachiyan 1989] Khachiyan, L. G. (1989). The problem of computing the volume of polytopes is np-hard. *Uspekhi Mat. Nauk 44*, 179–180. In Russian; translation in *Russian Math. Surveys* **44** (1989), no. 3, 199–200.

[Khachiyan 1993] Khachiyan, L. G. (1993). Complexity of polytope volume computation. In J. Pach (Ed.), *New trends in discrete and computational geometry*, Number 10 in Algorithms and Combinatorics, pp. 91–101. Springer.

[Lawrence 1991] Lawrence, J. (1991). Polytope volume computation. *Math. Comp. 57*(195), 259–271.

[Lenstra 1983] Lenstra, H. W. (1983). Integer programming with a fixed number of variables. *Math. Oper. Res. 8*, 538–548.

[Lovász 1986] Lovász, L. (1986). *An algorithmic theory of numbers, graphs and convexity*. Number 50 in CBMS–NSF Regional Conference Series in Applied Mathematics. Philadelphia: SIAM.

[Lovász and Simonovits 1990] Lovász, L. and M. Simonovits (1990). The mixing rate of markov chains, an isoperimetric inequality, and computing the volume. In 31*st Annual Symposium on Foundations of Computer Science*, Los Alamitos, CA, pp. 346–354. IEEE Comput. Soc. Press.

[Lovász and Simonovits 1992] Lovász, L. and M. Simonovits (1992). On the randomized complexity of volume and diameter. In 33*st Annual Symposium on Foundations of Computer Science*, Los Alamitos, CA, pp. 482–491. IEEE Comput. Soc. Press.

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

[Mihail 1989] Mihail, M. (1989). Conductance and convergence of Markov chains—a combinatorial treatment of expanders. In 30*th Annual Symposium on Foundations of Computer Science*, Los Alamitos, CA, pp. 526–531. IEEE Comput. Soc. Press.

[Payne and Weinberger 1960] Payne, L. E. and H. F. Weinberger (1960). An optimal poincaré inequality for convex domains. *Arch. Rational Mech. Anal. 5*, 286–292.

[Pisier 1989] Pisier, G. (1989). *The volume of convex bodies and Banach space geometry*. Cambridge Tracts in Mathematics, 94. Cambridge: Cambridge University Press.

[Ravi Kannan and Simonovits ] Ravi Kannan, L. L. and M. Simonovits. Random walks and an $o^*(n^5)$ volume algorithm for convex bodies. to appear.

[Santaló 1949] Santaló, L. (1949). Un invariante afin para los cuerpos convexos del espacio de $n$ dimensiones.

[Sinclair and Jerrum 1989] Sinclair, A. J. and M. R. Jerrum (1989). Approximate counting, uniform generation and rapidly mixing markov chains. *Inform. and Comput. 82*, 93–133.

[Vaaler 1979] Vaaler, J. D. (1979). A geometric inequality with applications to linear forms. *Pacific J. Math. 83*, 543–553.

[Vazirani 1991] Vazirani, U. (1991). Rapidly mixing markov chains. In B. Bollobás (Ed.), *Probabilistic Combinatorics and Its Applications*, Number 44 in Proc. Symp. Applied Math., Providence, pp. 99–121. Amer. Math. Soc.

## Index

Béla Bollobás
Department of Mathematics
University of Memphis
Institute for Advanced Study
bollobab@ibex.msci.memphis.edu