Dispersion of Mass and the Complexity of Geometric Problems

by

Luis Alexis Rademacher

Licenciado en Ciencias de la Ingeniería
Universidad de Chile (2002)

Ingeniero Civil Matemático
Universidad de Chile (2002)

Submitted to the Department of Mathematics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2007

Author . . . .
Department of Mathematics
May 1, 2007

Certified by . . . .
Santosh S. Vempala
Associate Professor of Applied Mathematics
Thesis Supervisor

Accepted by . . . .
Alar Toomre
Chairman, Applied Mathematics Committee

Accepted by . . . .
Pavel I. Etingof
Chairman, Department Committee on Graduate Students

Dispersion of Mass and the Complexity of Geometric Problems

by

Luis Alexis Rademacher

Submitted to the Department of Mathematics
on May 1, 2007, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

How much can randomness help computation? Motivated by this general question and by volume computation, one of the few instances where randomness provably helps, we analyze a notion of dispersion and connect it to asymptotic convex geometry. We obtain a nearly quadratic lower bound on the complexity of randomized volume algorithms for convex bodies in $\mathbb{R}^n$ (the current best algorithm has complexity roughly n^4, conjectured to be n^3). Our main tools, dispersion of random determinants and dispersion of the length of a random point from a convex body, are of independent interest and applicable more generally; in particular, the latter is closely related to the variance hypothesis from convex geometry. This geometric dispersion also leads to lower bounds for matrix problems and property testing.

We also consider the problem of computing the centroid of a convex body in $\mathbb{R}^n$. We prove that if the body is a polytope given as an intersection of half-spaces, then computing the centroid exactly is $\#P$-hard, even for order polytopes, a special case of 0–1 polytopes. We also prove that if the body is given by a membership oracle, then for any deterministic algorithm that makes a polynomial number of queries there exists a body satisfying a roundedness condition such that the output of the algorithm is outside a ball of radius $\sigma/100$ around the centroid, where σ^2 is the minimum eigenvalue of the inertia matrix of the body.

Finally, we consider the problem of determining whether a given set S in $\mathbb{R}^n$ is *approximately* convex, i.e., if there is a convex set $K \subseteq \mathbb{R}^n$ such that the volume of their symmetric difference is at most $\epsilon \operatorname{vol}(S)$ for some given ϵ. When the set is presented only by a membership oracle and a random oracle, we show that the problem can be solved with high probability using $\operatorname{poly}(n)(c/\epsilon)^n$ oracle calls and computation time. We complement this result with an exponential lower bound for the natural algorithm that tests convexity along "random" lines. We conjecture that a simple 2-dimensional version of this algorithm has polynomial complexity.

Thesis Supervisor: Santosh S. Vempala
Title: Associate Professor of Applied Mathematics

Acknowledgments

I would like to thank those that made this work possible.

To Santosh Vempala, my advisor, as this work is the result of our collaboration. To Madhu Sudan and Peter Shor, for being part of my thesis committee. To the Department of Mathematics and CSAIL at MIT and their members, for providing an environment that strongly stimulates research.

To my friends Amit Deshpande and Roberto Rondanelli, for interesting discussions around this work.

To my grandparents, father and mother, for the love and support that they showed.

Contents

List of Figures

Notation

This is a list of symbols used frequently with their descriptions and, in parentheses, the numbers of the pages in which they are introduced.

R_i	ith row (15)
R_{-i}	all rows but ith (15)
$\hat{R}$	normalized rows (15)
conv	convex hull (15)
vol	volume (15)
D, D'	distributions (17)
Q	membership oracle (16)
Q'	modified oracle (17)
$\mathrm{disp}_\mu(p)$	dispersion of a distribution μ (30)
$\mathcal{K}$	convex bodies and $\varnothing$ (60)
ϵ-convex	(61)

Introduction

Among the most intriguing questions raised by complexity theory is the following: how much can the use of randomness affect the computational complexity of algorithmic problems? At the present time, there are many problems for which randomized algorithms are simpler or faster than known deterministic algorithms but only a few known instances where randomness provably helps. As we will see, one of these known instances is the geometric problem of computing the volume of a convex body in $\mathbb{R}^n$ given by a membership oracle. One of the results of this work is another example where randomness provably helps: the problem of computing the centroid of a convex body.

The best known algorithm for approximating the volume of a convex body in $\mathbb{R}^n$ has query complexity $O(n^4)$ [29]. The second and main contribution of this work is a lower bound of $\Omega(n^2/\log n)$ for the complexity of approximating the volume. This lower bound actually holds for parallelotopes of the form $\{x : \|Ax\|_\infty \leq 1\}$, for an $n \times n$ matrix A. As the volume of such a parallelotope is proportional to $1/|\det A|$, what we actually give is a lower bound for the problem of approximating $|\det A|$ when A is accessed through an oracle that given $q \in \mathbb{R}^n$ decides whether $\|Aq\|_\infty \leq 1$. We also give similar lower bounds for the problem of approximating the product of the lengths of the rows of A, and approximating the length of a vector.

To prove this lower bounds we introduce a measure of dispersion of a distribution that models the probability of failure of an algorithm against a distribution on inputs. The computational lower bounds are then a consequence of two dispersion lower bounds that we prove: dispersion of the determinant of certain distributions on matrices, and dispersion of the length of a random vector from a polytope. This last results has interesting connections with an important open problem in asymptotic convex geometry, the variance hypothesis.

Finally, we study the problem of testing whether a set given by a membership oracle and a random oracle is approximately convex. We give an exponential time algorithm and we prove that a very natural algorithm that checks the convexity of the input set along random lines has exponential complexity.

The thesis is organized as follows:

Chapter 2 introduces some notation and basic results that will be used in the rest of this work.

Chapter 3 gives known deterministic lower bounds for the complexity of the volume and gives our deterministic lower bounds for the complexity of the centroid.

Chapter 4 defines our notion of dispersion of a distribution, proves a few basic properties of it and then proceeds to prove our two main dispersion results: dispersion of the length of a random vector from a polytope and dispersion of the determinant of a random matrix.

Chapter 5 uses the dispersion of the determinant to prove our lower bound for the complexity of approximating the volume.

Chapter 6 uses the dispersion of a random vector from a polytope to prove our other two randomized lower bounds, for the length of a vector and the product of the rows of a matrix.

Chapter 7 presents our results about the complexity of testing a set for convexity.

Chapter 1

Preliminaries

1.1 Notation and definitions

Let $S \subseteq \mathbb{R}^n$. If S is measurable, $\text{vol}(S)$ denotes the volume of S. The convex hull of S is denoted $\text{conv}(S)$.

Let $x \cdot y = \sum_{i=1}^n x_i y_i$, the usual inner product in $\mathbb{R}^n$.

A *parallelotope* is any set that results from an affine transformation of a hypercube. Generally we will consider parallelotopes of the form $\{x \in \mathbb{R}^n : \|Ax\|_\infty \leq 1\}$ specified by an $n \times n$ matrix A. $N(0,1)$ denotes the standard normal distribution, with mean 0 and variance 1.

The n-dimensional ball of radius 1 centered at the origin is denoted B_n. We define $\pi_V(u)$ to be the projection of a vector u to a subspace V. Given a matrix R, let R_i denote the i'th row of R, and let $\hat{R}$ be the matrix having the rows of R normalized to be unit vectors. Let $\tilde{R}_i$ be the projection of R_i to the subspace orthogonal to $R_1, \ldots, R_{i-1}$. For any row R_i of matrix R, let R_{-i} denote (the span of) all rows except R_i. So $\pi_{R_{-i}^\perp}(R_i)$ is the projection of R_i orthogonal to the subspace spanned by all the other rows of R.

1.2 Yao's lemma

We will need the following version of Yao's lemma. Informally, the probability of failure of a randomized algorithm ν on the worst input is at least the probability of failure of the best deterministic algorithm against some distribution μ.

Lemma 1.1. *Let μ be a probability measure on a set I (a "distribution on inputs") and let ν be a probability measure on a set of deterministic algorithms A (a "randomized algorithm"). Then*

$$\inf_{a \in A} \Pr(\textit{alg. } a \textit{ fails on measure } \mu) \leq \sup_{i \in I} \Pr(\textit{randomized alg. } \nu \textit{ fails on input } i).$$

The proof will use the following lemma and notation.

Let I be a set (a subset of the inputs of a computational problem, for example the set of all well-rounded convex bodies in $\mathbb{R}^n$ for some n). Let O be another set (the set of possible outputs of a computational problem, for example, real numbers that are an approximation to the volume of a convex body). Let A be a set of functions from I to O (these functions represent deterministic algorithms that take elements in I as inputs and have outputs in O). Let $C : I \times A \to \mathbb{R}$ (for $a \in A$ and $i \in I$, $C(i,a)$ is a measure of the badness of the algorithm a on input i, such as the indicator of a giving a wrong answer on i).

Lemma 1.2. *Let μ and ν be probability measures over I and A, respectively. Let $C : I \times A \to \mathbb{R}$ be integrable with respect to $\mu \times \nu$. Then*

$$\inf_{a\in A} \mathbb{E}_{\mu(i)} C(i,a) \leq \sup_{i\in I} \mathbb{E}_{\nu(a)} C(i,a)$$

Proof. By means of Fubini's theorem and the integrability assumption we have

$$\mathbb{E}_{\nu(a)} \mathbb{E}_{\mu(i)} C(i,a) = \mathbb{E}_{\mu(i)} \mathbb{E}_{\nu(a)} C(i,a).$$

Also

$$\mathbb{E}_{\nu(a)} \mathbb{E}_{\mu(i)} C(i,a) \geq \inf_{a\in A} \mathbb{E}_{\mu(i)} C(i,a)$$

and

$$\mathbb{E}_{\mu(i)} \mathbb{E}_{\nu(a)} C(i,a) \leq \sup_{i\in I} \mathbb{E}_{\nu(a)} C(i,a).$$

□

Proof (of Lemma 1.1). Let $C : I \times A \to \mathbb{R}$, where for $i \in I$, $a \in A$ we have

$$C(i,a) = \begin{cases} 1 & \text{if } a \text{ fails on } i \\ 0 & \text{otherwise.} \end{cases}$$

Then the consequence of Lemma 1.2 for this C is precisely what we want to prove. □

1.3 The query model and decision trees

We will denote by Q our standard query model: A *membership oracle* for a set $K \in \mathbb{R}^n$ takes any $q \in \mathbb{R}^n$ and outputs YES if $q \in K$ and NO otherwise. When K is a parallelotope of the form $\{x \in \mathbb{R}^n : \|Ax\|_\infty \leq 1\}$ specified by an $n \times n$ matrix A, the oracle outputs YES if $\|Aq\|_\infty \leq 1$ and NO otherwise.

It is useful to view the computation of a deterministic algorithm with an oracle as a decision tree representing the sequence of queries: the nodes (except the leaves) represent queries, the root is the first query made by the algorithm and there is one query subtree per answer. The leaves do not represent queries but instead the answers to the last query along every path. Any leaf l has a set P_l of inputs that are consistent with the corresponding path of queries and answers on the tree. Thus the set of inputs is partitioned by the leaves.

To prove our main lower bounds for the complexity of approximating the volume of parallelotopes, it will be convenient to consider a modified query model Q' that can output more information: Given $q \in \mathbb{R}^n$, the modified oracle outputs YES as before if $\|Aq\|_\infty \leq 1$; otherwise it outputs a pair (i, s) where i is the "least index among violated constraints", $i = \min\{j : |A_j q| > 1\}$, and $s \in \{-1, 1\}$ is the "side", $s = \operatorname{sign}(A_i q)$. An answer from Q' gives at least as much information as the respective answer from Q, and this implies that a lower bound for algorithms with access to Q' is also a lower bound for algorithms with access to Q. The following definition and lemma explain the advantage of Q' over Q.

Definition 1.3. *Let $\mathcal{M}$ be a set of $n \times n$ matrices. We say that $\mathcal{M}$ is a* product set along rows *if there exist sets $\mathcal{M}_i \subseteq \mathbb{R}^n$, $1 \leq i \leq n$,*

$$\mathcal{M} = \{M : \forall 1 \leq i \leq n, M_i \in \mathcal{M}_i\}.$$

Lemma 1.4. *If the set of inputs is a product set along rows, then the leaves of a decision tree in the modified query model Q' induce a partition of the input set where each part is itself a product set along rows.*

Proof. We start with $\mathcal{M}$, a product set along rows with components $\mathcal{M}_i$. Let us observe how this set is partitioned as we go down a decision tree. A YES answer imposes two additional constraints of the form $-1 \leq q \cdot x \leq 1$ on every set $\mathcal{M}_i$. For a NO answer with response (i, s), we get two constraints for all $\mathcal{M}_j$, $1 \leq j < i$, one constraint for the i'th set and no new constraints for the remaining sets. Given this information, a particular setting of any row (or subset of rows) gives no additional information about the other rows. Thus, the set of possible matrices at each child of the current query is a product set along rows. The lemma follows by applying this argument recursively. □

1.4 Distributions and concentration properties

We use two distributions on $n \times n$ matrices called D and D' for our randomized lower bounds. A random matrix from D is obtained by selecting each row independently and uniformly from the ball of radius $\sqrt{n}$. A random matrix from D' is obtained by selecting each entry of the matrix independently and uniformly from the interval $[-1, 1]$. In the analysis, we will also encounter random matrices where each entry is selected independently from $N(0, 1)$. We use the following properties.

Lemma 1.5. *Let σ be the minimum singular value of an $n \times n$ matrix G with independent entries from $N(0, 1)$. For any $t > 0$,*

$$\Pr\left(\sigma\sqrt{n} \leq t\right) \leq t.$$

Proof. To bound σ, we will consider the formula for the density of $\lambda = \sigma^2$ given in [14]:

$$f(\lambda) = \frac{n}{2^{n-1/2}} \frac{\Gamma(n)}{\Gamma(n/2)} \lambda^{-1/2} e^{-\lambda n/2} U\left(\frac{n-1}{2}, -\frac{1}{2}, \frac{\lambda}{2}\right)$$

where U is the Tricomi function, which satisfies for all $\lambda \geq 0$:

- $U(\frac{n-1}{2}, -\frac{1}{2}, 0) = \Gamma(3/2)/\Gamma((n+2)/2)$,
- $U(\frac{n-1}{2}, -\frac{1}{2}, \lambda) \geq 0$
- $\frac{d}{d\lambda} U(\frac{n-1}{2}, -\frac{1}{2}, \lambda) \leq 0$

We will now prove that for any n the density function of $t = \sqrt{n\lambda}$ is at most 1. To see this, the density of t is given by

$$g(t) = f\left(\frac{t^2}{n}\right)\frac{2t}{n} = 2f(\lambda)\sqrt{\frac{\lambda}{n}} = \frac{\sqrt{n}}{2^{n-3/2}}\frac{\Gamma(n)}{\Gamma(n/2)} e^{-\lambda n/2} U\left(\frac{n-1}{2}, -\frac{1}{2}, \frac{\lambda}{2}\right).$$

Now,

$$\frac{d}{dt}g(t) = \frac{\sqrt{n}}{2^{n-3/2}}\frac{\Gamma(n)}{\Gamma(n/2)} \times \\ \times \left[-\frac{n}{2}e^{-\lambda n/2} U\left(\frac{n-1}{2}, -\frac{1}{2}, \frac{\lambda}{2}\right) + e^{-\lambda n/2}\frac{d}{d\lambda} U\left(\frac{n-1}{2}, -\frac{1}{2}, \frac{\lambda}{2}\right)\right]\frac{2t}{n} \leq 0.$$

Thus, the maximum of g is at $t = 0$, and

$$g(0) = \frac{\sqrt{n}}{2^{n-3/2}}\frac{\Gamma(n)}{\Gamma(n/2)}\frac{\Gamma(3/2)}{\Gamma(\frac{n+2}{2})} \leq 1.$$

It follows that $\Pr(\sigma\sqrt{n} \leq \alpha) \leq \alpha$. □

Lemma 1.6. *Let X be a random n-dimensional vector with independent entries from $N(0,1)$. Then for $\epsilon > 0$*

$$\Pr\big(\|X\|^2 \geq (1+\epsilon)n\big) \leq \big((1+\epsilon)e^{-\epsilon}\big)^{n/2}$$

and for $\epsilon \in (0,1)$

$$\Pr\big(\|X\|^2 \leq (1-\epsilon)n\big) \leq \big((1-\epsilon)e^{\epsilon}\big)^{n/2}.$$

For a proof, see [37, Lemma 1.3].

Lemma 1.7. *Let X be a uniform random vector in the n-dimensional ball of radius r. Let Y be an independent random n-dimensional unit vector. Then,*

$$\mathbb{E}(\|X\|^2) = \frac{nr^2}{n+2} \quad \text{and} \quad \mathbb{E}\big((X\cdot Y)^2\big) = \frac{r^2}{n+2}.$$

Proof. For the first part, we have

$$\mathbb{E}(\|X\|^2) = \frac{\int_0^r t^{n+1}dt}{\int_0^r t^{n-1}dt} = \frac{nr^2}{n+2}.$$

For the second part, because of the independence and the symmetry we can assume that Y is any fixed vector, say $(1, 0, \dots, 0)$. Then $\mathbb{E}\big((X \cdot Y)^2\big) = \mathbb{E}(X_1^2)$. But

$$\mathbb{E}(X_1^2) = \mathbb{E}(X_2^2) = \cdots = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}(X_i^2) = \frac{\mathbb{E}(\|X\|^2)}{n} = \frac{r^2}{n+2}.$$

□

Lemma 1.8. *There exists a constant $c > 0$ such that if $P \subseteq \mathbb{R}^n$ compact and X is a random point in P then*

$$\mathbb{E}(\|X\|^2) \geq c(\operatorname{vol} P)^{2/n} n$$

Proof. For a given value of $\operatorname{vol} P$, the value $\mathbb{E}(\|X\|^2)$ is minimized when P is a ball centered at the origin. It is known that the volume of the ball of radius r is at least $c^{n/2} r^n / n^{n/2}$ for some $c > 0$. This implies that, for a given value of $\operatorname{vol} P$, the radius r of the ball of that volume satisfies

$$\frac{c^{n/2} r^n}{n^{n/2}} \geq \operatorname{vol} P. \tag{1.1}$$

On the other hand, Lemma 1.7 claims that for Y a random point in the ball of radius r, we have

$$\mathbb{E}(\|Y\|^2) = \frac{nr^2}{n+2}. \tag{1.2}$$

Combining (1.1), (1.2) and the minimality of the ball, we get

$$\left(\frac{c\,\mathbb{E}(\|X\|^2)(n+2)}{n^2}\right)^{n/2} \geq \operatorname{vol} P$$

and this implies the desired inequality. □

We conclude this section with two elementary properties of variance.

Lemma 1.9. *Let X, Y be independent real-valued random variables. Then*

$$\frac{\operatorname{var}(XY)}{(\mathbb{E}(XY))^2} = \left(1 + \frac{\operatorname{var} X}{(\mathbb{E}\, X)^2}\right)\left(1 + \frac{\operatorname{var} Y}{(\mathbb{E}\, Y)^2}\right) - 1 \geq \frac{\operatorname{var} X}{(\mathbb{E}\, X)^2} + \frac{\operatorname{var} Y}{(\mathbb{E}\, Y)^2}.$$

Lemma 1.10. *For real-valued random variables X, Y, $\operatorname{var} X = \mathbb{E}_Y \operatorname{var}(X \;/\; Y) + \operatorname{var}_Y \mathbb{E}(X \;/\; Y)$.*

Chapter 2

Deterministic lower bounds

2.1 Volume

About the possibility of exact computation of the volume, there are at least two negative answers. The first is that it is know to be $\#P$-hard when the polytope is given as a list of vertices or as an intersection of halfspaces [11]. Secondly, one can construct a polytope given by rational inequalities having rational volume a/b such that writing b needs a number of bits that is exponential in the bit-length of the input [23].

About the possibility of a deterministic approximation algorithm with a membership oracle, the answer is also negative. The complexity of an algorithm is measured by the number of such queries. The work of Elekes [15] and Bárány and Füredi [4] showed that any deterministic polynomial-time algorithm cannot estimate the volume to within an exponential (in n) factor. We quote their theorem below.

Theorem 2.1 ([4])**.** *For every deterministic algorithm that uses at most n^a membership queries and given a convex body K with $B_n \subseteq K \subseteq nB_n$ outputs two numbers A, B such that $A \leq \mathrm{vol}(K) \leq B$, there exists a body K' for which the ratio B/A is at least*

$$\left(\frac{cn}{a \log n}\right)^n$$

where c is an absolute constant.

We will see in Chapter 4 that there are randomized algorithms that approximate the volume in polynomial time, which shows that randomization provably helps in this problem.

2.2 Centroid

Given a convex body in $\mathbb{R}^n$, the centroid is a basic property that one may want to compute. It is a natural way of representing or summarizing the set with just a single point. There are also diverse algorithms that use centroid computation as a subroutine (for an example, see [5], convex optimization). The following non-trivial property illustrates the power of the centroid: Any hyperplane through the centroid of a convex body cuts it into two parts such that each has a volume that is at least a $1/e$ fraction of the volume of the body.

There are no known efficient deterministic algorithms for computing the centroid of a convex body exactly. We will see that this is natural by proving the following result:

Theorem 2.2. *It is $\#P$-hard to compute the centroid of an polytope given as an intersection of halfspaces, even if the polytope is an order polytope.*

(Order polytopes are defined in Subsection 2.2.1)

By centroid computation being $\#P$-hard we mean here that for any problem in $\#P$, there is a polynomial time Turing machine with an oracle for centroids of order polytopes that solves that problem.

On the other hand, there are efficient randomized algorithms for approximating the centroid of a convex body given by a membership oracle (See [5]. Essentially, take the average of $O(n)$ random points in the body. Efficient sampling from a convex body is achieved by a random walk, as explained in [28]). We will see that no deterministic algorithm can match this, by proving the following:

Theorem 2.3. *There is no polynomial time deterministic algorithm that when given access to a membership oracle of a convex body K such that*

$$\frac{1}{17n^2}B_n \subseteq K \subseteq 2nB_n$$

outputs a point at distance $\sigma/100$ of the centroid, where σ^2 is the minimum eigenvalue of the inertia matrix of K.

(The inertia matrix of a convex body is defined in Subsection 2.2.1)

That the centroid is hard to compute is in some sense folklore, but we are not aware of any rigorous analysis of its hardness. The hardness is mentioned in [5] and [20] without proof, for example.

2.2.1 Preliminaries

Let $K \subseteq \mathbb{R}^n$ be a compact set with nonempty interior. Let X be a random point in K. The *centroid* of K is the point $c = \mathbb{E}(X)$. The *inertia matrix* of K is the n by n matrix $\mathbb{E}\big((X-c)(X-c)^T\big)$.

For $K \subseteq \mathbb{R}^n$ bounded and a a unit vector, let $w_a(K)$, the width of K along a, be defined as:

$$w_a(K) = \sup_{x \in K} a \cdot x - \inf_{x \in K} a \cdot x.$$

By *canonical directions* in $\mathbb{R}^n$ we mean the set of vectors that form the columns of the n by n identity matrix.

For $a, b, c \in \mathbb{R}$, $a, b > 0$ and $c \geq 1$ we say that *a is within a factor of c of b* iff

$$\frac{1}{c}b \leq a \leq cb.$$

For a partial order $\prec$ of $[n] = \{1, \dots, n\}$, the *order polytope* associated to it is

$$P(\prec) = \{x \in [0,1]^n \, : \, x_i \leq x_j \text{ whenever } i \prec j\}.$$

In [9], it is proved that computing the volume of order polytopes (given the partial order or, equivalently, the facets of the polytope) is $\#P$-hard. We will use this result to prove Theorem 2.2.

We will also need the following result for isotropic convex bodies, which is in a sense folklore (A convex body $K \subseteq \mathbb{R}^n$ is in isotropic position iff for random $X \in K$ we have $\mathbb{E}(X) = 0$ and $\mathbb{E}(XX^T) = I$). A proof can be found in [21].

Theorem 2.4. *Let $K \subseteq \mathbb{R}^n$ be a convex body in isotropic position. Then*

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

2.2.2 Proofs

The idea of both proofs is to reduce volume computation to centroid computation, given that it is know in several senses that volume computation is hard.

A basic step in the proofs is the following *key idea*: if a convex body is cut into two pieces, then one can know the ratio between the volumes of the pieces if one knows the centroids of the pieces and of the convex body. Namely, if the body has centroid c and the pieces have centroids c_-, c_+, then the volumes of the pieces are in proportion $\|c - c_-\| / \|c - c_+\|$.

It is known that the volume of a polytope given as an intersection of halfspaces can have a bit-length that is exponential in the length of the input [23]. It is not hard to see that the centroid of a polytope given in that form may also need exponential space. Thus, to achieve a polynomial time reduction from volume to centroid, we need to consider a family of polytopes such that all the centroids that appear in the reduction have a length that is polynomial in the length of the input. To this end we consider the fact that it is $\#P$-hard to compute the volume of order polytopes.

Lemma 2.5. *Let P be an order polytope. Then the centroid of P and the volume of P have a bit-length that is polynomial in the bit-length of P.*

Proof. Call "total order polytope" an order polytope corresponding to a total order. Such a polytope is actually a simplex with 0–1 vertices, its volume is $1/n!$ and its centroid has polynomial bit-length. The set of total order polytopes forms a partition of $[0,1]^n$ into $n!$ parts, and any order polytope is a disjoint union of at most $n!$ total order polytopes. The lemma follows. □

Proof (of Theorem 2.2). Let $P \subseteq [0,1]^n$ be an order polytope, given as a set of halfspaces of the form $H_k = \{x \, : \, x_{i_k} \leq x_{j_k}\}$, $k = 1, \dots, K$. Suppose that we have access to an oracle that can compute the centroid of an order polytope. Then we can compute $\operatorname{vol} P$ in the following way: Consider the sequence of bodies that starts with $[0,1]^n$, and then adds one constraint at a time until we reach P. That is, $P_0 = [0,1]^n$,

$P_k = P_{k-1} \cap H_k$. In order to use the *key idea*, for every k, let $Q_k = P_{k-1} \setminus P_k$, compute the centroid c_k of P_k and the centroid d_k of Q_k. We have $P_{k-1} = P_k \uplus Q_k$ and

$$\frac{\operatorname{vol} Q_k}{\operatorname{vol} P_k} = \frac{\|c_{k-1} - d_k\|}{\|c_{k-1} - c_k\|}.$$

Thus,

$$\frac{\operatorname{vol} P_{k-1}}{\operatorname{vol} P_k} = \frac{\|c_{k-1} - d_k\|}{\|c_{k-1} - c_k\|} + 1.$$

This implies, multiplying over all k,

$$\operatorname{vol} P = \prod_{k=1}^{K} \left(\frac{\|c_{k-1} - d_k\|}{\|c_{k-1} - c_k\|} + 1 \right)^{-1}.$$

The reduction costs $2K$ centroid oracle calls. Even though some expressions involve norms, all the intermediate quantities are rational (as the volumes of order polytopes are rational). Moreover, the bit-length of the intermediate quantities is polynomial in n (Lemma 2.5). □

Proof (of Theorem 2.3). Suppose for a contradiction that there exists an algorithm that finds a point at distance $C\sigma$ of the centroid. Then the following algorithm would approximate the volume in a way that contradicts Theorem 2.1, for a value of C to be determined. Theorem 2.1 is actually proved for a family of convex bodies restricted in the following way: We can assume that the body contains the axis-aligned cross-polytope of diameter $2n$ and is contained in the axis-aligned hypercube of side $2n$. Let P be a convex body satisfying that constraint, given as a membership oracle.

Algorithm

```
1. Let $M = 1$, $i = 0$, $P_0 = P$.
2. For every canonical direction $a$:
   (a) While $w_a(P) \geq 1$:
       i. $i \leftarrow i + 1$.
       ii. Compute an approximate centroid $c_{i-1}$ of $P_{i-1}$. Let $H$ be the hyperplane through $c$ orthogonal to $a$.
       iii. Let $P_i$ be (as an oracle) the intersection of $P_{i-1}$ and the halfspace determined by $H$ containing the origin (if $H$ contains the origin, pick any halfspace).
       iv. Let $Q_i$ be (as an oracle) $P_{i-1} \setminus P_i$.
       v. Compute an approximate centroid $d_i$ of $Q_i$.
       vi. $M \leftarrow M \frac{\|c_{i-1} - c_i\|}{\|d_i - c_i\|}$
3. Let $V$ be the volume of $P_i$. Output $V/M$.
```

To see that the algorithm terminates, we will show that the "while" loop ends after $O(n \log n)$ iterations. Assuming that $C \leq 1/2$, at every iteration $w_a(P_i)$ decreases at most by a factor of $1/(4n)$ (Lemma 2.9). Thus, P_i always contains a hypercube of side $1/(4n)$, and $\operatorname{vol} P_i \geq 1/(4n)^n$. Initially, $\operatorname{vol} P_0 \leq (2n)^n$, and every iteration multiplies the volume by a factor of at most $1 - \frac{1}{e} + C$ (Lemma 2.8). Thus, the algorithm runs for at most

$$\frac{2n \log(n\sqrt{8})}{\log\left(1 - \frac{1}{e} + C\right)^{-1}}$$

iterations.

We will now argue that for all the centroids that the algorithm computes, it knows a ball contained in the corresponding body. Let σ_i^2 be the minimum eigenvalue of the inertia matrix of P_i. Initially, the algorithm knows that P_0 contains a ball of radius $\sqrt{n}$ around the origin. Also, Theorem 2.4 implies that for every i, P_i contains a ball of radius $\sigma_i \sqrt{(n+2)/n}$ around the centroid. Because P_i contains a hypercube of side $1/(4n)$, Theorem 2.4 also implies that $\sigma_i \geq 1/(8n\sqrt{n(n+2)})$. Thus, after we compute c_i, the algorithm knows that P_i contains a ball of radius

$$\left(\sqrt{\frac{n+2}{n}} - C\right)\sigma_i \geq \left(\sqrt{\frac{n+2}{n}} - C\right)\frac{1}{8n\sqrt{n(n+2)}} \geq \frac{1-C}{8n^2}$$

around c_i, and this implies that the algorithm knows that P_{i+1}, Q_{i+1} contain balls of radius $(1-C)/(16n^2)$ around known points.

At step 3, P_i contains the origin and has width at most 1 along all canonical directions. This implies that it is completely contained in the input body, as the input body contains the cross-polytope of diameter $2n$. Thus, the volume of P_i is easy to compute because it is a hypercube that we know explicitly at this point, the intersection of all the halfspaces chosen by the algorithm.

At every cut, $\|c_{i-1} - c_i\|$ is within a constant factor of the true value, as the following argument shows: Let δ_i^2 be the minimum eigenvalue of the inertia matrix of Q_i. Let $\bar{c}_i$, $\bar{d}_i$ be the centroids of P_i, Q_i, respectively. We have that $\|c_i - \bar{c}_i\| \leq C\sigma_i \leq C\sigma_{i-1}$ and $\|d_i - \bar{d}_i\| \leq C\delta_i \leq C\sigma_{i-1}$. That is,

$$\|\bar{c}_{i-1} - \bar{c}_i\| - 2C\sigma_{i-1} \leq \|c_{i-1} - c_i\| \leq \|\bar{c}_{i-1} - \bar{c}_i\| + 2C\sigma_{i-1}$$

and we also have that the true distance satisfies $\|\bar{c}_{i-1} - \bar{c}_i\| \geq \sigma_{i-1}/2$ (Lemma 2.6). Thus, the estimate satisfies:

$$(1-4C)\|\bar{c}_{i-1} - \bar{c}_i\| \leq \|c_{i-1} - c_i\| \leq (1+4C)\|\bar{c}_{i-1} - \bar{c}_i\|.$$

A similar argument shows:

$$(1-4C)\|\bar{d}_i - \bar{c}_i\| \leq \|d_i - c_i\| \leq (1+4C)\|\bar{d}_i - \bar{c}_i\|.$$

Thus, M, as an estimate of $V/\operatorname{vol} P$, is within a factor of

$$\left(\frac{1+4C}{1-4C}\right)^{\frac{2n\log(n\sqrt{8})}{\log(1-\frac{1}{e}+C)^{-1}}}$$

of the true value, and so is the estimate of the volume, V/M, with respect to $\operatorname{vol} P$. The choice of $C = 1/100$ would give the contradiction. □

Lemma 2.6 (centroid versus σ)**.** *Let $K \subseteq \mathbb{R}^n$ be a convex body with centroid at the origin. Let a be a unit vector. Let $K_+ = K \cap \{x : a \cdot x \geq 0\}$. Let X be random in K. Let c be the centroid of K_+. Let $\sigma^2 = \mathbb{E}\big((X \cdot a)^2\big)$. Then*

$$c \cdot a \geq \sigma/2.$$

Proof. Let X_+ be random in K_+, let $K_- = K \setminus K_+$, let X_- be random in K_-. Let $\sigma_+^2 = \mathbb{E}((X_+ \cdot a)^2)$, $\sigma_-^2 = \mathbb{E}((X_- \cdot a)^2)$. Lemma 2.7 implies $c \cdot a \geq \sigma_+/\sqrt{2}$. To relate σ and σ_+, we observe that σ is between σ_+ and σ_-, and we use Lemma 2.7 again:

$$\sigma_+ \geq \mathbb{E}(X_+ \cdot a) = -\mathbb{E}(X_- \cdot a) \geq \sigma_-/\sqrt{2}.$$

This implies $\sigma_+ \geq \sigma/\sqrt{2}$ and the lemma follows. □

The following is a particular case of Lemma 5.3 (c) in [30].

Lemma 2.7 ($\mathbb{E}(X)$ versus $\mathbb{E}(X^2)$)**.** *Let X be a non-negative random variable with logconcave density function $f : \mathbb{R}^+ \to \mathbb{R}$. Then*

$$2(\mathbb{E}\, X)^2 \geq \mathbb{E}(X^2).$$

The next lemma follows from the proof of Theorem 1 in [5]:

Lemma 2.8 (volume lemma)**.** *Let $K \subseteq \mathbb{R}^n$ be a convex body with centroid at the origin, let σ^2 be the minimum eigenvalue of the inertia matrix of K, let $c \in \mathbb{R}^n$. Let a be a unit vector. Let $K_+ = K \cap \{x : a \cdot x \geq a \cdot c\}$. Then*

$$\operatorname{vol} K_+ \geq \left(\frac{1}{e} - \frac{|c \cdot a|}{\sigma}\right) \operatorname{vol} K.$$

Lemma 2.9 (width lemma)**.** *Let $K \subseteq \mathbb{R}^n$ be a convex body with centroid at the origin, let σ^2 be the minimum eigenvalue of the inertia matrix of K, let $c \in \mathbb{R}^n$. Let a be a unit vector. Let $K_+ = K \cap \{x : a \cdot x \geq a \cdot c\}$. Then*

$$w_a(K_+) \geq \left(1 - \frac{|c \cdot a|}{\sigma}\right) \frac{w_a(K)}{2n}.$$

Proof. In view of Theorem 2.4, consider an ellipsoid E centered at the origin such that $E \subseteq K \subseteq nE$. Theorem 2.4 implies that we can choose E so that $\frac{1}{2} w_a(E) \geq \sigma$.

Then

$$\begin{aligned} w_a(K^+) &\geq \frac{1}{2} w_a(E) - |c \cdot a| \\ &\geq \left(1 - \frac{|c \cdot a|}{\sigma}\right) \frac{1}{2} w_a(E) \\ &\geq \left(1 - \frac{|c \cdot a|}{\sigma}\right) \frac{1}{2n} w_a(K). \end{aligned}$$

□

2.2.3 Discussion

We proved two hardness results for the computation of the centroid of a convex body. Some open problems:

- Find a substantial improvement of Theorem 2.3, that is, is the centroid hard to approximate even within a ball of radius superlinear in σ?
- Prove a lower bound on the query complexity of any randomized algorithm that approximates the centroid. A possible approach may be given by the lower bound for volume approximation in Chapter 4.

Chapter 3

Dispersion of mass

For the lower bounds in Chapter 5 (length, product of lengths), the main tool in the analysis is a geometric dispersion lemma that is of independent interest in asymptotic convex geometry. Before stating the lemma, we give some background and motivation. There is an elegant body of work that studies the distribution of a random point X from a convex body K [2, 7, 8, 31]. A convex body K is said to be in *isotropic* position if $\text{vol}(K) = 1$ and for a random point X we have

$$\mathbb{E}(X) = 0, \quad \text{and} \quad \mathbb{E}(XX^T) = \alpha I \text{ for some } \alpha > 0.$$

We note that there is a slightly different definition of isotropy (more convenient for algorithmic purposes) which does not restrict $\text{vol}(K)$ and replaces the second condition above by $\mathbb{E}(XX^T) = I$. Any convex body can be put in isotropic position by an affine transformation. A famous conjecture (*isotropic constant*) says that α is bounded by a universal constant for every convex body and dimension. It follows that $\mathbb{E}(\|X\|^2) = O(n)$. Motivated by the analysis of random walks, Lovász and Vempala made the following conjecture (under either definition). If true, then some natural random walks are significantly faster for isotropic convex bodies.

Conjecture 3.1. *For a random point X from an isotropic convex body,*

$$\text{var}(\|X\|^2) = O(n).$$

The upper bound of $O(n)$ is achieved, for example, by the isotropic cube. The isotropic ball, on the other hand, has the smallest possible value, $\text{var}(\|X\|^2) = O(1)$. The variance lower bound we prove in this work (Theorem 3.5) directly implies the following: for an isotropic convex polytope P in $\mathbb{R}^n$ with at most $\text{poly}(n)$ facets,

$$\text{var}(\|X\|^2) = \Omega\left(\frac{n}{\log n}\right).$$

Thus, the conjecture is nearly tight for not just the cube, but *any* isotropic polytope with a small number of facets. Intuitively, our lower bound shows that the length of a random point from such a polytope is *not concentrated* as long as the volume is reasonably large. Roughly speaking, this says that in order to determine the length, one would have to localize the entire vector in a small region.

Returning to the analysis of algorithms, one can view the output of a randomized algorithm as a distribution. Proving a lower bound on the complexity is then equivalent to showing that the output distribution after some number of steps is *dispersed.* To this end, we define a simple parameter of a distribution:

Definition 3.2. *Let μ be a probability measure on $\mathbb{R}$. For any $0 < p < 1$, the* p-dispersion *of μ is*

$$\mathrm{disp}_\mu(p) = \inf\{|a-b| \,:\, a, b \in \mathbb{R}, \mu([a,b]) \geq 1-p\}.$$

Thus, for any possible output z, and a random point X, with probability at least p, $|X-z| \geq \mathrm{disp}_\mu(p)/2$.

We begin with two simple cases in which large variance implies large dispersion.

Lemma 3.3. *Let X be a real random variable with finite variance σ^2.*

a. If the support of X is contained in an interval of length M then $\mathrm{disp}_X(\frac{3\sigma^2}{4M^2}) \geq \sigma$.

b. If X has a logconcave density then $\mathrm{disp}_X(p) \geq (1-p)\sigma$.

Proof. Let $a, b \in \mathbb{R}$ be such that $b - a < \sigma$. Let $\alpha = \Pr(X \notin [a,b])$. Then

$$\operatorname{var} X \leq (1-\alpha)\left(\frac{b-a}{2}\right)^2 + \alpha M^2.$$

This implies

$$\alpha > \frac{3\sigma^2}{4M^2}.$$

For the second part, Lemma 5.5(a) from [30] implies that a logconcave density with variance σ^2 is never greater than $1/\sigma$. This implies that if $a, b \in \mathbb{R}$ are such that $\Pr(X \in [a,b]) \geq p$ then we must have $b - a \geq p\sigma$. □

Lemma 3.4. *Let X, Y be real-valued random variables and Z be a random variable that is generated by setting it equal to X with probability α and equal to Y with probability $1-\alpha$. Then,*

$$\mathrm{disp}_Z(\alpha p) \geq \mathrm{disp}_X(p).$$

3.1 Variance of polytopes

The next theorem states that the length of a random point from a polytope with few facets has large variance. This is a key tool in our lower bounds. It also has a close connection to the variance hypothesis (which conjectures an upper bound for all isotropic convex bodies), suggesting that polytopes might be the limiting case of that conjecture.

Theorem 3.5. *Let $P \subseteq \mathbb{R}^n$ be a polytope with at most n^k facets and contained in the ball of radius n^q. For a random point X in P,*

$$\operatorname{var} \|X\|^2 \geq \operatorname{vol}(P)^{\frac{4}{n}+\frac{3c}{n\log n}} e^{-c(k+3q)} \frac{n}{\log n}$$

where c is a universal constant.

Thus, for a polytope of volume at least 1 contained in a ball of radius at most poly(n), with at most poly(n) facets, we have $\operatorname{var} \|X\|^2 = \Omega(n/\log n)$. In particular this holds for any isotropic polytope with at most poly(n) facets. The proof of Theorem 3.5 is given later in this section.

Let $X \in K$ be a random point in a convex body K. Consider the parameter σ_K of K defined as

$$\sigma_K^2 = \frac{n \operatorname{var} \|X\|^2}{\left(\mathbb{E} \|X\|^2\right)^2}.$$

It has been conjectured that if K is isotropic, then $\sigma_K^2 \leq c$ for some universal constant c independent of K and n (the *variance hypothesis*). Together with the isotropic constant conjecture, it implies Conjecture 3.1. Our lower bound (Theorem 3.5) shows that the conjecture is nearly tight for isotropic polytopes with at most poly(n) facets and they might be the limiting case.

We now give the main ideas of the proof of Theorem 3.5. It is well-known that polytopes with few facets are quite different from the ball. Our theorem is another manifestation of this phenomenon: the width of an annulus that captures most of a polytope is much larger than one that captures most of a ball. The idea of the proof is the following: if $0 \in P$, then we bound the variance in terms of the variance of the cone induced by each facet. This gives us a constant plus the variance of the facet, which is a lower-dimensional version of the original problem. This is the recurrence in our Lemma 3.6. If $0 \notin P$ (which can happen either at the beginning or during the recursion), we would like to translate the polytope so that it contains the origin without increasing $\operatorname{var} \|X\|^2$ too much. This is possible if certain technical conditions hold (case 3 of Lemma 3.6). If not, the remaining situation can be handled directly or reduced to the known cases by partitioning the polytope. It is worth noting that the first case ($0 \in P$) is not generic: translating a convex body that does not contain the origin to a position where the body contains the origin may increase $\operatorname{var} \|X\|^2$ substantially. The next lemma states the basic recurrence used in the proof.

Lemma 3.6 (recurrence)**.** *Let $T(n, f, V)$ be the infimum of* $\operatorname{var} \|X\|^2$ *among all polytopes in $\mathbb{R}^n$ with volume at least V, with at most f facets and contained in the ball of radius $R > 0$. Then there exist constants $c_1, c_2, c_3 > 0$ such that*

$$T(n, f, V) \geq \left(1 - \frac{c_1}{n}\right) T\left(n-1, f+2, \frac{c_2}{nR^2}\left(\frac{V}{Rf}\right)^{1+\frac{2}{n-1}}\right) + \frac{c_3}{R^{8/(n-1)}} \left(\frac{V}{Rf}\right)^{\frac{4}{n-1}+\frac{8}{(n-1)^2}}.$$

Proof. Let P be a polytope as in the statement (not necessarily minimal). Let U be the nearest point to the origin in P. We will use more than one argument, depending on the case:

Case 1: (origin) $0 \in P$.

For every facet F of P, consider the cone C_F obtained by taking the convex hull of the facet and the origin. Consider the affine hyperplane H_F determined by F. Let U be the nearest point to the origin in H_F. Let Y_F be a random point in C_F, and decompose it into a random point $X_F + U$ in F and a scaling factor $t \in [0, 1]$ with a density proportional to t^{n-1}. That is, $Y_F = t(X_F + U)$. We will express $\operatorname{var} \|Y_F\|^2$ as a function of $\operatorname{var} \|X_F\|^2$.

We have that $\|Y_F\|^2 = t^2(\|U\|^2 + \|X_F\|^2)$. Then,

$$\operatorname{var} \|Y_F\|^2 = (\mathbb{E}\, t^4) \operatorname{var} \|X_F\|^2 + (\operatorname{var} t^2)\Big(\|U\|^4 + (\mathbb{E}\, \|X_F\|^2)^2 + 2\|U\|^2\, \mathbb{E}\, \|X_F\|^2\Big) \tag{3.1}$$

Now, for $k \geq 0$

$$\mathbb{E}\, t^k = \frac{n}{n+k}.$$

and

$$\operatorname{var} t^2 = \frac{4n}{(n+4)(n+2)^2} \geq \frac{c_1}{n^2}$$

for $c_1 = 1/2$ and $n \geq 3$. This in (3.1) gives

$$\begin{aligned} \operatorname{var} \|Y_F\|^2 &\geq \frac{n}{n+4} \operatorname{var} \|X_F\|^2 + \frac{c_1}{n^2} \left(\|U\|^4 + (\mathbb{E}\, \|X_F\|^2)^2 + 2\|U\|^2\, \mathbb{E}\, \|X_F\|^2\right) \\ &\geq \frac{n}{n+4} \operatorname{var} \|X_F\|^2 + \frac{c_1}{n^2} \left(\mathbb{E}\, \|X_F\|^2\right)^2. \end{aligned} \tag{3.2}$$

Now, by means of Lemma 1.8, we have that

$$\mathbb{E}\, \|X_F\|^2 \geq c_2 V_{n-1}(F)^{2/(n-1)}(n-1)$$

and this in (3.2) implies for some constant $c_3 > 0$ that

$$\operatorname{var} \|Y_F\|^2 \geq \frac{n}{n+4} \operatorname{var} \|X_F\|^2 + c_3 V_{n-1}(F)^{4/(n-1)}.$$

Using this for all cones induced by facets we get

$$\begin{aligned} \operatorname{var} \|X\|^2 &\geq \frac{1}{\operatorname{vol} P} \sum_{F \text{ facet}} \operatorname{vol} C_F \operatorname{var} \|Y_F\|^2 \\ &\geq \frac{1}{\operatorname{vol} P} \sum_{F \text{ facet}} \operatorname{vol} C_F \left(\frac{n}{n+4} \operatorname{var} \|X_F\|^2 + c_3 V_{n-1}(F)^{4/(n-1)} \right) \end{aligned} \tag{3.3}$$

Now we will argue that $\operatorname{var} \|X_F\|^2$ is at least $T(n-1, f, \frac{V}{Rf})$ for most facets. Because the height of the cones is at most R, we have that the volume of the cones associated to facets having $V_{n-1}(F) \leq \operatorname{vol} P/\alpha$ is at most

$$f\frac{1}{n}R\frac{\operatorname{vol} P}{\alpha}$$

That is, the cones associated to facets having $V_{n-1}(F) > \operatorname{vol} P/\alpha$ are at least a

$$1 - \frac{Rf}{\alpha n}$$

fraction of P. For $\alpha = Rf$ we have that a $1 - 1/n$ fraction of P is composed of cones having facets with $V_{n-1}(F) > \operatorname{vol} P/(Rf)$. Let $\mathcal{F}$ be the set of these facets. The number of facets of any facet F of P is at most f, which implies that for $F \in \mathcal{F}$ we have

$$\operatorname{var} \|X_F\|^2 \geq T(n-1, f, \frac{V}{Rf}).$$

Then (3.3) becomes

$$\begin{aligned}
\operatorname{var} \|X\|^2 &\geq \frac{1}{\operatorname{vol} P} \sum_{F\in\mathcal{F}} \operatorname{vol} C_F \left(\frac{n}{n+4} \operatorname{var} \|X_F\|^2 + c_3 V_{n-1}(F)^{4/(n-1)}\right) \\
&\geq \frac{1}{\operatorname{vol} P} \sum_{F\in\mathcal{F}} \operatorname{vol} C_F \left(\frac{n}{n+4} T\left(n-1, f, \frac{V}{Rf}\right) + c_3 \left(\frac{V}{Rf}\right)^{4/(n-1)}\right) \\
&\geq \left(1 - \frac{1}{n}\right) \left(\frac{n}{n+4} T\left(n-1, f, \frac{V}{Rf}\right) + c_3 \left(\frac{V}{Rf}\right)^{4/(n-1)}\right) \\
&\geq \left(1 - \frac{c_5}{n}\right) T\left(n-1, f, \frac{V}{Rf}\right) + c_4 \left(\frac{V}{Rf}\right)^{4/(n-1)}
\end{aligned}$$

for some constants $c_5, c_4 > 0$.

Case 2: (slicing)

$$\operatorname{var} \mathbb{E}\big(\|X\|^2 \,/\, X \cdot U\big) \geq \beta = \frac{c_4}{16} \left(\frac{V}{Rf}\right)^{4/(n-1)}.$$

In this case, using Lemma 1.10,

$$\begin{aligned}
\operatorname{var} \|X\|^2 &= \mathbb{E} \operatorname{var}\big(\|X\|^2 \,/\, X \cdot U\big) + \operatorname{var} \mathbb{E}\big(\|X\|^2 \,/\, X \cdot U\big) \\
&\geq \mathbb{E} \operatorname{var}\big(\|X\|^2 \,/\, X \cdot U\big) + \beta
\end{aligned} \tag{3.4}$$

Call the set of points $X \in P$ with some prescribed value of $X \cdot U$ a slice. Now we will argue that the variance of a slice is at least $T\big(n-1, f, \frac{V}{2nR}\big)$ for most slices. Because the width of P is at most $2R$, we have that the volume of the slices S having $V_{n-1}(S) \leq V/\alpha$ is at most $2RV/\alpha$. That is, the slices having $V_{n-1}(S) > V/\alpha$ are at least a $1 - 2R/\alpha$ fraction of P. For $\alpha = 2nR$, we have that a $1 - 1/n$ fraction of P are slices with $V_{n-1}(S) > V/(2nR)$. Let $\mathcal{S}$ be the set of these slices. The number of facets of a slice is at most f, which implies that for $S \in \mathcal{S}$ we have $\operatorname{var}\bigl(\|X\|^2 \,/\, X \in S\bigr) \geq T\bigl(n-1, f, \frac{V}{2nR}\bigr)$. Then (3.4) becomes

$$\operatorname{var}\|X\|^2 \geq \left(1 - \frac{1}{n}\right) T\left(n-1, f, \frac{V}{2nR}\right) + \frac{c_4}{16}\left(\frac{V}{Rf}\right)^{4/(n-1)}.$$

Case 3: (translation) $\operatorname{var}(X \cdot U) \leq \beta$ and $\operatorname{var}\mathbb{E}\bigl(\|X\|^2 \,/\, X \cdot U\bigr) < \beta$.

Let $X_0 = X - U$. We have,

$$\operatorname{var}\|X\|^2 = \operatorname{var}\|X_0\|^2 + 4\operatorname{var} X \cdot U + 4\operatorname{cov}(X \cdot U, \|X_0\|^2). \tag{3.5}$$

Now, Cauchy-Schwartz inequality and the fact that $\operatorname{cov}(A, B) = \operatorname{cov}(A, \mathbb{E}(B \,/\, A))$ for random variables A, B, give

$$\begin{aligned}\operatorname{cov}(X \cdot U, \|X_0\|^2) &= \operatorname{cov}(X \cdot U, \|X\|^2 - 2X \cdot U + \|U\|^2)\\ &= \operatorname{cov}(X \cdot U, \|X\|^2) - 2\operatorname{var} X \cdot U\\ &= \operatorname{cov}(X \cdot U, \mathbb{E}(\|X\|^2 \,/\, X \cdot U)) - 2\operatorname{var} X \cdot U\\ &\geq -\sqrt{\operatorname{var} X \cdot U}\sqrt{\operatorname{var}\mathbb{E}(\|X\|^2 \,/\, X \cdot U)} - 2\operatorname{var} X \cdot U.\end{aligned}$$

This in (3.5) gives

$$\begin{aligned}\operatorname{var}\|X\|^2 &\geq \operatorname{var}\|X_0\|^2 - 4\operatorname{var} X \cdot U - 4\sqrt{\operatorname{var} X \cdot U}\sqrt{\operatorname{var}\mathbb{E}\bigl(\|X\|^2 \,/\, X \cdot U\bigr)}\\ &\geq \operatorname{var}\|X_0\|^2 - 8\beta.\end{aligned}$$

Now, X_0 is a random point in a translation of P containing the origin, and thus case 1 applies, giving

$$\operatorname{var}\|X\|^2 \geq \left(1 - \frac{c_5}{n}\right) T\left(n-1, f, \frac{V}{Rf}\right) + \frac{c_4}{2}\left(\frac{V}{Rf}\right)^{4/(n-1)}.$$

Case 4: (partition) otherwise:

In order to control $\operatorname{var} X \cdot U$ for the third case, we will subdivide P into parts so that one of previous cases applies to each part. Let $P_1 = P$, let U_i be the nearest point to the origin in P_i (or, if P_i is empty, the sequence stops),

$$Q_i = P_i \cap \left\{x \,:\, \|U_i\| \leq \hat{U}_i \cdot x \leq \|U_i\| + \sqrt{\beta}/R\right\},$$

and $P_{i+1} = P_i \setminus Q_i$. Observe that $\|U_{i+1}\| \geq \|U_i\| + \sqrt{\beta}/R$ and $\|U_i\| \leq R$, this implies that $i \leq R^2/\sqrt{\beta}$ and the sequence is always finite.

For any i and by definition of Q_i we have $\operatorname{var}(X \cdot U_i \,/\, X \in Q_i) = \|U_i\|^2 \operatorname{var}(X \cdot \hat{U}_i \,/\, X \in Q_i) \leq \beta$.

The volume of the parts Q_i having $\operatorname{vol} Q_i \leq V/\alpha$ is at most $\frac{VR^2}{\alpha\sqrt{\beta}}$. That is, the parts having $\operatorname{vol} Q_i > V/\alpha$ are at least a $1 - \frac{R^2}{\alpha\sqrt{\beta}}$ fraction of P. For $\alpha = nR^2/\sqrt{\beta}$ we have that a $1-1/n$ fraction of P are parts with $\operatorname{vol}(Q_i) > V\sqrt{\beta}/(nR^2)$. Let $\mathcal{Q}$ be the set of these parts. The number of facets of a part is at most $f+2$. Thus, applying one of the three previous cases to each part in $\mathcal{Q}$, and using that $f \geq n$,

$$\begin{aligned}
&\operatorname{var} \|X\|^2 \\
&\geq \frac{1}{\operatorname{vol} P} \sum_{Q \in \mathcal{Q}} \operatorname{vol} Q \operatorname{var}(\|X\|^2 \,/\, X \in Q) \\
&\geq \left(1 - \frac{1}{n}\right)\left(\left(1 - \frac{c_5}{n}\right) T\left(n-1, f+2, \frac{V\sqrt{\beta}}{nR^3 \max\{f, 2n\}}\right) + \frac{c_4}{16}\left(\frac{V\sqrt{\beta}}{nR^3 f}\right)^{4/(n-1)}\right) \\
&\geq \left(1 - \frac{1}{n}\right)\left(\left(1 - \frac{c_5}{n}\right) T\left(n-1, f+2, \frac{V\sqrt{\beta}}{2fnR^3}\right) + \frac{c_4}{16}\left(\frac{V\sqrt{\beta}}{nR^3 f}\right)^{4/(n-1)}\right).
\end{aligned}$$

In any of these cases,

$$\begin{aligned}
&\operatorname{var} \|X\|^2 \\
&\geq \left(1 - \frac{c_6}{n}\right) T\left(n-1, f+2, \frac{V}{2Rf} \min\Bigl(1, \frac{\sqrt{\beta}}{nR^2}\Bigr)\right) + c_7\left(\frac{V}{Rf} \min\Bigl(1, \frac{\sqrt{\beta}}{nR^2}\Bigr)\right)^{4/(n-1)}.
\end{aligned} \tag{3.6}$$

Now, by assumption, $V \leq 2^n R^n$, and this implies by definition that

$$\frac{\sqrt{\beta}}{nR^2} \leq O\left(\frac{1}{n}\right).$$

That is,

$$\min\left(1, \frac{\sqrt{\beta}}{nR^2}\right) = O\left(\frac{\sqrt{\beta}}{nR^2}\right)$$

and the lemma follows, after replacing the value of β in Equation (3.6). □

Proof (of Theorem 3.5). Assume that $\operatorname{vol} P = 1$; the inequality claimed in the theorem can be obtained by a scaling, without loss of generality. For $n \geq 13$, this implies that $R \geq 1$. We use the recurrence lemma in a nested way $t = n/\log n$ times[1]. The radius R stays fixed, and the number of facets involved is at most $f + 2t \leq 3f$. Each time, the volume is raised to the power of at most $1 + \frac{2}{n-t}$ and divided by at most

$$c' nR^2 \bigl(R(f+2t)\bigr)^{1+\frac{2}{n-t}} > 1,$$

for $c' = \max(c_2^{-1}, 1)$. That is, after t times the volume is at least (using the fact that

[1] To force t to be an integer would only add irrelevant complications that we omit.

$(1+\frac{2}{n-t})^t = O(1))$

$$\left(c'nR^2\big(R(f+2t)\big)^{1+\frac{2}{n-t}}\right)^{-t(1+\frac{2}{n-t})^t} \geq 1/(3c'nR^3f)^{O(t)}$$

That means that from the recurrence inequality we get (we ignore the expression in "?", as we will discard that term):

$$\begin{aligned} T(n,f,1) \geq \left(1-\frac{c_1}{n}\right)^t T(n-t,f+2t,?) + \\ + c_3 t\left(1-\frac{c_1}{n}\right)^{t-1} \frac{1}{R^{8/(n-t-1)}} \left(\frac{1}{3Rf}\frac{1}{(3c'nR^3f)^{O(t)}}\right)^{\frac{4}{n-1}+\frac{8}{(n-1)^2}}. \end{aligned}$$

We discard the first term and simplify to get,

$$T(n,f,1) \geq \frac{n}{\log n}\left(\frac{1}{R^3f}\right)^{O(1/\log n)}$$

Thus, for a polytope of arbitrary volume we get by means of a scaling that there exists a universal constant $c>0$ such that

$$\operatorname{var}\|X\|^2 \geq (\operatorname{vol}P)^{4/n}\left(\frac{(\operatorname{vol}P)^{3/n}}{R^3f}\right)^{c/\log n}\frac{n}{\log n}.$$

The theorem follows. □

3.2 Dispersion of the determinant

In our proof of the volume lower bound, we begin with a distribution on matrices for which the determinant is dispersed. The main goal of the proof is to show that even after considerable conditioning, the determinant is still dispersed. The notion of a product set along rows (Definition 1.3) will be useful in describing the structure of the distribution and how it changes with conditioning.

Lemma 3.7. *There exists a constant $c>0$ such that for any partition $\{\mathcal{A}^j\}_{j\in N}$ of $(\sqrt{n}B_n)^n$ into $|N| \leq 2^{n^2-2}$ parts where each part is a product set along rows, there exists a subset $N' \subseteq N$ such that*

a. $\operatorname{vol}(\bigcup_{j\in N'}\mathcal{A}^j) \geq \frac{1}{2}\operatorname{vol}\big((\sqrt{n}B_n)^n\big)$ *and*

b. for any $u>0$ and a random point X from $\mathcal{A}^j$ for any $j\in N'$, we have

$$\Pr\big(|\det X| \notin [u,u(1+c)]\big) \geq \frac{1}{2^7n^6}.$$

The proof of this lemma will be postponed until Chapter 4, because it uses some intuition from the volume problem.

3.3 Discussion

It is an open problem as to whether the logarithmic factor in the variance of polytopes with few facets can be removed. The lower bounds in Chapter 5 would improve if the variance lower bound is improved.

Chapter 4

Lower bound for randomized computation of the volume

In striking contrast with the exponential lower bound for deterministic algorithms given in Section 2.1, the celebrated paper of Dyer, Frieze and Kannan [13] gave a polynomial-time randomized algorithm to estimate the volume to arbitrary accuracy (the dependence on n was about n^{23}). This result has been much improved and generalized in subsequent work (n^{16}, [25]; n^{10}, [24, 1]; n^8, [12]; n^7, [26]; n^5, [22]; n^4, [29]); the current fastest algorithm has complexity that grows as roughly $O(n^4/\epsilon^2)$ to estimate the volume to within relative error $1+\epsilon$ with high probability (for recent surveys, see [35, 36]). Each improvement in the complexity has come with fundamental insights and lead to new isoperimetric inequalities, techniques for analyzing convergence of Markov chains, algorithmic tools for rounding and sampling logconcave functions, etc.

These developments lead to the question: what is the best possible complexity of any randomized volume algorithm? A lower bound of $\Omega(n)$ is straightforward. Here we prove a nearly quadratic lower bound: there is a constant $c > 0$ such that any randomized algorithm that approximates the volume to within a $(1+c)$ factor needs $\Omega(n^2/\log n)$ queries. The formal statement appears in Theorem 4.1.

For the more restricted class of randomized nonadaptive algorithms (also called "oblivious"), an exponential lower bound is straightforward (Section 4.3). Thus, the use of full-fledged adaptive randomization is crucial in efficient volume estimation, but cannot improve the complexity below $n^2/\log n$.

In fact, the quadratic lower bound holds for a restricted class of convex bodies, namely parallelotopes. A parallelotope in $\mathbb{R}^n$ centered at the origin can be compactly represented using a matrix as $\{x : \|Ax\|_\infty \leq 1\}$, where A is an $n \times n$ nonsingular matrix; the volume is simply $2^n |\det(A)|^{-1}$. One way to interpret the lower bound theorem is that in order to estimate $|\det(A)|$ one needs almost as many bits of information as the number of entries of the matrix. The main ingredient of the proof is a dispersion lemma (Theorem 3.7) which shows that the determinant of a random matrix remains dispersed even after conditioning the distribution considerably. We discuss other consequences of the lemma in Section 4.4.

We now state our lower bound for randomized algorithms. Its proof is given in

Section 4.2. Besides the dimension n, the complexity also depends on the "roundness" of the input body. This is the ratio R/r where $rB_n \subseteq K \subseteq RB_n$. To avoid another parameter in our results, we ensure that R/r is bounded by a polynomial in n.

Theorem 4.1 (volume)**.** *Let K be a convex body given by a membership oracle such that $B_n \subseteq K \subseteq O(n^8)B_n$. Then there exists a constant $c > 0$ such that any randomized algorithm that outputs a number V such that $(1-c)\operatorname{vol}(K) \leq V \leq (1+c)\operatorname{vol}(K)$ holds with probability at least $1 - 1/n$ has complexity $\Omega(n^2/\log n)$.*

We note that the lower bound can be easily extended to any algorithm with success probability $p > 1/2$ with a small overhead [19]. The theorem actually holds for parallelotopes with the same roundness condition. We restate the theorem for this case.

Theorem 4.2 (determinant)**.** *Let A be an matrix with entries in $[-1, 1]$ and smallest singular value at least $2^{-12}n^{-7}$ that can be accessed by the following oracle: for any x, the oracle determines whether $\|Ax\|_\infty \leq 1$ is true or false. Then there exists a constant $c > 0$ such that any randomized algorithm that outputs a number V such that*

$$(1-c)|\det(A)| \leq V \leq (1+c)|\det(A)|$$

holds with probability at least $1 - 1/n$, has complexity $\Omega(n^2/\log n)$.

4.1 Preliminaries

Throughout this chapter we assume that $n > 12$ to avoid trivial complications.

4.2 Proof of the volume lower bound

We will use the distribution D on parallelotopes (or matrices, equivalently). Recall that a random $n \times n$ matrix R is generated by choosing its rows $R_1, \dots, R_n$ uniformly and independently from the ball of radius $\sqrt{n}$. The convex body corresponding to R is a parallelotope having the rows of R as facets' normals:

$$\{x \in \mathbb{R}^n : (\forall i)|R_i \cdot x| \leq 1\}$$

Its volume is $V : \mathbb{R}^{n \times n} \to \mathbb{R}$ given (a.s.) by $V(R) = 2^n |\det R|^{-1}$.

At a very high level, the main idea of the lower bound is the following: after an algorithm makes all its queries, the set of inputs consistent with those queries is a product set along rows (in the oracle model Q'), while the level sets of the function that the algorithm is trying to approximate, $|\det(\cdot)|$, are far from being product sets. In the partition of the set of inputs induced by any decision tree of height $O(n^2/\log n)$, all parts are product sets of matrices and most parts have large volume, and therefore V is dispersed in most of them. To make this idea more precise, we first examine the structure of a product set along rows, all matrices with *exactly* the same determinant. This abstract "hyperbola" has a rather sparse structure.

Theorem 4.3. *Let $R \subseteq \mathbb{R}^{n\times n}$ be such that $R = \prod_{i=1}^n R_i$, $R_i \subseteq \mathbb{R}^n$ convex and there exists $c > 0$ such that $|\det M| = c$ for all $M \in R$. Then, for some ordering of the R_i's, $R_i \subseteq S_i$, with S_i an $(i-1)$-dimensional affine subspace, $0 \notin S_i$ and satisfying: S_i is a translation of the linear hull of S_{i-1}.*

Proof. By induction on n. It is clearly true for $n = 1$. For arbitrary n, consider the dimension of the affine hull of each R_i, and let R_1 have minimum dimension. Let $a \in R_1$. There will be two cases:

If $R_1 = \{a\}$, then let A be the hyperplane orthogonal to a. If we denote T_i the projection of R_i onto A, then we have that $T = \prod_{i=1}^{n-1} T_i$ satisfies the hypotheses in $A \cong \mathbb{R}^{n-1}$ with constant $c/\|a\|$ and the inductive hypothesis implies that, for some ordering, the $T_2, \ldots, T_n$ are contained in affine subspaces not containing 0 of dimensions $0, \ldots, n-2$ in A, that is, $R_2, \ldots, R_n$ are contained in affine subspaces not containing 0 of dimensions $1, \ldots, n-1$.

If there are $a, b \in R_1$, $b \neq a$, then there is no zero-dimensional R_i. Also, because of the condition on the determinant, b is not parallel to a. Let $x_\lambda = \lambda a + (1-\lambda)b$ and consider the argument of the previous paragraph applied to x_λ and its orthogonal hyperplane. That is, for every λ there is some region T_i in A that is zero-dimensional. In other words, the corresponding R_i is contained in a line. Because there are only $n-1$ possible values of i but an infinite number of values of λ, we have that there exists one region R_i that is picked as the zero-dimensional for at least two different values of λ. That is, R_i is contained in the intersection of two non-parallel lines, and it must be zero-dimensional, which is a contradiction. □

Now we need to extend this to an approximate hyperbola, i.e., a product set along rows with the property that for most of the matrices in the set, the determinant is restricted in a given interval. This extension is the heart of the proof and is captured in Lemma 3.7. We will need a bit of preparation for its proof.

We define two properties of a matrix $R \in \mathbb{R}^{n\times n}$:

- Property $P_1(R,t)$: $\prod_{i=1}^n \|\pi_{R_{-i}^\perp}(R_i)\| \leq t$ ("short 1-D projections").

- Property $P_2(R,t)$: $|\det \hat{R}| \geq t$ ("angles not too small").

Lemma 4.4. *Let R be drawn from distribution D. Then for any $\alpha > 1$,*

a. $\Pr\big(P_1(R,\alpha^n)\big) \geq 1 - \frac{1}{\alpha^2}$,

b. there exists $\beta > 1$ (that depends on α) such that $\Pr(P_2\big(R, 1/\beta^n)\big) \geq 1 - \frac{1}{n^\alpha}$.

Proof. For part (a), by the AM-GM inequality and Lemma 1.7 we have

$$\mathbb{E}\left(\Big(\prod_i \|\pi_{R_{-i}^\perp}(R_i)\|^2\Big)^{1/n}\right) \leq \frac{1}{n}\sum_i \mathbb{E}\,\|\pi_{R_{-i}^\perp}(R_i)\|^2 = \frac{n}{n+2}.$$

Thus, by Markov's inequality,

$$\Pr\left(\prod_i \|\pi_{R_{-i}^{\perp}}(R_i)\| \geq c^n\right) = \Pr\left(\left(\prod_i \|\pi_{R_{-i}^{\perp}}(R_i)\|^2\right)^{1/n} \geq c^2\right) \leq \frac{1}{c^2}.$$

For part (b), we can equivalently pick each entry of R independently as $N(0,1)$. In any case,

$$\det \hat{R} = \frac{\det R}{\prod_i \|R_i\|} = \frac{\prod_i \|\tilde{R}_i\|}{\prod_i \|R_i\|}.$$

We will find an upper bound for the denominator and a lower bound for the numerator.

For the denominator, concentration of a Gaussian vector (Lemma 1.6) gives

$$\Pr(\|R_i\|^2 \geq 4n) \leq 2^{-n}$$

which implies

$$\Pr\left(\prod_{i=1}^{n} \|R_i\|^2 \geq 4^n n^n\right) \leq \Pr\left((\exists i)\|R_i\|^2 \geq 4^n n^n\right) \leq n2^{-n} \leq e^{-\Omega(n)}. \tag{4.1}$$

For the numerator, let $\mu_i = \mathbb{E}\,\|\tilde{R}_i\|^2 = n-i+1$, let $\mu = \mathbb{E}\prod_{i=1}^{n}\|\tilde{R}_i\|^2 = \prod_{i=1}^{n}\mu_i = n!$.

Now, concentration of a Gaussian vector (Lemma 1.6) also gives

$$\Pr\bigl(\|\tilde{R}_i\|^2 \geq \mu_i/2\bigr) \geq 1 - 2^{(n-i+1)/8} \tag{4.2}$$

Alternatively, for $t \in (0,1)$

$$\Pr\bigl(\|\tilde{R}_i\|^2 \geq t\mu_i\bigr) \geq 1 - \sqrt{t\mu_i}(n-i+1). \tag{4.3}$$

Let $c > 0$ be such that $2^{(n-i+1)/8} \leq 1/(2n^{\alpha+1})$ for $i \leq n - c\log n$. Using inequality 4.2 for $i \leq n - c\log n$ and 4.3 for the rest with

$$t = \frac{1}{2n^{2\alpha}(c\log n)^{5/2}}$$

we get

$$\begin{aligned}\Pr&\left(\prod_{i=1}^{n}\|\tilde{R}_i\|^2 \geq \frac{\mu}{2^{n-c\log n}t^{c\log n}}\right)\\ &\geq \prod_{i=1}^{n-c\log n}\Pr\Bigl(\|\tilde{R}_i\|^2 \geq \frac{\mu_i}{2}\Bigr)\prod_{i=n-c\log n}^{n}\Pr\bigl(\|\tilde{R}_i\|^2 \geq t\mu_i\bigr)\\ &\geq 1 - \frac{1}{n^{\alpha}}\end{aligned} \tag{4.4}$$

where, for some $\gamma > 1$ we have $2^{n-c\log n} t^{c \log n} \leq \gamma^n$. The result follows from equations 4.4 and 4.1. □

Proof (of Lemma 3.7). The idea of the proof is the following: If we assume that $|\det(\cdot)|$ of most matrices in a part fits in an interval $[u, u(1+\epsilon)]$, then for most choices R_{-n} of the first $n-1$ rows in that part we have that most choices Y of the last row in that part have $|\det(R_{-n}, Y)|$ in that interval. Thus, in view of the formula[1] $|\det(R_{-n}, Y)| = \|\tilde{Y}\| \prod_{i=1}^{n-1} \|\tilde{R}_i\|^{-1}$ we have that, for most values of Y,

$$\|\tilde{Y}\| \in \big[u, u(1+\epsilon)\big] \prod_{i=1}^{n-1} \|\tilde{R}_i\|^{-1}$$

where $\tilde{Y}$ is the projection of Y to the line orthogonal to $R_1, \dots, R_{n-1}$. In other words, most choices of the last row are forced to be contained in a set of the form $\{x : b \leq |a \cdot x| \leq c\}$, that we call a double band, and the same argument works for the other rows. In a similar way, we get a pair of double bands of "complementary" widths for every pair of rows. These constraints on the part imply that it has small volume, giving a contradiction. This argument only works for parts containing mostly "matrices that are not too singular"—matrices that satisfy P_1 and P_2—, and we choose the parameters of these properties so that at least half of $(\sqrt{n}B_n)^n$ satisfies them.

We will firstly choose N' as the family of large parts that satisfy properties P_1 and P_2 for suitable parameters so that (a) is satisfied. We will say "probability of a subset of $(\sqrt{n}B_n)^n$" to mean its probability with respect to the uniform probability measure on $(\sqrt{n}B_n)^n$. The total probability of the parts having probability at most α is at most $\alpha|N|$. Thus, setting $\alpha = 1/(4|L|)$, the parts having probability at least $1/4|L| \geq 1/2^{n^2}$ have total probability at least 3/4. Since $\operatorname{vol} \cup_{j\in N} \mathcal{A}^j \geq 2^{n^2}$, each of those parts has volume at least 1. Let these parts be indexed by $N'' \subseteq N$. Lemma 4.4 (with $\alpha = 8$ for part (a), $\alpha = 2$ for part (b)) implies that at most 1/4 of $(\sqrt{n}B_n)^n$ does not satisfy $P_1(\cdot, 8^n)$ or $P_2(\cdot, 1/\beta^n)$, and then at least 3/4 of the parts in probability satisfy $P_1(\cdot, 8^n)$ and $P_2(\cdot, 1/\beta^n)$ for at least half of the part in probability. Let $N''' \subseteq N$ be the set of indices of these parts. Let $N' = N'' \cap N'''$. We have that $\cup_{j\in N'} \mathcal{A}^j$ has probability at least 1/2.

We will now prove (b). Let $A = \prod_{i=1}^n A_i$ be one of the parts indexed by N'. Let X be random in A. Let ϵ be a constant and $p_1(n)$ be a function of n both to be fixed later. Assume for a contradiction that there exists u such that

$$\Pr\big(|\det X| \notin [u, u(1+\epsilon)]\big) < p_1(n). \tag{4.5}$$

Let $G \subseteq A$ be the set of $M \in A$ such that $|\det M| \in [u, u(1+\epsilon)]$. Let $p_2(n)$, $p_3(n)$ be functions of n to be chosen later. Consider the subset of points $R \in G$ satisfying:

I. $P_1(R, 8^{n/2})$ and $P_2(R, 1/\beta^n)$,

[1] Recall that $\tilde{R}_i$ is the projection of R_i to the subspace orthogonal to $R_1, \dots, R_{i-1}$.

II. for any $i \in \{1, \dots, n\}$, for at most a $p_2(n)$ fraction of $Y \in A_i$ we have $(Y, R_{-i}) \notin G$, and

III. for any $i, j \in \{1, \dots, n\}$, $i \neq j$, for at most a $p_3(n)$ fraction of $(Y, Z) \in A_i \times A_j$ we have $(Y, Z, R_{-ij}) \notin G$.

Because of the constraints, such a subset is a

$$1 - \Pr(X \notin G) - \Pr(X \text{ not as I, II and III}) \geq 1 - p_1(n) - \frac{1}{2} - n\frac{p_1(n)}{p_2(n)} - n^2\frac{p_1(n)}{p_3(n)} \quad (4.6)$$

fraction of A. The function $p_1(n)$ will be chosen at the end so that the right hand side is positive. Fix a matrix $R = (R_1, \dots, R_n)$ in that subset.

The constraints described in the first paragraph of the proof are formalized in Lemma 4.5, which, for all i, j, gives sets B_{ij} (double bands, of the form $\{x : b \leq |a \cdot x| \leq c\}$), such that most of A_i is contained in $\cap_{j=1}^n B_{ij}$. Lemma 4.5 is invoked in the following way: For each pair i, j with $i < j$, let E be the two-dimensional subspace orthogonal to all the rows of R except i, j. We set X_1 (respectively X_2) distributed as the marginal in E of the uniform probability measure on A_i (respectively A_j). We also set $a_1 = \pi_E(R_i)$, $a_2 = \pi_E(R_j)$, $\alpha = p_3(n)$, $\beta = p_2(n)$ and u and ϵ as here, while γ will be chosen later.

Let l_{ij} be the width of (each component of) the double band B_{ij}. Then, according to Lemma 4.5, the following relations hold:

$$\begin{aligned} l_{ii} &\leq \epsilon \|\pi_{R_{-i}^\perp}(R_i)\| && \text{for any } i, \\ l_{ij} &\leq 4\epsilon \|\pi_{R_{-i}^\perp}(R_i)\| \|\pi_{R_{-j}^\perp}(R_j)\| / l_{ji} && \text{for } i > j. \end{aligned}$$

Since each double band has two components, the intersection of all the n bands associated to a particular region A_i, namely $\cap_{j=1}^n B_{ij}$, is the union of 2^n congruent parallelotopes. Thus, using properties P_1 and P_2 of R and fixing ϵ as a sufficiently small constant, the "feasible region" defined by the double bands, $B = \prod_{i=1}^n \cap_{j=1}^n B_{ij}$, satisfies:

$$\begin{aligned} \operatorname{vol} B &\leq 2^{n^2} \frac{\prod_{i,j=1}^n l_{ij}}{|\det \hat{R}|^n} \\ &\leq 2^{n^2} \frac{\prod_{i=1}^n \left(\epsilon \|\pi_{R_{-i}^\perp}(R_i)\| \prod_{j=2}^i 4\epsilon \|\pi_{R_{-i}^\perp}(R_i)\| \|\pi_{R_{-j}^\perp}(R_j)\| \right)}{|\det \hat{R}|^n} \\ &= 2^{n^2} \frac{\epsilon^{\binom{n}{2}} 4^{\binom{n-1}{2}} \prod_i \|\pi_{R_{-i}^\perp}(R_i)\|^n}{|\det \hat{R}|^n} \\ &\leq 1/4^n. \end{aligned}$$

Each region A_i is not much bigger than the intersection of the corresponding double bands $B_i = \cap_{j=1}^n B_{ij}$ as follows: restricting to the double band B_{ii} removes at most a $p_2(n)$ fraction of A_i, each double band B_{ij} for $j < i$ removes at most a γ fraction of

A_i, and each double band B_{ij} for $j > i$ removes a $p_2(n) + (p_3(n)/\gamma)$ fraction of A_i. We set $\gamma = 1/4n^2$, $p_2(n) = 1/(4n^2)$ and $p_3(n) = 1/(16n^4)$ so that, as a fraction of $\operatorname{vol} A_i$, $\operatorname{vol} B_i$ is no less than

$$1 - np_2(n) - \binom{n}{2}\gamma - \binom{n}{2}\left(p_2(n) + \frac{p_3(n)}{\gamma}\right) \geq 1/2.$$

Thus, $\operatorname{vol} A \leq 2^n \operatorname{vol} B \leq 1/2^n$, which is a contradiction. The condition on $p_1(n)$ given by Equation (4.6) is satisfied for $p_1(n) = 1/(2^7 n^6)$. □

Lemma 4.5 (2-D lemma)**.** *Let X_1, X_2 be two independent random vectors in $\mathbb{R}^2$ with bounded support (not necessarily with the same distribution). Let X be a random matrix with rows X_1, X_2. Assume that there exist $u > 0$, $0 < \epsilon \leq 1$ such that*

$$\Pr\bigl(|\det X| \notin [u, u(1+\epsilon)]\bigr) < \alpha.$$

Let $G = \{M \in \mathbb{R}^{2\times 2} : |\det M| \in [u, u(1+\epsilon)]\}$. Let $a_1, a_2 \in \mathbb{R}^2$ be such that $(a_1, a_2) \in G$ and

$$\Pr(X_1 : (X_1, a_2) \notin G) \leq \beta, \qquad \Pr(X_2 : (X_2, a_1) \notin G) \leq \beta.$$

Let $\gamma > \alpha/(1-\beta)$. Then there exist double bands $B_{ij} \subseteq \mathbb{R}^2$, $b_{ij} \geq 0$, $i, j \in \{1, 2\}$, $l \geq 0$,

$$\begin{aligned}
B_{11} &= \Bigl\{x : |a_2^\perp \cdot x| \in \bigl[b_{11}, b_{11} + \epsilon\|\pi_{a_2^\perp}(a_1)\|\bigr]\Bigr\} \\
B_{22} &= \Bigl\{x : |a_1^\perp \cdot x| \in \bigl[b_{22}, b_{22} + \epsilon\|\pi_{a_1^\perp}(a_2)\|\bigr]\Bigr\} \\
B_{12} &= \Bigl\{x : |a_1^\perp \cdot x| \in \bigl[b_{12}, b_{12} + l\bigr]\Bigr\} \\
B_{21} &= \Bigl\{x : |a_2^\perp \cdot x| \in \bigl[b_{21}, b_{21} + 4\epsilon\|\pi_{a_2^\perp}(a_1)\|\|\pi_{a_1^\perp}(a_2)\|/l\bigr]\Bigr\}
\end{aligned}$$

such that

$$\begin{aligned}
&\Pr(X_1 \notin B_{11}) \leq \beta && \Pr(X_1 \notin B_{12}) \leq \beta + (\alpha/\gamma) \\
&\Pr(X_2 \notin B_{21}) \leq \gamma && \Pr(X_2 \notin B_{22}) \leq \beta.
\end{aligned}$$

Proof. The proof refers to Figure 4-1 which depicts the bands under consideration.

A double band of the form $\{x : |a \cdot x| \in [u, v]\}$ has (additive or absolute) width $v - u$ and relative (or multiplicative) width v/u. Consider the expansion $|\det X| = \|X_2\|\|\pi_{X_2^\perp}(X_1)\|$ and the definition of a_2 to get

$$\Pr\bigl(\|\pi_{a_2^\perp}(X_1)\| \notin \|a_2\|^{-1}[u, u(1+\epsilon)]\bigr) \leq \beta.$$

That is, with probability at most β we have X_1 outside of a double band of relative width $1 + \epsilon$:

$$B_{11} = \bigl\{x : \|\pi_{a_2^\perp}(x)\| \in \|a_2\|^{-1}[u, u(1+\epsilon)]\bigr\}.$$

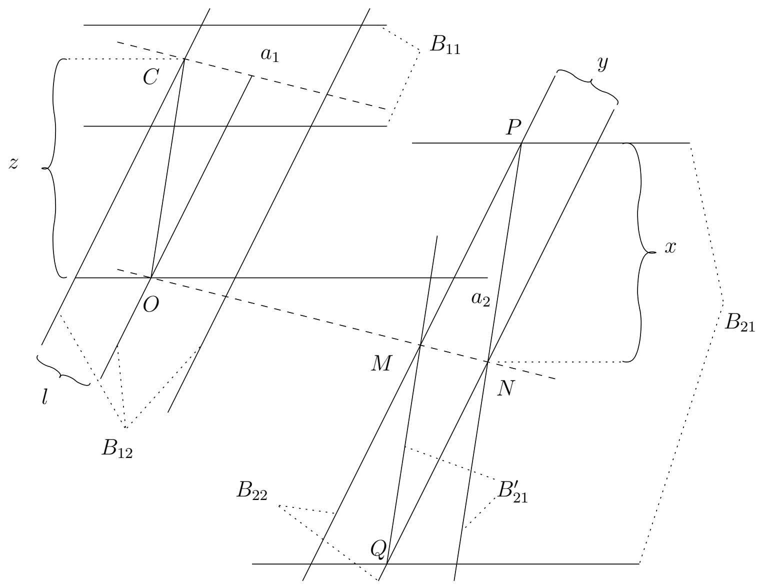

Figure 4-1: The 2-D argument.

Because $a_1 \in B_{11}$, the absolute width is at most $\epsilon\|\pi_{a_2^\perp}(a_1)\|$. If we exchange the roles of a_1 and a_2 in the previous argument, we get a double band B_{22}.

Let $\mathcal{A}$ be the set of $a \in \mathbb{R}^2$ satisfying: $(a, a_2) \in G$ and with probability at most γ over X_2 we have $(X_2, a) \notin G$. We have that

$$\Pr(X_1 \in \mathcal{A}) \geq 1 - \beta - \frac{\alpha}{\gamma}.$$

Consider a point $C \in \mathcal{A}$ that maximizes the distance to the span of a_1. Similarly to the construction of B_{11}, by definition of $\mathcal{A}$ and with probability at most γ we have X_2 outside of a double band of relative width $1 + \epsilon$. We denote it B'_{21}. In order to have better control of the angles between the bands, we want to consider a bigger double band parallel to B_{11}, the minimum such a band that contains the intersection of B_{22} and B'_{21}. Call this band B_{21}. The width of this band is at most $2x$, and the triangles Oa_1C and PMN are similar. Then,

$$\frac{x}{z} = \frac{y}{l},$$

where $l = \|\pi_{a_1^\perp}(C)\|$ is the width of a band imposed on $\mathcal{A}$ by definition of C, y is the width of B_{22}, $y \leq \epsilon\|\pi_{a_1^\perp}(a_2)\|$, and z is the distance between C and the span of a_2, that is,

$$z = \|\pi_{a_2^\perp}(C)\| \leq (1 + \epsilon)\|\pi_{a_2^\perp}(a_1)\| \leq 2\|\pi_{a_2^\perp}(a_1)\|.$$

Thus, $x \leq 2\epsilon\|\pi_{a_2^\perp}(a_1)\|\|\pi_{a_1^\perp}(a_2)\|/l$. Let B_{12} be the band imposed on $\mathcal{A}$ by definition of C. □

We are now ready to prove the complexity lower bounds.

Proof of Theorem 4.2. In view of Yao's lemma, it is enough to prove a lower bound on the complexity of deterministic algorithms against a distribution and then a lower bound on the minimum singular value of matrices according to that distribution. The deterministic lower bound is a consequence of the dispersion of the determinant proved in Theorem 3.7, the bound on the minimum singular value is an easy adaptation of a bound on the minimum singular value of a Gaussian matrix given by Lemma 1.5. These two claims are formalized below.

Claim 1: Let R be a random input according to distribution D. Then there exists a constant $c > 0$ such that any deterministic algorithm that outputs a number V such that

$$(1-c)|\det R| \leq V \leq (1+c)|\det R|$$

with probability at least $1 - 1/(2^8 n^6)$, makes more than

$$\frac{n^2-2}{\log_2(2n+1)}$$

queries in the oracle model Q'.

Claim 2: Let A be an $n \times n$ random matrix from distribution D. Let σ be the minimum singular value of A. Then for any $t \geq 0$

$$\Pr(\sigma\sqrt{n} \leq t) \leq 4t + \frac{n}{2^{n-1}}$$

(the choice of $t = 1/(2^{12}n^6)$ proves Theorem 4.2).

Proof of Claim 1: For a deterministic algorithm and a value of n, consider the corresponding decision tree. Let

$$h \leq \frac{n^2-2}{\log_2(2n+1)}$$

be the height and L be the set of leaves of this tree. Let $(P_l)_{l\in L}$ be the partition on the support of D induced by the tree.

Every query has at most $2n+1$ different answers, and every path has height at most h. Thus,

$$|L| \leq (2n+1)^h = 2^{n^2-2}.$$

The sets P_l are convex and Lemma 1.4 guarantees that they are also product sets of matrices, and hence by Lemma 3.7 we have that there exists a constant $c > 0$ such that with probability at least $1/(2^8 n^6)$ and for any $a > 0$ we have that $|\det R|$ is outside of $[a, (1+c)a]$. Claim 1 follows.

Proof of Claim 2: We will bound $\|A^{-1}\|_2 = 1/\sigma$. To achieve this, we will reduce the problem to the case where the entries of the matrix are $N(0,1)$ and independent. We write $A = GDE$, where G has its entries independently as $N(0,1)$, D is the diagonal matrix that normalizes the rows of G and E is another random diagonal matrix independent of (G, D) that scales the rows of GD to give them the length distribution of a random vector in $\sqrt{n}B_n$. We have

$$\|A^{-1}\|_2 \le \|D^{-1}\|_2 \|E^{-1}\|_2 \|G^{-1}\|_2. \tag{4.7}$$

Now, with probability at least $1 - n/2^n$ the diagonal entries of E are at least $\sqrt{n}/2$. Thus, except for an event that happens with probability $n/2^n$,

$$\|E^{-1}\|_2 \le 2/\sqrt{n} \tag{4.8}$$

On the other hand, Lemma 1.6 (with $\epsilon = 3$) implies that with probability at least $1 - n/2^n$ the diagonal entries of D^{-1} are at most $2\sqrt{n}$. Thus, except for an event that happens with probability $n/2^n$,

$$\|D^{-1}\|_2 \le 2\sqrt{n}. \tag{4.9}$$

From (4.7), (4.8) and (4.9), we get $\|A^{-1}\|_2 \le 4\|E^{-1}\|$. Using Lemma 1.5 which bounds the singular values for a Gaussian matrix, Claim 2 follows. □

Finally, Theorem 4.1 is a simple consequence.

Proof of Theorem 4.1. It remains to prove that a parallelotope given by a random matrix A contains $B_n/\sqrt{n}$ and is contained in $\frac{\sqrt{n}}{\sigma}B_n$ whenever $\sigma > 0$, where σ is the minimum singular value of A. The first inclusion is evident since the entries must be from $[-1, 1]$. It is sufficient to prove the second inclusion for the vertices of the parallelotope, i.e., solutions to $Ax = b$ for any $b \in \{-1, 1\}^n$. That is, $x = A^{-1}b$ and therefore

$$\|x\| \le \|A^{-1}\|_2 \|b\| \le \sqrt{n}/\sigma.$$

□

4.3 Nonadaptive volume algorithms

An algorithm is *nonadaptive* if its queries are independent of the input.

Theorem 4.6 (nonadaptive lower bound)**.** *Let K be a convex body given by a membership oracle such that $B_n \subseteq K \subseteq 2nB_n$. Then any nonadaptive randomized algorithm that outputs a number V such that $.9\,\mathrm{vol}(K) \le V \le 1.1\,\mathrm{vol}(K)$ holds with probability at least $3/4$ has complexity at least $\frac{1}{4}(4n)^{n/2}$.*

Proof. Consider the distribution on parallelotopes induced by the following procedure: first, with equal probability choose one of the following bodies:

- (“brick”) $\big\{x \in \mathbb{R}^n : (\forall i \in \{2, \ldots, n\})\ |x_i| \le 1\big\} \cap nB_n$
- (“double brick”) $\big\{x \in \mathbb{R}^n : (\forall i \in \{2, \ldots,\ n\})\ |x_i| \le 1\big\} \cap 2nB_n$

and then, independently of the first choice, apply a random rotation.

We will prove the following claim, from which the desired conclusion can be obtained by means of Yao’s lemma.

Claim: Let K be a parallelotope according to the previous distribution. Then any nonadaptive deterministic algorithm that outputs a number V such that

$$.9 \operatorname{vol}(K) \le V \le 1.1 \operatorname{vol}(K) \tag{4.10}$$

holds with probability more than $\frac{1}{2} + Q(\frac{2}{e\pi n})^{n/2}$ has complexity at least Q.

Proof of Claim: To satisfy Equation (4.10), the algorithm has to actually distinguish between the brick and the double brick. Let the *bad surface* be the intersection between the input and the sphere of radius n. In order to distinguish between the two bodies, the algorithm has to make at least one query whose ray hits the bad surface. We will prove that the probability of this event is no more than $2Q(2/e\pi n)^{n/2}$. To see this, observe that the probability of a query hitting the bad surface is at most the volume of the bad surface divided by the volume of the sphere of radius n. The former can be bounded in the following way: Let $x = (x_2, \ldots, x_n)$ be the coordinates along the normals to the $n-1$ facets of the body. Parameterize one of the hemispheres determined by the hyperplane containing those normals as $F(x_2, \ldots, x_n) = \sqrt{n^2 - x_2^2 - \cdots - x_n^2}$.

We have that

$$\frac{d}{dx_i} F(x) = \frac{x_i}{F(x)}.$$

In the domain of integration $[-1,1]^{n-1}$ we have $\|x\|^2 \le n$ and this implies that in that domain

$$\|\nabla F(x)\|^2 = \frac{\|x\|^2}{n^2 - \|x\|^2} \le \frac{1}{n-1}.$$

The volume of the bad surface is given by

$$2 \int_{[-1,1]^{n-1}} \sqrt{1 + \|\nabla F(x)\|^2}\, dx \le 2^n \sqrt{1 + \frac{1}{n-1}} \le 2^{n+1}$$

The volume of the sphere of radius n is

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

Thus, the probability that a particular query hits the bad surface is at most

$$\frac{2^{n+1}}{(2e\pi n)^{n/2}} = 2\left(\frac{2}{e\pi n}\right)^{n/2}.$$

Therefore the algorithm gives the wrong answer with probability at least

$$\frac{1}{2}\left(1 - 2Q\left(\frac{2}{e\pi n}\right)^{n/2}\right).$$

□

4.4 Discussion

Related earlier work includes [6, 10], showing lower bounds for linear decision trees (i.e., every node of the tree tests whether an affine function of the input is nonnegative). [6] considers the problem of deciding whether given n real numbers, some k of them are equal, and they prove that it has complexity $\Theta(n \log(n/k))$. [10] proves that the n-dimensional knapsack problem has complexity at least $n^2/2$.

The results for determinant/volume hold with the following stronger oracle: we can specify any $k \times k$ submatrix A' of A and a vector $x \in \mathbb{R}^k$ and ask whether $\|A'x\|_\infty \leq 1$. In particular, this allows us to query individual entries of the matrix. More specifically, consider the oracle that takes indices i, j and $a \in \mathbb{R}$ and returns whether $A_{ij} \leq a$. Using this oracle, our proof (Lemma 3.7) yields the following result: there is a constant $c > 0$ such that any randomized algorithm that approximates the determinant to within a $(1 + c)$ factor has complexity $\Omega(n^2)$. In the property testing framework, this rules out sublinear (in the input size) methods for estimating the determinant, even with randomized (adaptive) access to arbitrary entries of the input matrix.

For the volume problem itself, the best known algorithm has complexity roughly $O(n^4)$ but the complexity of that algorithm is conjectured to be n^3. It is conceivable that our lower bound for membership oracle queries can be improved to n^3, although one would have to use bodies other than parallelotopes. Also, it is an open problem to give a faster algorithm using a separation oracle.

Chapter 5

Other lower bounds

Our lower bound for randomized volume approximation is nearly the best possible for our restricted class of parallelotopes. Using $O(n^2 \log n)$ queries, we can find a close approximation to the entire matrix A and therefore any reasonable function of its entries. This naturally raises the question of what other parameters require a quadratic number of queries. We prove that estimating the product of the lengths of the rows of an unknown matrix A to within a factor of about $(1+1/\log n)$ also requires $\Omega(n^2/\log n)$ queries. The simplest version of this problem is the following: given a membership oracle for any unknown halfspace $a \cdot x \leq 1$, estimate $\|a\|$, the Euclidean length of the normal vector a (alternatively, estimate the distance of the hyperplane from the origin). This problem can be solved deterministically using $O(n \log n)$ oracle queries. We prove that any randomized algorithm that estimates $\|a\|$ to within an additive error of about $1/\sqrt{\log n}$ requires $\Omega(n)$ oracle queries.

Also, in the previous chapter we saw a lower bound for the complexity of approximating the absolute value of the determinant of a matrix. A slightly weaker lower bound holds for estimating the product of the lengths of the rows. The proof is in Section 5.2.

We now state this lower bounds in a precise way.

Theorem 5.1 (product). *Let A be an unknown matrix that can be accessed by the following oracle: for any x, the oracle determines whether $\|Ax\|_\infty \leq 1$ is true or false. Then there exists a constant $c > 0$ such that any randomized algorithm that outputs a number L such that*

$$\left(1 - \frac{c}{\log n}\right) \prod_{i=1}^{n} \|A_i\| \leq L \leq \left(1 + \frac{c}{\log n}\right) \prod_{i=1}^{n} \|A_i\|$$

with probability at least $1 - 1/n$ has complexity $\Omega(n^2/\log n)$.

When A has only a single row, we get a stronger bound. In this case, the oracle is simply a membership oracle for a halfspace.

Theorem 5.2 (length). *Let a be a vector in $[-1,1]^n$ with $\|a\| \geq \sqrt{n} - 4\sqrt{\log n}$ and $a \cdot x \leq 1$ be the corresponding halfspace in $\mathbb{R}^n$ given by a membership oracle. Then* *there exists a constant $c > 0$ such that any randomized algorithm that outputs a number l such that*

$$\|a\| - \frac{c}{\sqrt{\log n}} \leq l \leq \|a\| + \frac{c}{\sqrt{\log n}}$$

with probability at least $1 - 1/n$ has complexity at least $n - 1$.

The restrictions on the input in all the above theorems ("roundness") only make them stronger. For example, the bound on the length of a above implies that it only varies in an interval of length $4\sqrt{\log n}$. To pin it down in an interval of length $c/\sqrt{\log n}$ (which is $O(\log\log n)$ bits of information) takes $\Omega(n)$ queries. This result is in the spirit of hardcore predicates [17].

It is worth noting that a very simple algorithm can approximate the length as in the theorem with probability at least $3/4$ and $O(n\log^2 n)$ queries: the projection of a onto a given vector b can be computed up to an additive error of $1/\operatorname{poly}(n)$ in $O(\log n)$ queries (binary search along the line spanned by b). If b is random in S_{n-1}, then $\mathbb{E}((a \cdot b)^2) = \|a\|^2/n$. A Chernoff-type bound gives that the average of $O(n \log n)$ random projections allows the algorithm to localize $\|a\|$ in an interval of length $O(1/\sqrt{\log n})$ with probability at least $3/4$.

5.1 Proof of the lower bound for length estimation

In this section, we prove Theorem 5.2. Let a be uniform random vector from $[-1, 1]^n$. By Lemma 1.6, $\|a\| \geq \sqrt{n} - 4\sqrt{\log n}$ as required by the theorem with probability at least $1 - 1/n^2$. We will prove that there exists a constant $c > 0$ such that any deterministic algorithm that outputs a number l such that

$$\|a\| - \frac{c}{\sqrt{\log n}} \leq l \leq \|a\| + \frac{c}{\sqrt{\log n}}$$

with probability at least $1 - O(1/n \log n)$ makes at least $n - 1$ halfspace queries. Along with Yao's lemma this proves the theorem.

Our access to a is via a membership oracle for the halfspace $a \cdot x \leq 1$. Consider the decision tree of height h for some deterministic algorithm. This will be a binary tree. The distribution at a leaf l is uniform over the intersection of $[-1, 1]^n$ with the halfspaces given by the path (queries, responses) to the leaf l from the root r, i.e., uniform over a polytope P_l with at most $2n + h$ facets.

The volume of the initial set is 2^n. The volume of leaves with $\operatorname{vol}(P_l) < 1$ is less than $|L| = 2^h$ and so the total volume of leaves with $\operatorname{vol}(P_l) \geq 1$ is at least $2^n - 2^h$. Setting $h = n - 1$, this is 2^{n-1} and so with probability at least $1/2$, $\operatorname{vol}(P_l) \geq 1$. For a random point X from any such P_l, Theorem 3.5 implies that $\operatorname{var} \|X\|^2 \geq cn/\log n$ for some absolute constant $c > 0$. Now by Lemma 3.3(a), and the fact that the support of $\|X\|^2$ is an interval of length n, we get that for any b,

$$\Pr\left(\left|\|X\|^2 - b\right| \geq \frac{1}{2}\sqrt{\frac{cn}{\log n}}\right) \geq \frac{3c}{4n\log n}.$$

It follows that $\|X\|$ is dispersed after $n-1$ queries. We note that the lower bound can be extended to any algorithm that succeeds with probability $1-1/n^\epsilon$ by a standard trick to boost the success probability: we repeat the algorithm $O(1/\epsilon)$ times and use the median of the results.

5.2 Proof of the lower bound for the product

Lemma 5.3. *Let $f:[0,M]\to\mathbb{R}_+$ be a density function with mean μ and variance σ^2. Suppose the distribution function of f is logconcave. Then f can be decomposed into a convex combination of densities g and h, i.e., $f(x)=\alpha g(x)+(1-\alpha)h(x)$, where g is uniform over an interval $[a,b]$, with $a\ge\mu$, $\alpha(a-b)^2=\Omega\big(\sigma^2/\log(M/\sigma)\big)$ and $\alpha=\Omega\big(\sigma^2/M^2\log(M/\sigma)\big)$.*

Proof. Let the distribution function be $F(t)=\Pr(X\le t)=e^{g(t)}$ for some concave function g and the density is $f(t)=g'(t)e^{g(t)}$ where $g'(t)$ is nonincreasing. First, we observe that logconcavity implies that $F(\mu)\ge 1/4$. To see this, let $\mu-l$ be the point where $F(\mu-l)=F(\mu)/2$. Then, $F(\mu-il)\le F(\mu)/2^i$ and

$$\begin{aligned}\int_0^\mu(\mu-x)f(x)\,dx &\le\sum_{i\ge1}\big(F(\mu-(i-1)l)-F(\mu-il)\big)(il)\\ &\le F(\mu)l+\sum_{i>1}F(\mu-il)\big((i+1)-i\big)l\\ &\le F(\mu)l\sum_{i\ge0}\frac{1}{2^i}=2lF(\mu).\end{aligned}$$

On the other hand (assuming $F(\mu)\le1/4$, otherwise, there is nothing to prove),

$$\int_\mu^\infty(x-\mu)f(x)\,dx\ge\sum_{i=1}^{\lfloor\log(1/F(\mu))\rfloor}(2^i-2^{i-1})F(\mu)(i-1)l\ge\frac{\log\big(1/F(\mu)\big)}{2}.$$

Therefore, we must have $2F(\mu)\ge\log(1/F(\mu))/2$ which implies $F(\mu)\ge1/4$.

Next,

$$\int_0^\mu(\mu-x)f(x)\,dx\ge\int_0^{\mu-l}(\mu-x)f(x)\,dx\ge F(\mu-l)l\ge\frac{l}{8}.$$

Therefore, since μ is the mean,

$$\int_\mu^\infty(x-\mu)f(x)\,dx\ge\frac{l}{8}.$$

It follows that

$$\int_\mu^\infty(x-\mu)^2f(x)\,dx\ge\frac{l^2}{64}.\tag{5.1}$$

Suppose $l < \sigma/4$. Then,

$$\begin{aligned}\int_0^{\mu} (x-\mu)^2 f(x)\,dx &\le \sum_{i\ge 1}\bigl(F(\mu-(i-1)l) - F(\mu-il)\bigr)(il)^2 \\ &\le F(\mu)l^2 + \sum_{i>1} F(\mu-il)\bigl((i+1)^2-i^2\bigr)l^2 \\ &\le F(\mu)l^2 \sum_{i\ge 1}\frac{2i+1}{2^i} = 5l^2F(\mu) \le \sigma^2/2.\end{aligned}$$

Since

$$\sigma^2 = \int_0^{\infty}(x-\mu)^2 f(x)\,dx = \int_0^{\mu}(x-\mu)^2 f(x)\,dx + \int_{\mu}^{\infty}(x-\mu)^2 f(x)\,dx,$$

we must have

$$\int_{\mu}^{\infty}(x-\mu)^2 f(x)\,dx \ge \frac{\sigma^2}{2}.$$

Using this and (5.1), we have (regardless of the magnitude of l),

$$\int_{\mu}^{\infty}(x-\mu)^2 f(x) \ge \frac{\sigma^2}{2^{10}}. \tag{5.2}$$

Now we consider intervals to the right of μ. Let $J_0 = (\mu, x_0]$ where x_0 is the smallest point to the right of μ for which $f(x_0) \le 1$ (J_0 could be empty). Let J_i, for $i = 1, 2, \ldots, m = 2\log(M/\sigma) + 12$ be $[x_{i-1}, x_i]$ where x_i is the smallest point for which $f(x_i) \le 1/2^i$. For any $t \ge t' \ge \mu$, $f(t') \ge f(t)F(t')/F(t) \ge f(t)F(\mu) \ge f(t)/4$. Therefore, the function f is approximately constant in any interval J_i for $i \ge 1$. If $x_0 > \mu + \sigma/64$, then we can use the interval $[\mu, \mu+\sigma/64]$. Otherwise, $\int_{J_0}(x-\mu)^2 f(x)\,dx \le \sigma^2/2^{12}$. Also, $\int_{x_m}^{\infty}(x-\mu)^2 f(x)\,dx \le \sigma^2/2^{12}$. Finally, consider intervals whose individual mass is less than

$$\frac{\sigma^2}{2^{12}M^2\bigl(2\log(M/\sigma)+12\bigr)}.$$

Their contribution to $\int_{\mu}^{\infty}(x-\mu)^2 f(x)\,dx$ is at most $\sigma^2/2^{12}$. Therefore, from (5.2), one of the remaining intervals J_i, $i \ge 1$, must have

$$\int_{J_i}(x-\mu)^2 f(x)\,dx \ge \frac{\sigma^2}{2^{12}\log(M/\sigma)} \text{ and } \int_{J_i} f(x)\,dx \ge \frac{\sigma^2}{2^{12}M^2\bigl(2\log(M/\sigma)+12\bigr)}.$$

□

Proof of Theorem 5.1. For this lower bound, we use the distribution D' on matrices. Let R be an $n \times n$ random matrix having each entry uniformly and independently in $[-1, 1]$. On input R from distribution D' having rows $(R_1, \ldots, R_n)$ and with probability at least 1/2 over the inputs, we consider algorithms that output an approximation to $f(R) = \prod_i \|R_i\|$. The next claim for deterministic algorithms, along with Yao's lemma, proves Theorem 5.1.

Claim: Suppose that a deterministic algorithm makes at most

$$h := \frac{\frac{n^2}{2} - 1}{\log_2(2n+1)}$$

queries on any input R and outputs V. Then there exists a constant $c > 0$ such that the probability of the event

$$\left(1 - \frac{c}{\log n}\right) f(R) \leq V \leq \left(1 + \frac{c}{\log n}\right)$$

is at most $1 - O(1/n)$.

To prove the claim, we consider a decision tree corresponding to a deterministic algorithm. Let P_l be the set of matrices associated with a leaf l. By Lemma 1.4, we have that the set P_l is a product set along rows, that is $P_l = \prod_i \mathcal{R}_i$, where $\mathcal{R}_i \subseteq \mathbb{R}^n$ is the set of possible choices of the row R_i consistent with l. The conditional distribution of R at a leaf l consists of *independent*, uniform choices of the rows from their corresponding sets. Moreover, the sets $\mathcal{R}_i$ are polytopes with at most $f = 2n + 2h$ facets. Every query has at most $2n+1$ different answers, and every path has height at most h. Thus, $|L| \leq (2n+1)^h = 2^{\frac{n^2}{2}-1}$. The total probability of the leaves having probability at most α is at most $\alpha|L|$. Thus, setting $\alpha = 1/(2|L|)$, the leaves having probability at least

$$\frac{1}{2|L|} \geq \frac{1}{2^{n^2/2}}$$

have total probability at least 1/2. Because $\operatorname{vol} \cup_{l \in L} P_l = 2^{n^2}$, we have that those leaves have volume at least $2^{n^2/2}$. Further, since $P_l = \prod_i \mathcal{R}_i$, we have that for such P_l at least $n/2$ of the $\mathcal{R}_i$'s have volume at least 1. Theorem 3.5 implies that for those $\operatorname{var} \|R_i\|^2 \geq \Omega(n/\log n)$. Along with the fact that $\|R_i\| \leq \sqrt{n}$ and Lemma 1.9, for a random matrix R from such a P_l, we get

$$\frac{\operatorname{var}\bigl(f(R)^2\bigr)}{\bigl(\mathbb{E}(f(R)^2)\bigr)^2} \geq \sum_i \frac{\operatorname{var}(\|R_i\|^2)}{\bigl(\mathbb{E}(\|R_i\|^2)\bigr)^2} = \Omega\left(\frac{1}{\log n}\right).$$

Thus, the variance of $f(R)$ is large. However, this does not directly imply that $f(R)$ is dispersed since the support of $f(R)$ could be of exponential length and its distribution is not logconcave.

Let $X = \prod_{i=1}^n X_i$ where $X_i = \|R_i\|^2$. To prove the lower bound, we need to show that $\operatorname{disp}_X(p)$ is large for p at least inverse polynomial in n. For i such that $\operatorname{vol}(\mathcal{R}_i) \geq 1$, we have $\operatorname{var} X_i = \Omega(n/\log n)$ by Theorem 3.5. As remarked earlier at least $n/2$ sets satisfy the volume condition and we will henceforth focus our attention on them. We also get $\mathbb{E}(X_i) \geq n/16$ from this. The distribution function of each X_i is logconcave (although not its density) and its support is contained in $[0, n]$. So by Lemma 5.3, we can decompose the density f_i of each X_i as $f_i(x) = p_i g_i(x) + (1 - p_i) g_i'(x)$. where g_i is the uniform distribution over an interval $[a_i, b_i]$ of length L_i and

$$p_i L_i^2 = \Omega\left(\frac{n}{\log^2 n}\right) \qquad \text{and} \qquad p_i = \Omega\left(\frac{1}{n \log n}\right).$$

We will assume that $p_i L_i^2 = cn/\log^2 n$ and $p_i = \Omega(1/n^2)$. This can be achieved by noting that L_i is originally at most n and truncating the interval suitably. Let X_i' be a random variable drawn uniformly from the interval $[a_i, b_i]$. Let $Y_i = \log X_i'$, I be a subset of $\{1, 2, \ldots, n\}$ and $Y_I = \sum_{i \in I} \log X_i'$. The density of Y_i is $h_i(t) = e^t/L_i$ for $\log a_i \le t \le \log b_i$ and zero outside this range. Thus Y_i has a logconcave density and so does Y_I (the sum of random variables with logconcave density also has a logconcave density). Also, $\text{var}(Y_I) = \sum_{i \in I} \text{var}(Y_i)$. To bound the variance of Y_i, we note that since $a_i \ge \mathbb{E}(X_i)$ by Lemma 5.3, we have $b_i \le 16 a_i$ and so $h_i(t)$ varies by a factor of at most 16. Thus, we can decompose h_i further into h_i' and h_i'' where h_i' is uniform over $[\log a_i, \log b_i]$ and

$$h_i(x) = \frac{1}{16} h_i'(x) + \frac{15}{16} h_i''(x).$$

Let Y_i' have density h_i'. Then

$$\text{var}(Y_i) \ge \frac{1}{16} \text{var}(Y_i') = \frac{(\log b_i - \log a_i)^2}{192}.$$

Therefore

$$\text{var}(Y_I) \ge \frac{1}{192} \sum_{i \in I} (\log b_i - \log a_i)^2$$

From this we get a bound on the dispersion of Y_I using the logconcavity of Y_I and Lemma 3.3(b). The bound depends on the set I of indices that are chosen. This set is itself a random variable defined by the decompositions of the X_i's. We have

$$\mathbb{E}_I\big(\text{var}(Y_I)\big) \ge \frac{1}{192} \sum_{i=1}^n p_i (\log b_i - \log a_i)^2 \ge \frac{1}{192} \sum_{i=1}^n p_i \frac{L_i^2}{(8a_i)^2} \ge \frac{c_1}{\log^2 n}$$

On the other hand,

$$\begin{aligned} \text{var}_I\big(\text{var}(Y_I)\big) &\le \frac{1}{144} \sum_{i=1}^n p_i (\log b_i - \log a_i)^4 \\ &\le \frac{1}{144} \sum_{i=1}^n p_i \frac{L_i^4}{a_i^4} \\ &\le \frac{16^4}{144 n^4} \sum_{i=1}^n \frac{p_i^2 L_i^4}{p_i} \\ &= \frac{16^4}{144 n^4} \frac{c^2 n^2}{\log^4 n} \sum_{i=1}^n \frac{1}{p_i}. \end{aligned}$$

Suppose $p_i \geq c_2/n$ for all i. Then we get,

$$\operatorname{var}_I\bigl(\operatorname{var}(Y_I)\bigr) \leq \frac{c_2'}{\log^4 n}$$

and for c_2 large enough, $\operatorname{var}_I\bigl(\operatorname{var}(Y_I)\bigr) \leq \bigl(\mathbb{E}_I \operatorname{var}(Y_I)\bigr)^2/4$. Hence, using Chebychev's inequality, with probability at least 1/4, $\operatorname{var}(Y_I) \geq c_1/4\log^2 n$. By Lemma 3.3(b), with probability at least 1/4, we have $\operatorname{disp}_{Y_I}(1/2) \geq \frac{\sqrt{c_1}}{4\log n}$. This implies that for any u,

$$\Pr\left(X \in \left[u, u\Bigl(1+\frac{\sqrt{c_1}}{4\log n}\Bigr)\right]\right) \leq \frac{7}{8}.$$

Finally, if for some i, $p_i < c_2/n$, then for that Y_i, $L_i^2 = \Omega(n^2/\log^2 n)$ and using just that i, we get $\operatorname{disp}_{Y_i}(p_i/2) \geq \sqrt{L_i^2/a_i^2} = \Omega(1/\log^2 n)$ and once again X is dispersed as well (recall that $p_i = \Omega(1/n^2)$). □

Chapter 6

Testing convexity

Geometric convexity has played an important role in algorithmic complexity theory. Fundamental problems (sampling, optimization, etc.) that are intractable in general can be solved efficiently with the assumption of convexity. The algorithms developed for these problems assume that the input is a convex set and are often not well-defined for arbitrary sets. Nevertheless, sampling-based approaches for optimization might be extendable to approximately convex sets, since there is hope that approximately convex sets can be sampled efficiently. This raises a basic question: How can we test if a given compact set in $\mathbb{R}^n$ is convex? Similarly, do short proofs of convexity or non-convexity of a set exist? Can one find these proofs efficiently?

To address these questions, we first need to decide how the set (called S henceforth) is specified. At the least, we need a membership oracle, i.e., a blackbox that takes as input a point $x \in \mathbb{R}^n$ and answers YES or NO to the question "Does x belong to S?" This is enough to prove that a set is not convex. We find 3 points $x, y, z \in \mathbb{R}^n$ such that $x, z \in S$, $y \in [x, z]$ and $y \notin S$. Since a set is convex iff it is convex along every line, such a triple constitutes a proof of non-convexity.

On the other hand, how can we prove that a set *is* convex? Imagine the perverse situation where a single point is deleted from a convex set. We would have to test an uncountable number of points to detect the non-convexity. So we relax the goal to determining if a set is approximately convex. More precisely, given $0 < \epsilon \leq 1$, either determine that S is not convex or that there is a convex set K such that

$$\mathrm{vol}(S \setminus K) + \mathrm{vol}(K \setminus S) \leq \epsilon\, \mathrm{vol}(S) \ .$$

In words, the condition above says that at most an ϵ fraction of S has to be changed to make it convex. We will call this the problem of testing approximate convexity.

This formulation of the problem fits the *property testing* framework developed in the literature [16]. In fact, there has been some work on testing convexity of discrete 1-dimensional functions [33], but the general problem is open.

Testing approximate convexity continues to be intractable if S is specified just by a membership oracle. Consider the situation where a small part of S is very far from the rest. How do we find it? To counter this, we assume that we also have access to uniform random points in S, i.e., a random oracle[1]. (There are other alternatives, but we find this to be the cleanest). In this chapter, we address the question of testing approximate convexity of a set given by a membership oracle and a random oracle. The complexity of an algorithm is measured by the number of calls to these oracles and the additional computation time.

We begin with a proof that the problem is well-defined, i.e., there exists a closest convex set. Then we give a simple algorithm with complexity $\text{poly}(n)(c/\epsilon)^n$ for any set S in $\mathbb{R}^n$. The algorithm uses random sampling from a convex polytope as a subroutine. Next, we consider what is perhaps the most natural algorithm for testing approximate convexity: get a pair of random points from the set and test if the intersection of the line through them with S is convex. This is motivated by the following conjecture: If the intersection of S with "most" lines is convex, then S itself is approximately convex. Many property testing algorithms in the literature have this flavor, i.e., get a random subset and test if the subset has the required property. Surprisingly, it turns out that the number of tests needed can be *exponential* in the dimension. We construct an explicit family of sets for which the lines through most (all but an exponentially small fraction) pairs of points have convex intersections with the set (i.e., they intersect S in intervals), yet the set is far from convex. Finally, we conjecture that if "most" 2-dimensional sections of a set S are convex, then S is approximately convex.

6.1 Preliminaries

The following notation will be used.

Let $A, B \subseteq \mathbb{R}^n$ be measurable sets. The *symmetric difference measure distance* (or simply, *distance*) between A and B is

$$d(A, B) = \text{vol}(A \Delta B) .$$

Let $\mathcal{K}$ denote the set of all compact convex sets in $\mathbb{R}^n$ with nonempty interior, and the empty set.

Proposition 6.1. *Let $S \subseteq \mathbb{R}^n$ compact. Then $\inf_{C \in \mathcal{K}} d(S, C)$ is attained.*

Proof. The set $\mathcal{K}$ with distance d is a metric space. The selection theorem of Blaschke (see the appendix at the end of this chapter) implies that $\{C \in \mathcal{K}, C \subseteq \text{conv}\, S\}$ is

[1] A non-trivial example where testing approximate convexity makes sense and the oracles are naturally available is testing approximate convexity of the union of m convex bodies given by membership oracles. In this case, the individual membership oracles give a membership oracle for the union. Also, the membership oracles can simulate random oracles for every convex set (approximately, see [27]), and allow us to approximate the volumes of the convex bodies. Finally, by using a technique similar to the one used to approximate the number of satisfying assignments of a DNF formula (see [32], for example), one can simulate a random oracle for the union (approximately) by means of the individual membership and random oracles and the individual volumes, in time polynomial in m and the other parameters.

compact. Moreover, $d(S,\cdot):\mathcal{K}\to\mathbb{R}$ is continuous. Also, it is sufficient to consider convex sets contained in $\operatorname{conv} S$, that is,

$$\inf_{C\in\mathcal{K}} d(S,C) = \inf_{\substack{C\in\mathcal{K}\\ C\subseteq \operatorname{conv} S}} d(S,C) \ .$$

The last expression is the infimum of a continuous function on a compact set, thus it is attained. □

Definition 6.2. *Given $S\subseteq\mathbb{R}^n$ compact, a set $C\in\operatorname{argmin}_{C\in\mathcal{K}} d(S,C)$ is called a* closest convex set *of S. S is said to be* ϵ-convex *iff $d(S,C)\le\epsilon\operatorname{vol}(S)$.*

6.2 Algorithms for testing approximate convexity

We are interested in the following algorithmic problem:

Let $S\subseteq\mathbb{R}^n$ be compact. We are given a *membership oracle* that given $x\in\mathbb{R}^n$ answers "YES" if $x\in S$ and "NO" if $x\notin S$; we also have access to a *random oracle* that when called gives a uniformly sampled random point from S. For any given $\epsilon>0$, our goal is to determine either that S is ϵ-convex (output "YES") or that S is not convex (output "NO").

In this section, we will give a randomized algorithm for the problem. We will prove that the algorithm works with probability at least $3/4$. This can be easily boosted to any desired $1-\delta$ while incurring an additional factor of $O(\ln(1/\delta))$ in the complexity.

6.2.1 The one-dimensional case

One-dimensional algorithm
INPUT: Access to membership and random oracles of $S\subseteq\mathbb{R}$.

1. Get $12/\epsilon$ points from the random oracle. Let C be their convex hull (the interval containing them).

2. Choose $12/\epsilon$ random points in C. Check if they are all in S using the membership oracle. If so, output "YES", else output "NO".

Theorem 6.3. *With probability at least $3/4$, the one-dimensional algorithm determines that S is not convex or that S is ϵ-convex.*

Proof. Clearly, if S is convex then the algorithm answers "YES". So assume that S is not ϵ-convex. We say that the first step succeeds if we get at least one point in the leftmost $\epsilon/4$ fraction of S and another point in the rightmost $\epsilon/4$ fraction of S. The first step fails with probability at most $2(1-\epsilon/4)^{12/\epsilon}\le 2/e^3$. Suppose the first step succeeds. Then,

$$\operatorname{vol}(S\setminus C)\le\operatorname{vol}(S)\frac{\epsilon}{2}\ .$$

This implies that

$$\operatorname{vol}(C\setminus S)\ge\operatorname{vol}(S)\frac{\epsilon}{2}\ .$$

From this, we get

$$\begin{aligned} \operatorname{vol}(C \setminus S) &\geq \max\left\{\frac{\epsilon}{2}\operatorname{vol}(S), \operatorname{vol}(C) - \operatorname{vol}(S)\right\} \\ &= \operatorname{vol}(C)\max\left\{\frac{\epsilon}{2}\frac{\operatorname{vol}(S)}{\operatorname{vol}(C)}, 1 - \frac{\operatorname{vol}(S)}{\operatorname{vol}(C)}\right\} . \end{aligned} \tag{6.1}$$

Given that $\epsilon > 0$, the expression

$$\max\left\{\frac{\epsilon}{2}\alpha, 1 - \alpha\right\}$$

is minimized as a function of α when $\frac{\epsilon}{2}\alpha = 1 - \alpha$, i.e., for $\alpha = \frac{2}{\epsilon+2}$. Thus, from Equation (6.1) we get

$$\operatorname{vol}(C \setminus S) \geq \frac{\epsilon}{2+\epsilon}\operatorname{vol}(C) .$$

That is, conditioned on the success of the first step, with probability at least $1 - (1 - \epsilon/3)^{12/\epsilon} \geq 1 - 1/e^4$ the algorithm answers "NO". Thus, overall the algorithm answers "NO" with probability at least $(1 - 1/e^4)(1 - 2/e^3) \geq 3/4$. □

6.2.2 The general case

Here we consider the problem in $\mathbb{R}^n$. It is not evident that the time complexity of the problem can be made independent of the given set S (that is, depending only on ϵ and the dimension). The following algorithm shows such independence ($m = m(\epsilon, n)$ will be chosen later).

```
$n$-dimensional algorithm
INPUT: Access to membership and random oracles of $S \subseteq \mathbb{R}^n$.

  1. Get $m$ random points from $S$. Let $C$ be their convex hull.
  2. Get $4/\epsilon$ random points from $S$. If any of them is not in $C$, output "NO".
  3. Get $6/\epsilon$ random points from $C$. If each of them is in $S$ according to the
     membership oracle, then output "YES", else output "NO".
```

Checking if a point y belongs to C is the same as answering whether y can be expressed as a convex combination of the m points that define C. This can be done by solving a linear program. The third step requires random points from C, which is a convex polytope. Sampling convex bodies is a well-studied algorithmic problem and can be done using $O^*(n^3)$ calls to a membership oracle (see [27], for example).

To prove the correctness of the algorithm we will use the following lemmas (the first is from [18] and the second is paraphrased from [3]).

Lemma 6.4. *Let $C = \operatorname{conv}\{X_1, \dots, X_m\}$, where the X_i's are independent uniform random samples from a convex body K. Then for any integer $t > 0$ we have $\mathbb{E}\big((\operatorname{vol}(C)/\operatorname{vol}(K))^t\big)$ is minimized iff K is an ellipsoid.*

Lemma 6.5. *Let $B_n \subseteq \mathbb{R}^n$ be the unit ball. Let $C = \operatorname{conv}\{X_1, \dots, X_m\}$, where the X_i's are independent uniform random samples from B_n. There exists a constant c such that, for $m = (cn/\epsilon)^n$,*

$$\mathbb{E}\big(\operatorname{vol}(B_n \setminus C)\big) \le \epsilon \operatorname{vol}(B_n) \ .$$

Theorem 6.6. *Using $m = (224cn/\epsilon)^n$ random points and* $\operatorname{poly}(n)/\epsilon$ *membership calls, the n-dimensional algorithm determines with probability at least* $3/4$ *that S is not convex or that S is ϵ-convex.*

Proof. First, assume that S is convex. We want to show that the algorithm outputs "YES" with probability at least $3/4$. Let $X = \operatorname{vol}(S \setminus C)/\operatorname{vol}(S)$. Then by Lemma 6.4, $\mathbb{E}(X)$ is maximized when K is a ball and using Lemma 6.5 with our choice of m, we get that

$$\mathbb{E}(X) \le \frac{\epsilon}{224n} \ .$$

By Markov's inequality, with probability at least $6/7$,

$$\operatorname{vol}(S \setminus C) \le \frac{\epsilon}{32} \operatorname{vol}(S) \ .$$

Given this, Markov's inequality implies that the algorithm will not stop at step 2 with probability at least $3/4$: in step 2, if we let Y be the number of points not in C then

$$\mathbb{E}(Y) \le \frac{\epsilon}{32}\frac{4}{\epsilon} = \frac{1}{8} \ ,$$

and therefore, by Markov's inequality,

$$\mathbb{P}(\text{algorithm outputs “NO” in step 2}) = \mathbb{P}(Y \ge 1) = \mathbb{P}\big(Y \ge 8\,\mathbb{E}(Y)\big) \le \frac{1}{8} \ .$$

Thus, the algorithm outputs "YES" with probability at least $\frac{6}{7}\frac{7}{8} = \frac{3}{4}$.

Next, if S is not ϵ-convex, the analysis can be divided into two cases after the first step: either $\operatorname{vol}(S \setminus C) \ge \operatorname{vol}(S)\epsilon/2$ or $\operatorname{vol}(S \setminus C) < \operatorname{vol}(S)\epsilon/2$. In the first case, step 2 outputs "NO" with probability at least $1 - \big(1 - \frac{\epsilon}{2}\big)^{4/\epsilon} \ge 1 - \frac{1}{e^2} \ge \frac{3}{4}$. In the second case we have

$$\operatorname{vol}(C \setminus S) \ge \frac{\epsilon}{2} \operatorname{vol}(S)$$

and by the same analysis as the one-dimensional case, $\operatorname{vol}(C \setminus S) \ge \frac{\epsilon}{3} \operatorname{vol}(C)$. Thus, step 3 outputs "NO" with probability at least $1 - (1 - \frac{\epsilon}{3})^{6/\epsilon} \ge 3/4$. □

Note that, unlike the one-dimensional case, this algorithm has two-sided error. The complexity of the algorithm is independent of S and depends only on n and ϵ. It makes an exponential number of calls to the random oracle and this dependency is unavoidable for this algorithm. It is known for example that the convex hull of any subset of fewer than c^n points of the ball, contains less than half its volume [4].

The one-dimensional algorithm suggests another algorithm for the general case: let $\ell(x, y)$ be the line through x and y,

Lines-based algorithm
INPUT: Access to membership and random oracles of $S \subseteq \mathbb{R}^n$ compact.

Generate m pairs of random points (x, y) and test if $\ell(x, y) \cap S$ is convex.

How large does m need to be? Somewhat surprisingly, we show in the next section that this algorithm also has an exponential complexity. Testing if $\ell(x, y) \cap S$ is convex is not a trivial task (note that we have a membership oracle for $\ell(x, y) \cap S$ from the oracle for S, but simulating a random oracle is not so simple). However, for the purpose of showing a lower bound in m we will assume that the one-dimensional algorithm checks *exactly* whether $\ell(x, y) \cap S$ is convex (that is, it is an interval).

6.3 The lines-based algorithm is exponential

In this section, we construct an explicit family of compact sets each of which has the following properties: (i) the set is far from convex, and (ii) for all but an exponentially small fraction of pairs of points from the set, the line through the pair of points has a convex intersection with the set. This implies that the lines-based algorithm (described at the end of Subsection 6.2.2) has exponential worst-case complexity. Thus, although exact convexity is characterized by "convex along every line," the corresponding reduction of approximate convexity to "convex along most lines" is not efficient.

The proof of the lower bound is in two parts, first we show that the algorithm needs many tests and then that the test family is far from convex (i.e., ϵ is large).

6.3.1 The family of sets: the cross-polytope with peaks

The n-dimensional cross-polytope is an n-dimensional generalization of the octahedron and can be defined as the unit ball with the norm $\|x\|_1 = \sum_{i=1}^n |x_i|$. Let T_n be the "cross-polytope with peaks", that is, the union of the cross-polytope and, for each of its facets $i \in \{1, \dots, 2^n\}$, the convex hull of the facet and a point $v_i = \lambda d$, where d is the unit outer normal to the facet and $\lambda \geq 1/\sqrt{n}$ is a parameter (that may depend on the dimension). Informally, one adds an n-dimensional simplex on top of each facet of the cross-polytope. The volume of the cross-polytope is a $\frac{1}{\lambda\sqrt{n}}$ fraction of the volume of T_n. We will choose $\lambda = \frac{\sqrt{n}}{n-2}$. In that case, the cross-polytope as a convex set shows that T_n is $O(\frac{1}{n})$-convex. We will prove that T_n is not $\frac{1}{18n^2}$-convex, i.e., for any convex set K, we have $d(K, T_n) > \frac{1}{18n^2} \operatorname{vol} T_n$.

6.3.2 The non-convexity of the family cannot be detected by the lines-based algorithm

Proposition 6.7. *If $\lambda \leq \frac{\sqrt{n}}{n-2}$ then the one-dimensional test has an exponentially low probability of detecting the non-convexity of the cross-polytope with peaks.*

Proof. First, we will prove the following claim:

> Under the hypothesis, every peak is contained in the intersection of the half-spaces determining the n facets of the cross-polytope adjacent to the peak.

It is enough to see that the point $v_i = \lambda d$ (a vertex of the peak) is contained in that intersection. Because of the symmetry, we can concentrate on any particular pair of adjacent facets, say those having normals $d = (1, 1, \dots, 1)/\sqrt{n}$ and $d' = (-1, 1, \dots, 1)/\sqrt{n}$. The halfspace determining the facet with normal d is given by $\{x \in \mathbb{R}^n : x \cdot d' \leq 1/\sqrt{n}\}$. Then $v_i = \lambda d$ is contained in the halfspace associated to the facet with normal d' (which is sufficient) if $\lambda d \cdot d' \leq 1/\sqrt{n}$. That is, $\lambda \leq \sqrt{n}/(n-2)$. This proves the claim.

It is sufficient to note that, for the algorithm to answer "NO", we need to choose a line whose intersection with T_n is not convex. Suppose that a line L shows non-convexity. Then it does not intersect the cross-polytope part of T_n a. s. (almost surely), otherwise L intersects exactly 2 facets of the cross-polytope a. s., and intersects only the peaks that are associated to those facets, because of the claim (if one follows the line after it leaves the cross-polytope through one of the facets, it enters a peak, and that peak is the only peak on that side of the facet, because of the claim), and thus $L \cap T_n$ would be convex. Now, while intersecting a peak, L intersects two of its facets at two points that are not at the same distance of the cross-polytope, a.s. The half of L that leaves the peak through the farthest point cannot intersect any other peak because of the claim (the halfspace determined by the respective facet of the cross-polytope containing this peak contains only this peak, and this half of L stays in this halfspace). The half of L that leaves the peak through the closest point will cross the hyperplane determined by one of the adjacent peaks[2] before intersecting any other peak, a.s.; after crossing that hyperplane it can intersect only one peak, namely, the peak associated to that hyperplane, because of the claim. Thus, L has to intersect exactly 2 peaks that have to be adjacent a. s., and L does not intersect the cross-polytope. In other words, the two random points that determine L are in the same peak or in adjacent peaks. The probability of this event is no more than $\frac{n+1}{2^n}$. □

6.3.3 The sets in the family are far from convex

To prove that T_n is far from being convex, we will prove that a close convex set must substantially cover most peaks, and because of this, a significant volume of a close convex set must lie between pairs of adjacent substantially covered peaks, outside of T_n, adding to the symmetric difference. The following lemma will be useful for this part. For $A \subseteq \mathbb{R}^n$, H a hyperplane and $v \in \mathbb{R}^n$ a unit normal for H, let

$$w_H(A) = \sup_{x \in A} v \cdot x - \inf_{x \in A} v \cdot x \ .$$

[2] "The hyperplane determined by a peak" is the unique hyperplane that contains the facet of the cross-polytope associated to the peak.

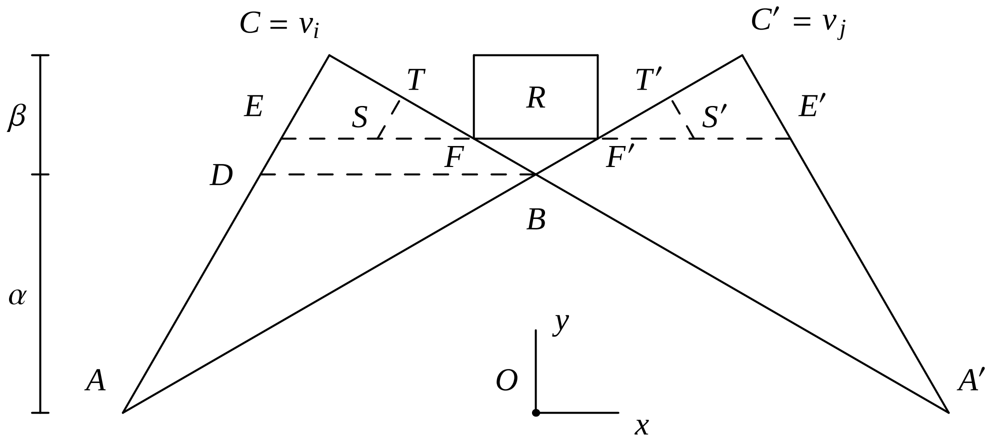

Figure 6-1: Projection of the peaks (i, j) of the cross-polytope with peaks onto v_i, v_j for $n = 4$.

Lemma 6.8. *Let $A, B \subseteq \mathbb{R}^n$ compact. Let H be a separating hyperplane*[3] *for A, B. Let $C = H \cap \operatorname{conv}(A \cup B)$. Then*

$$V_{n-1}(C) \geq \min\left\{\frac{\operatorname{vol} A}{w_H(A)}, \frac{\operatorname{vol} B}{w_H(B)}\right\} .$$

Proof. There exist sections, parallel to H, of A and B that have $(n-1)$-volumes at least $(\operatorname{vol} A)/w_H(A)$ and $(\operatorname{vol} B)/w_H(B)$, respectively. That is, there exist $a, b \in \mathbb{R}^n$ such that $A' = (H + a) \cap A$, $B' = (H + b) \cap B$ satisfy $V_{n-1}(A') \geq (\operatorname{vol} A)/w_H(A)$ and $V_{n-1}(B') \geq (\operatorname{vol} B)/w_H(B)$. Clearly $H \cap \operatorname{conv}(A' \cup B') \subseteq C$ and therefore

$$\begin{aligned} V_{n-1}(C) &\geq V_{n-1}(H \cap \operatorname{conv}(A' \cup B')) \\ &\geq \min\{V_{n-1}(A'), V_{n-1}(B')\} \\ &\geq \min\left\{\frac{\operatorname{vol} A}{w_H(A)}, \frac{\operatorname{vol} B}{w_H(B)}\right\} . \end{aligned}$$

□

This bound is sharp: consider a cylinder with a missing slice in the middle, that is, consider in the plane as A a rectangle with axis-parallel sides and non-adjacent vertices $(1, 0)$ and $(2, 1)$, as B the reflection of A with respect to the y-axis and as the separating line, the y-axis.

Lemma 6.9. *For $\lambda = \frac{\sqrt{n}}{n-2}$, T_n is not $\frac{1}{18n^2}$-convex.*

Proof. Let C_n be a closest convex set to T_n.

Consider a pair of adjacent peaks (i, j). Figure 6-1 shows the projection of the pair onto the plane containing the vertices v_i, v_j and the origin. B is the projection

[3] That is, a set of the form $H = \{x \in \mathbb{R}^n : x \cdot y = \alpha\}$ for some $y \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$, such that for all $x \in A$ we have $x \cdot y \leq \alpha$ and for all $x \in B$ we have $x \cdot y \geq \alpha$.

of the intersection of the two peaks, an $(n-2)$-dimensional simplex. Points A and C are the other two vertices of one of the peaks, A' and C' are the respective vertices of the other peak. The plane is orthogonal to the two respective facets of the cross-polytope, the segment AB is the projection of one of them and $A'B$ is the projection of the other facet. D is such that DB is orthogonal to OB, where O is the origin.

The idea of the proof is that if C_n is close to T_n, then it covers a good fraction of the preimage (with respect to the projection) of the triangles SFT and $S'F'T'$. Lemma 6.8 applied to this covered parts implies that a certain fraction of C_n lies in the preimage of R, contributing to $C_n \setminus T_n$, that is, to the symmetric difference, and this is happening between most pairs of peaks.

First, we will compute α and β. Without loss of generality we can assume that v_i is parallel to $(-1, 1, \dots, 1)$ and v_j is parallel to $(1, \dots, 1)$. Then $(0, \frac{1}{n-1}, \dots, \frac{1}{n-1})$ is a vector in the preimage of B that is in the projection plane, and α is the norm of that vector, that is, $\alpha = 1/\sqrt{n-1}$. An orthonormal basis of the projection plane corresponding to the x, y axes of Figure 6-1 is

$$\{(1, 0, \dots, 0), (0, 1/\sqrt{n-1}, \dots, 1/\sqrt{n-1})\} \,.$$

Then, $\alpha + \beta$ is the length of the projection of v_j onto $(0, 1/\sqrt{n-1}, \dots, 1/\sqrt{n-1})$, that is, $\alpha + \beta = \frac{\sqrt{n-1}}{n-2}$ and $\beta = \frac{1}{(n-2)\sqrt{n-1}}$.

EF is a segment parallel to DB and at a distance $\beta\delta$ from it (δ will be chosen later). S is a point on EF such that the triangle SFT is a scaled and translated version of EFC, with scaling factor δ.

Let $\overline{DB}$, $\overline{AA'}$ be the lengths of the segments DB, AA', respectively. We have that $\overline{DB} = \overline{AA'}\frac{\beta}{\alpha+\beta} = 2/(n-1)$, and $V_{n-2}(\text{preimage of } B)$ as a fraction of the volume of a peak is given by the following identity:

$$\begin{aligned} \text{vol(one peak)} &= \frac{1}{n}(\alpha+\beta)V_{n-1}(\text{preimage of } DB) \\ &= \frac{1}{n}(\alpha+\beta)\frac{1}{n-1}V_{n-2}(\text{preimage of } B)\overline{DB} \,, \end{aligned}$$

that is, $V_{n-2}(\text{preimage of } B)$ is a $\frac{n(n-1)^2(n-2)}{2\sqrt{n-1}}$ fraction of the volume of a peak.

The preimage of any point in SFT has volume at least $V_{n-2}(\text{preimage of } B)(1-\delta)^{2(n-2)}$. The area of SFT is $[(1-\delta)\delta]^2\beta\overline{DB}/2$. Then,

$$\text{vol(preimage of } SFT) \geq V_{n-2}(\text{preimage of } B)(1-\delta)^{2(n-2)}[(1-\delta)\delta]^2\frac{\beta\overline{DB}}{2},$$

that is, fixing $\delta = 1/n$ and using the previously computed values of β, $\overline{DB}$ and $V_{n-2}(\text{preimage of } B)$,

$$\begin{aligned} \text{vol(preimage of } SFT) &\geq \text{vol(peak)}\left(1 - \frac{1}{n}\right)^{2n-2}\frac{1}{2n} \\ &\geq \text{vol(peak)}\frac{1}{2ne^2} \,. \end{aligned} \tag{6.2}$$

Given a particular peak, we will say that it is *substantially covered* (by C_n) iff the volume of the intersection of C_n and the peak is at least a $1 - \frac{1}{4ne^2}$ fraction of the volume of the peak. Thus, by means of Equation (6.2), if a peak is substantially covered, then at least a $\frac{1}{4ne^2}$ fraction of its volume is covered in the preimage of the triangle SFT.

Now we will prove that every pair of adjacent substantially covered peaks contributes to $C_n \setminus T_n$ with at least a $\frac{1}{4ne^2}$ fraction of the volume of a peak, disjoint from the contribution of other pairs. To see this, let U be the subset of C_n intersected with peak i that projects onto SFT and let V be the subset of C_n intersected with peak j that projects onto $F'S'T'$. We will apply Lemma 6.8 to U, V and every hyperplane which is a preimage of a vertical line intersecting the rectangle R. Moreover, for any such hyperplane H we have that $w_H(U)$ and $w_H(V)$ are no more than $\delta\overline{DB} = \frac{2}{n(n-1)}$. Certainly $W = R \cap \operatorname{conv}(U \cup V)$ is contained in C_n and disjoint from T_n. Because of the choice of EF, the width of the rectangle R is a $1/n$ fraction of the distance between C and C', that is, $\frac{2}{n(n-2)}$. Also, $\operatorname{vol} U$ and $\operatorname{vol} V$ are no less than a $\frac{1}{4ne^2}$ fraction of the volume of a peak. Lemma 6.8 gives that

$$\begin{aligned}\frac{\operatorname{vol} W}{\operatorname{vol}(\text{one peak})} &\geq (\text{width of } R) \min\left\{\frac{\operatorname{vol} U}{\frac{2}{n(n-1)}}, \frac{\operatorname{vol} V}{\frac{2}{n(n-1)}}\right\} \frac{1}{\operatorname{vol}(\text{one peak})} \\ &\geq \frac{2}{n(n-2)} \frac{n(n-1)}{2} \frac{1}{4ne^2} = \frac{n-1}{n-2} \frac{1}{4ne^2} \\ &\geq \frac{1}{4ne^2} \ .\end{aligned}$$

Let $\epsilon(n) = d(C_n, T_n)$. We claim that the number of peaks that are not substantially covered is a fraction that is at most $en^2\epsilon(n)$ of the total number of peaks. To see this, let $q(n)$ be the fraction of the volume of T_n that the peaks contain. Clearly

$$q(n) = \frac{\lambda - \frac{1}{\sqrt{n}}}{\lambda} = \frac{2}{n} \ .$$

Let X be fraction of the number of peaks that are not substantially covered. Then,

$$X \frac{1}{4ne^2} q(n) \leq \epsilon(n) \ ,$$

that is,

$$X \leq 2n^2e^2\epsilon(n) \ . \tag{6.3}$$

We will see now that eventually (as n grows) the number of pairs of adjacent peaks that are substantially covered is a substantial fraction of the total number of adjacent pairs. For a contradiction, assume that, for some subsequence, $\epsilon(n) < \frac{1}{18n^2}$. For that subsequence, $2n^2e^2\epsilon(n) < 1/4$. The number of peaks is 2^n; the number of (unordered) pairs of adjacent peaks is $n2^{n-1}$. A peak that is not substantially covered can participate in at most n pairs of adjacent peaks. Because of (6.3), there are at most $\frac{1}{4}2^n = 2^{n-2}$ peaks that are not substantially covered (for the subsequence).

That way, all the peaks that are not substantially covered can participate in at most $n2^{n-2} = \frac{1}{2}n2^{n-1}$ pairs of adjacent peaks. Thus, at least half of the pairs of adjacent peaks involve only substantially covered peaks. For γ equal to the volume of the contribution to $C_n \setminus T_n$ of a pair of substantially covered peaks, this implies that

$$\begin{aligned}\epsilon(n) &\geq \frac{\operatorname{vol}(C_n \setminus T_n)}{\operatorname{vol} T_n} = \frac{\operatorname{vol}(C_n \setminus T_n)}{\operatorname{vol}(\text{all peaks})}\frac{\operatorname{vol}(\text{all peaks})}{\operatorname{vol} T_n} \\ &\geq \frac{\frac{1}{2}n2^{n-1}\gamma}{2^n \operatorname{vol}(\text{one peak})}q(n) \geq \frac{n}{4}\frac{1}{4ne^2}\frac{2}{n} \\ &= \frac{1}{8ne^2}\end{aligned}$$

which is a contradiction. □

6.4 An algorithm based on planes

In this section, we state a conjecture about approximate convexity. Let S be a compact subset of $\mathbb{R}^n$ whose center of gravity is the origin. For a pair of points $x, y \neq 0$ in $\mathbb{R}^n$ let the subspace spanned by them be $H(x, y)$ and define $P(x, y) = S \cap H(x, y)$ to be the part of S on this subspace. Our conjecture is the following:

Conjecture. Let μ be the distribution on 2-dimensional sections $P(x, y)$ obtained by picking x and y uniformly at random from S. If

$$\mathbb{P}_\mu\big(P(x, y) \text{ is convex}\big) > 1 - \epsilon\ ,$$

then S is $O(n\epsilon)$-convex.

The conjecture motivates the following algorithm (here $p(\cdot)$ and $q(\cdot)$ are fixed polynomials):

Repeat $p(n, 1/\epsilon)$ times

1. Get random points x, y from S.

2. Test if $P(x, y)$ is $q(1/n, \epsilon)$-convex.

Appendix

For the Hausdorff metric or the symmetric difference volume metric, we have (see [34], Theorem 4.18, for example):

Theorem 6.10 (Blaschke's selection theorem)**.** *In $\mathbb{R}^n$, any bounded sequence $(C_k)_{k\in\mathbb{N}}$ of nonempty, convex sets has a subsequence converging to some nonempty, compact, convex set C.*

Bibliography

[1] D. Applegate and R. Kannan. Sampling and integration of near log-concave functions. *Proceedings of the twenty-third annual ACM symposium on Theory of computing*, pages 156–163, 1991.

[2] K. Ball. Normed spaces with a weak-Gordon-Lewis property. In *Functional analysis (Austin, TX, 1987/1989)*, volume 1470 of *Lecture Notes in Math.*, pages 36–47. Springer, Berlin, 1991.

[3] I. Bárány and C. Buchta. Random polytopes in a convex polytope, independence of shape, and concentration of vertices. *Math. Ann.*, 297(3):467–497, 1993.

[4] I. Bárány and Z. Füredi. Computing the volume is difficult. *Discrete Comput. Geom.*, 2(4):319–326, 1987.

[5] D. Bertsimas and S. Vempala. Solving convex programs by random walks. *J. ACM*, 51(4):540–556, 2004.

[6] A. Björner, L. Lovász, and A. C. C. Yao. Linear decision trees: volume estimates and topological bounds. In *STOC '92: Proceedings of the twenty-fourth annual ACM symposium on Theory of computing*, pages 170–177, New York, NY, USA, 1992. ACM Press.

[7] S. G. Bobkov and A. Koldobsky. On the central limit property of convex bodies. In *Geometric aspects of functional analysis*, volume 1807 of *Lecture Notes in Math.*, pages 44–52. Springer, Berlin, 2003.

[8] J. Bourgain. On the distribution of polynomials on high dimensional convex sets. *Lecture Notes in Mathematics*, 1469:127–137, 1991.

[9] G. Brightwell and P. Winkler. Counting linear extensions. *Order*, 8(3):225–242, 1991.

[10] D. Dobkin and R. Lipton. A lower bound of $\frac{1}{2}n^2$ on linear search programs for the knapsack problem. *J. Comput. Syst. Sci.*, 16:413–417, 1978.

[11] M. Dyer and A. Frieze. On the complexity of computing the volume of a polytope. *SIAM J. Comput*, 17:967–974, 1988.

[12] M. Dyer and A. Frieze. Computing the volume of convex bodies: a case where randomness provably helps. Probabilistic Combinatorics and its applications. *Proceedings of the AMS Symposia in Applied Mathematics*, 44:123–170, 1991.

[13] M. Dyer, A. Frieze, and R. Kannan. A random polynomial-time algorithm for approximating the volume of convex bodies. *J. Assoc. Comput. Mach.*, 38(1):1–17, 1991.

[14] A. Edelman. Eigenvalues and condition numbers of random matrices. *SIAM Journal on Matrix Analysis and Applications*, 9:543, 1988.

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

[16] O. Goldreich, S. Goldwasser, and D. Ron. Property testing and its connection to learning and approximation. *J. ACM*, 45(4):653–750, 1998.

[17] O. Goldreich and L. A. Levin. A hard-core predicate for all one-way functions. *Proceedings of the Twenty-First Annual ACM Symposium on Theory of Computing*, pages 25–32, 1989.

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

[19] M. R. Jerrum, L. G. Valiant, and V. V. Vazirani. Random generation of combinatorial structures from a uniform distribution. *Theoret. Comput. Sci.*, 43(2-3):169–188, 1986.

[20] M. J. Kaiser, T. L. Morin, and T. B. Trafalis. Centers and invariant points of convex bodies. In *Applied geometry and discrete mathematics*, volume 4 of *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.*, pages 367–385. Amer. Math. Soc., Providence, RI, 1991.

[21] R. Kannan, L. Lovász, and M. Simonovits. Isoperimetric problems for convex bodies and a localization lemma. *Discrete Comput. Geom.*, 13(3-4):541–559, 1995.

[22] R. Kannan, L. Lovász, and M. Simonovits. Random walks and an $O^*(n^5)$ volume algorithm for convex bodies. *Random Structures and Algorithms*, 11(1):1–50, 1997.

[23] J. Lawrence. Polytope volume computation. *Mathematics of Computation*, 57(195):259–271, 1991.

[24] L. Lovász. How to compute the volume? Jber. d. Dt. Math. *Verein, Jubilaumstagung*, pages 138–151, 1990.

[25] L. Lovász and M. Simonovits. The mixing rate of Markov chains, an isoperimetric inequality, and computing the volume. In *31st Annual Symposium on Foundations of Computer Science, Vol. I, II (St. Louis, MO, 1990)*, pages 346–354. IEEE Comput. Soc. Press, Los Alamitos, CA, 1990.

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

[27] L. Lovász and S. Vempala. Hit-and-run from a corner. In *STOC '04*, pages 310–314. ACM, New York, 2004.

[28] L. Lovász and S. Vempala. Fast algorithms for logconcave functions: Sampling, rounding, integration and optimization. *Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS'06)*, pages 57–68, 2006.

[29] L. Lovász and S. Vempala. Simulated annealing in convex bodies and an $O^*(n^4)$ volume algorithm. *Journal of Computer and System Sciences*, 72(2):392–417, 2006.

[30] L. Lovász and S. Vempala. The geometry of logconcave functions and sampling algorithms. *Random Structures and Algorithms*, 30(3):307–358, 2007.

[31] V. Milman and A. Pajor. Isotropic position and inertia ellipsoids and zonoids of the unit ball of a normed n-dimensional space. *Lecture Notes in Mathematics*, 1376:64–104, 1989.

[32] R. Motwani and P. Raghavan. *Randomized Algorithms*. Cambridge University Press, 1995.

[33] M. Parnas, D. Ron, and R. Rubinfeld. On testing convexity and submodularity. *SIAM J. Comput.*, 32(5):1158–1184 (electronic), 2003.

[34] R. T. Rockafellar and R. J. Wets. *Variational Analysis*. Number 317 in Grundlehren der mathematischen Wissenschaften. Springer-Verlag, 1998.

[35] M. Simonovits. How to compute the volume in high dimension? *Mathematical Programming*, 97(1):337–374, 2003.

[36] S. Vempala. Geometric Random Walks: A Survey. *MSRI volume on Combinatorial and Computational Geometry*, 2005.

[37] S. S. Vempala. *The Random Projection Method*. American Mathematical Society, 2004.