




How to Compute the Volume in Low Dimension?

Arjan Cornelissen   

Simons Institute, UC Berkeley, CA, USA
Université Paris Cité, CNRS, IRIF, Paris, France

Simon Apers   

Université Paris Cité, CNRS, IRIF, Paris, France

Sander Gribling   

Tilburg University, The Netherlands

Abstract

Estimating the volume of a convex body is a canonical problem in theoretical computer science. Its study has led to major advances in randomized algorithms, Markov chain theory, and computational geometry. In particular, determining the query complexity of volume estimation to a membership oracle has been a longstanding open question. Most of the previous work focuses on the high-dimensional limit. In this work, we tightly characterize the deterministic, randomized and quantum query complexity of this problem in the high-precision limit, i.e., when the dimension is constant.

2012 ACM Subject Classification Theory of computation → Quantum query complexity; Theory of computation → Computational geometry

Keywords and phrases Query complexity, computational geometry, quantum computing, volume estimation, high-precision limit

Digital Object Identifier 10.4230/LIPIcs.ICALP.2025.61

Category Track A: Algorithms, Complexity and Games

Related Version *Full Version*: <https://arxiv.org/abs/2503.02207>

Funding *Arjan Cornelissen*: Supported by a Simons-CIQC postdoctoral fellowship through NSF QLCI Grant No. 2016245.

Simon Apers: Supported in part by the European QuantERA project QOPT (ERA-NET Cofund 2022-25), the French PEPR integrated projects EPiQ (ANR-22-PETQ-0007) and HQI (ANR-22-PNCQ-0002), and the French ANR project QUOPS (ANR-22-CE47-0003-01).

1 Introduction

The volume estimation problem is a widely-studied problem, that lies on the interface between convex geometry and computational complexity. The typical setting is the following, see for instance [33]. We consider a convex body $K \subseteq \mathbb{R}^d$ that satisfies $B_d \subseteq K \subseteq RB_d$, where $B_d = \{x \in \mathbb{R}^d : \|x\| \leq 1\}$ denotes the unit ball in d dimensions, and RB_d the radius- R ball in d dimensions. Denoting the volume of K by $\text{Vol}(K)$, the aim of volume estimation is to output $\tilde{V} \geq 0$ such that $|\tilde{V} - \text{Vol}(K)| < \varepsilon \text{Vol}(K)$, for a given relative precision $\varepsilon > 0$.

In this work, we are interested in the *query complexity* of volume estimation. Specifically, we will assume access to K through a *membership oracle* which, after querying a point $x \in \mathbb{R}^d$, returns whether $x \in K$. The key question is then how many membership queries (as a function of d , R and ε) are required to solve the volume estimation problem. We analyze this question in three *computational models*, namely the deterministic, randomized and quantum models, each of which is more powerful than the last.

Previous work mostly studied query complexity in the high-dimensional setting, which tracks the dependency of the query complexity on the dimension d . The title of this work is a reference to the early survey by Simonovits called “How to compute the volume in



© Arjan Cornelissen, Simon Apers, and Sander Gribling;
licensed under Creative Commons License CC-BY 4.0

52nd International Colloquium on Automata, Languages, and Programming (ICALP 2025).

Editors: Keren Censor-Hillel, Fabrizio Grandoni, Joël Ouaknine, and Gabriele Puppis

Article No. 61; pp. 61:1–61:18



Leibniz International Proceedings in Informatics

Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany



high dimension?” [33], and we refer the interested reader to this survey for a more in-depth discussion of this setting. In this work, however, we focus on the low-dimensional limit, or high-precision limit, i.e., the limit where the dimension d is fixed and the precision ε goes to zero. While we are not aware of any literature targeting this regime in the membership oracle setting, there are numerous algorithms in different access models targeting this regime. We summarize the most relevant of these below, and refer the interested reader to the excellent survey by Brunel [11].

1.1 Low-dimensional computational geometry

Before turning to volume estimation, we discuss a related problem that has been well studied in the low-dimensional limit. In the *convex set estimation* problem, one tries to output a convex set $\tilde{K} \subseteq \mathbb{R}^d$ that estimates the original convex set K .

One way to measure the error in this problem is through the *relative Nikodym metric*, defined as the volume of the symmetric difference $\tilde{K} \Delta K$ divided by the volume of K . An interesting sequence of results in this setting starts with a paper by Schütt [32]. They considered a bounded convex set $K \subseteq RB_d \subseteq \mathbb{R}^d$ with $R \in O(1)$, and showed that the convex hull of n points sampled uniformly at random from K , approximates it up to relative Nikodym distance $O(n^{-2/(d+1)})$. Later, Brunel proved a matching lower bound [10], implying that in order to achieve relative error ε , one requires $\Theta(\varepsilon^{-(d+1)/2})$ samples from the uniform distribution over K to solve the convex set estimation problem. Surprisingly, Baldin and Reiß [4] showed that the volume estimation problem is significantly easier in this setting, i.e., they show that $\Theta(\varepsilon^{-2(d+1)/(d+3)})$ uniform samples from K are necessary and sufficient to estimate its volume up to relative error ε .

Another line of research considers a qualitatively stronger error metric for the convex set estimation problem. To that end, suppose that a convex body $\tilde{K} \subseteq \mathbb{R}^d$ is an approximation of a convex body $K \subseteq \mathbb{R}^d$. For any $\varepsilon \in (0, 1)$, Agarwal, Har-Peled and Varadarajan [2] define \tilde{K} to be an ε -kernel of K if $\tilde{K} \subseteq K$, and

$$\forall u \in S_{d-1}, \quad \max_{x \in \tilde{K}} u^T x - \min_{x \in K} u^T x \geq (1 - \varepsilon) \left[\max_{x \in K} u^T x - \min_{x \in K} u^T x \right],$$

where $S_{d-1} = \partial B_d$ is the unit sphere in d dimensions. If \tilde{K} is an ε -kernel of K , then the Hausdorff distance between \tilde{K} and K , i.e., the maximum distance between a point in K from \tilde{K} and vice versa, is at most $\varepsilon \operatorname{diam}(K)$ (see Lemma 7). Being an ε -kernel is invariant under taking arbitrary linear transformations, just like the Nikodym metric, and any ε -kernel of K approximates it up to relative Nikodym distance $O(\varepsilon)$ too (see Proposition 8), yet the converse does not hold.

The starting point towards algorithmically constructing such ε -kernels is an early paper by Dudley [18, 23], which proves that every convex body $K \subseteq B_d$ has an approximate polytope $\tilde{K} \subseteq K$ with Hausdorff distance at most ε and only $O(\varepsilon^{-(d-1)/2})$ faces. Independently, Brohsteyn and Ivanov [9] proved that a similar approximate polytope exists with only $O(\varepsilon^{-(d-1)/2})$ vertices. Subsequently, in the more restricted setting where K is well-rounded, i.e., $B_d \subseteq K \subseteq RB_d$, Agarwal, Har-Peled and Varadarajan [2] used Brohsteyn and Ivanov’s construction to produce an ε -kernel with $O((R/\varepsilon)^{(d-1)/2})$ vertices, and their algorithm was subsequently improved independently by [13] and [34]. The referenced resources only explicitly analyze their algorithms’ performance in terms of the runtime in the case where K is the convex hull of n points, with [34] ultimately achieving a runtime of $O(n + \varepsilon^{-1/2})$ for $d = 2$ and $\tilde{O}(n + \varepsilon^{d-2})$ for $d \geq 3$. In this work, we build on these algorithmic ideas and port them to the membership oracle setting.

1.2 Our results

In this paper, we tightly characterize the membership oracle query complexity for the convex set estimation and volume estimation problems, in the low-dimensional limit, and in the deterministic, randomized and quantum models. The resulting complexities are displayed in Table 1.

■ **Table 1** Overview of the query complexity results obtained in this paper for the convex set estimation and volume estimation problems. We only track the dependence on $1/\varepsilon$ and R , the prefactors are allowed to depend exponentially on the dimension d . The tilde hides polylogarithmic factors in $1/\varepsilon$ and R .

Problem Error metric	Convex set estimation		Volume estimation
	Constructing ε -kernel	Rel. Nikodym distance ε	Rel. error ε
Deterministic	$\tilde{\Theta}(\varepsilon^{-\frac{d-1}{2}})$	$\tilde{\Theta}(\varepsilon^{-\frac{d-1}{2}})$	$\tilde{\Theta}(\varepsilon^{-\frac{d-1}{2}})$
Randomized	$\tilde{\Theta}(\varepsilon^{-\frac{d-1}{2}})$	$\tilde{\Theta}(\varepsilon^{-\frac{d-1}{2}})$	$\tilde{\Theta}(\varepsilon^{-\frac{2(d-1)}{d+3}})$
Quantum	$\tilde{\Theta}(\varepsilon^{-\frac{d-1}{2}})$	$\tilde{\Theta}(\varepsilon^{-\frac{d-1}{2}})$	$\tilde{\Theta}(\varepsilon^{-\frac{d-1}{d+1}})$

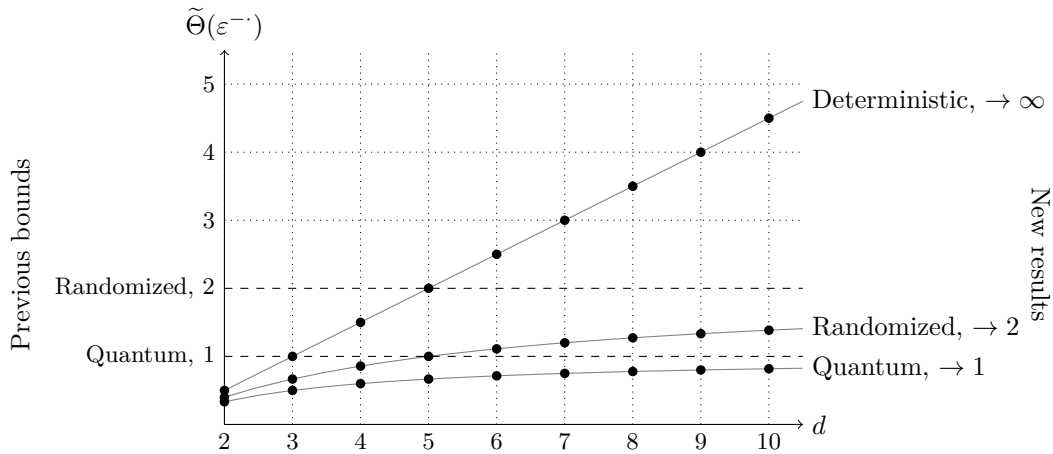
For the convex set estimation problem, we consider both the problem of constructing an ε -kernel, as well as estimating K up to relative Nikodym distance ε . We obtain that the query complexities of both problems are $\tilde{\Theta}(\varepsilon^{-(d-1)/2})$ in all three computational models. This shows in particular that randomness and quantumness alike do not provide any benefit over the deterministic model. Furthermore, for estimating a convex set up to relative Nikodym error, having access to a membership oracle is strictly more powerful than uniform sampling from the convex body, as we beat the $\Theta(\varepsilon^{-(d+1)/2})$ samples that are required in that setting [10].

For the volume estimation problem, we plot the obtained complexities in Figure 1. For any fixed dimension d , we beat the previously best-known algorithms in the randomized and quantum settings, that respectively make $O(1/\varepsilon^2)$ and $O(1/\varepsilon)$ queries. The exponents in our complexities converge to these naive bounds as d increases, showing that the obtained advantage becomes ever smaller with increasing d . Moreover, the gap between the randomized and quantum query complexity is small when d is small, and becomes bigger as d increases, converging to a full quadratic separation in the regime where d is large. Like in the uniform sampling setting, the volume estimation problem is significantly easier than the convex set estimation problem in the randomized model.

1.3 Techniques

Our algorithms are based on the approaches taken in [2, 13, 34], and they all proceed in a similar manner. First, we apply an existing, deterministic rounding procedure [22, Theorem 4.6.1] that makes $\tilde{O}(1)$ queries and transforms the convex body into a “well-rounded” one, i.e., satisfying $B_d \subseteq K \subseteq RB_d$ with $R \in O(1)$ when d is fixed.

Then, we follow [34], and take a set $\{v_j\}_{j=1}^n$ of n roughly equally spaced points on the boundary of $(R+1)B_d$, with $n \in \Theta(\varepsilon^{-(d-1)/2})$. Every point v_j is subsequently projected onto the convex body, and the convex hull of the resulting points forms an ε -kernel of K , as was shown in [2]. Since these projections cannot be implemented directly in our model, we formulate every projection operation as a convex optimization problem, and use existing solvers from [22] to obtain an approximate projection. It was already shown in [13] and [34, Theorem 1] that this suffices to obtain an ε -kernel \tilde{K} of K . It also follows that \tilde{K} is an ε -precise approximation of K in the relative Nikodym metric.



■ **Figure 1** Graph of the exponents in the query complexities for the volume estimation problem. d is fixed, and the asymptotic limit is for $R \rightarrow \infty$ and $\epsilon \downarrow 0$. The tilde hides polylogarithmic factors in $1/\epsilon$ and R . The dashed lines represent the previously best-known results, and the solid ones connect the newly-found complexities.

Next, we find a speed-up of this approach for the volume estimation problem. To that end, we first run the convex set estimation procedure to some precision $\epsilon' > \epsilon$, to obtain an ϵ' -kernel \underline{K} of K . We slightly enlarge the inner approximation $\underline{K} \subseteq K$ to generate an outer approximation $\overline{K} \supseteq K$ of K , with the property that $\text{Vol}(\overline{K} \setminus \underline{K}) / \text{Vol}(K) \in O(\epsilon')$. Then, in the randomized setting, we sample $\Theta((\epsilon'/\epsilon)^2)$ times uniformly from the region $\overline{K} \setminus \underline{K}$, and compute the fraction of samples ξ that is inside K . We then use this as a refinement to our estimate of $\text{Vol}(K)$, i.e., we output $\text{Vol}(\underline{K}) + \xi \text{Vol}(\overline{K} \setminus \underline{K})$. In the quantum case, we employ the same idea, but this time we use quantum amplitude estimation to quadratically speed up the computation of ξ . In both cases, balancing the number of queries made in the first and second step yields the complexity bounds from Table 1.

The matching lower bounds for the query complexities follow from a construction that dates back to Dudley [18], where it was used to prove a lower bound on the number of facets needed for approximating convex bodies by polytopes. The construction packs n disjoint, identical spherical caps on the boundary of the unit ball through a δ -net, for some suitable value of δ that depends on d and ϵ . Then, a convex body $K_x \subseteq B_d$ is associated to a bitstring x of length n by including the j th spherical cap in K_x if and only if $x_j = 1$. See Figure 2 for an illustration.

The core observation is that estimating K_x , or its volume, becomes equivalent to learning a fraction of the entries of the bitstring x , or its Hamming weight, respectively. Additionally, a membership query to K_x can be simulated with a single query to x . Recovering a constant fraction of bits of an n -bit string is known to cost $\Theta(n)$ queries in all three models that we consider. Similarly, the problem of estimating the Hamming weight of a bitstring is also known as the approximate counting problem and its (deterministic, randomized and quantum) query complexity has been tightly characterized in earlier work. We show that by carefully choosing the size of the spherical caps, this construction provides matching lower bounds for the query complexities from Table 1 up to polylogarithmic factors in R and ϵ (and large factors in the dimension d).

1.4 History in the high-dimensional setting

For completeness, we also comment on volume estimation in the high-dimensional setting, which tracks the dependency of the query complexity on the dimension d . The title of this work is a reference to the early survey by Simonovits called “How to compute the volume in high dimension?” [33], and we refer the interested reader to this survey for a more in-depth discussion of this setting. The main takeaways in this setting are the following:

- In the *deterministic* setting, any deterministic algorithm must make a number of queries that is exponential in d [20, 5].
- In the *randomized* setting, there famously exists a Markov chain Monte Carlo algorithm that makes a number of queries polynomially in d and $1/\varepsilon$ [19]. While the initial bound scaled as $\tilde{O}(d^{23}/\varepsilon^2)$, a long of research led to the current best bound that is $\tilde{O}(d^3/\varepsilon^2)$. This bound follows from combining algorithms in [17, 25] with recent breakthroughs on the so-called KLS-constant (see e.g. [27]), and it approaches a lower bound of $\tilde{\Omega}(d^2)$ on the randomized query complexity [31].
- In the *quantum* setting, Chakrabarti, Childs, Hung, Li, Wang and Wu [12] gave an improved bound of the form $\tilde{O}(d^3 + d^{2.5}/\varepsilon)$.¹ This was later improved to $\tilde{O}(d^3 + d^{2.25}/\varepsilon)$ by Cornelissen and Hamoudi [15]. Both use quantum algorithms to speed up the Markov chain Monte Carlo method. The authors of [12] also prove that the query complexity is $\Omega(\sqrt{d})$ when ε is constant, and $\Omega(1/\varepsilon)$ in the regime where $1/d \leq \varepsilon \leq 1/3$, which in turn implies $\Omega(d)$ when $\varepsilon = 1/d$.²

1.5 Organization

In Section 2, we fix notation, formally define algorithmic models and computational problems, and state results from (computational) geometry. Subsequently, in Section 3 we develop the algorithms, and in Section 4 we prove the corresponding lower bounds.

2 Preliminaries

2.1 Notation

We start by fixing some notation. Let \mathbb{N} be the set of all positive integers. For all $n \in \mathbb{N}$, let $[n] = \{1, \dots, n\}$, and $[n]_0 = \{0\} \cup [n]$. For any $k \in \mathbb{N}$ (or $x \in \mathbb{R}$), we let $\mathbb{N}_{\geq k}$ (resp. $\mathbb{R}_{\geq x}$) be the set of all integers (resp. reals) that are bigger than or equal to k (resp. x). We denote the Euclidean norm by $\|\cdot\|$. For every $d \in \mathbb{N}$, we denote the unit ball in d dimensions by $B_d = \{x \in \mathbb{R}^d : \|x\| \leq 1\}$. We write $\Gamma_d = \text{Vol}(B_d)$ for the volume of B_d . We denote the unit sphere in d dimensions by $S_{d-1} = \partial B_d$. For two sets A and B we let $A \Delta B = (A \setminus B) \cup (B \setminus A)$ be the symmetric difference of A and B .

¹ In [12, Page 20:11], the authors claim that recent breakthroughs in the KLS-conjecture [14, 24, 26] improve the analysis of their volume estimation algorithm to $\tilde{O}(d^{2.5+o(1)} + d^{2+o(1)}/\varepsilon)$ membership oracle queries. They argue that the mixing time of the hit-and-run walk is $\tilde{O}(d^2 \psi^2)$, where ψ is the KLS-constant. This indeed appears to be a correct upper bound on the mixing time of the *ball walk*, see [25, Theorem 2.7], but, as far as we are aware, the mixing time of the hit-and-run walk is not known to be improved due to the resolution of the KLS-conjecture. Adapting the algorithm of [12] to make use of the ball walk is highly non-trivial. Consequently, the question of whether volume estimation can be done in $\tilde{O}(d^{2.5+o(1)} + d^{2+o(1)}/\varepsilon)$ quantum queries to the membership oracle is still open.

² In [12], the authors additionally informally suggest a lower bound of $\Omega(1/\varepsilon)$ in general, but the corresponding theorem statement excludes the limit $\varepsilon \downarrow 0$. A lower bound of $\Omega(1/\varepsilon)$ for fixed d and in the limit where $\varepsilon \downarrow 0$ would indeed contradict our results.

A set $K \subseteq \mathbb{R}^d$ is convex if for any two points $x, y \in K$, the straight line segment between x and y is fully contained in K , i.e., for every $\lambda \in [0, 1]$, $\lambda x + (1 - \lambda)y \in K$. For two convex sets $K, K' \subseteq \mathbb{R}^d$, we recall that $K \cap K'$ and $K + K' = \{x + y : x \in K, y \in K'\}$ are again convex. For ease of notation, we typically make an implicit assumption that all convex bodies are closed.

We use big- O -notation to hide constant factors. That is, for two functions $f, g : \mathbb{R}_{\geq 0}^d \rightarrow \mathbb{R}_{\geq 0}$, we write $f \in O(g)$ if there exist $C, r > 0$ such that $f(x) \leq Cg(x)$ whenever $\|x\| \geq r$. We write $f \in \Omega(g)$ if $g \in O(f)$, and we write $f \in \Theta(g)$ if $f \in O(g) \cup \Omega(g)$. Sometimes, the limit in which the big- O -notation is to be interpreted plays an important role. For instance if the big- O -notation holds in the limit where $\varepsilon \downarrow 0$, then we interpret the function f, g as functions in the variable $1/\varepsilon$.

Similarly, we use the big- O -tilde-notation to hide polylogarithmic factors. That is, if $f, g : \mathbb{R}_{\geq 0}^d \rightarrow \mathbb{R}_{\geq 0}$, then we write $f \in \tilde{O}(g)$ if there exist constants $C, k, r \geq 0$ such that $f(x) \leq Cg(x) \prod_{j=1}^d \log^k(x_j)$ whenever $\|x\| \geq r$. We write $f \in \tilde{\Omega}(g)$ if $g \in \tilde{O}(f)$, and $f \in \tilde{\Theta}(g)$ if $f \in \tilde{O}(g) \cup \tilde{\Omega}(g)$.

2.2 Access model and computational models

The input to the algorithms we design in this work are a dimension $d \in \mathbb{N}$, outer radius $R \geq 1$, precision $\varepsilon \in (0, 1)$, and a convex body $K \subseteq \mathbb{R}^d$. The *access model* specifies how the algorithm can access this input. Throughout, we will assume that the parameters d, R and ε are set beforehand, and hence are known to the algorithm. As for the convex body K , many ways to access it have been considered in the literature. We refer to [22, Chapter 2] for an overview of several different access models. In this work, we assume that it can be accessed by means of a (strong) *membership oracle*, in line with [22, Definition 2.1.5].

► **Definition 1** (Membership oracle). *Let $d \in \mathbb{N}$, and $K \subseteq \mathbb{R}^d$ be a convex body. A membership oracle to K is a procedure that, given a point $x \in \mathbb{R}^d$, outputs whether $x \in K$.*

In general, an *oracle* is a subroutine that the algorithm can run only as a black box, i.e., it has no more information about what happens when it is run. One such call to the oracle is referred to as a query. The *query complexity* of a problem is the minimum number of queries any algorithm needs to make to solve said problem.

Additional to the access model, we also specify *computational models* that define the operations that the algorithms are allowed to use. We consider three different computational models in this work. In the *deterministic model*, the algorithm is completely described beforehand and follows a deterministic computational path. In the *randomized model*, the algorithm is allowed to use randomness during the execution of the algorithm, and in particular let the oracle's inputs depend on it. It is required to output a correct answer with probability at least $2/3$. In the *quantum model*, the algorithm is additionally allowed to perform operations on a quantum state, and supply the oracle's inputs in a superposition, which it answers coherently. For a more detailed introduction to quantum algorithms, we refer to [30].

Throughout the execution of our algorithms, we additionally assume that we can store vectors in \mathbb{R}^d precisely in memory in all three models, and that we can perform basic arithmetic with them exactly. We also assume in the randomized model that we can sample from any distribution on \mathbb{R}^d as long as we have a classical description of it, and we use this to sample uniformly from an arbitrary region in Theorem 13. Similarly, in the quantum model, we assume that we can generate a uniform superposition over an arbitrary region

in \mathbb{R}^d , and we use this in Theorem 14. Both of these operations do not depend on K , so in an implementation of the algorithms presented in this work, one can suppress the finite-precision errors that arise from these assumptions arbitrarily without increasing the number of membership oracle queries.

2.3 Rounding convex bodies

We start by describing a routine that “rounds” the convex body. A similar idea appears in [2, 13, 34], but in contrast to their approach, we use inner and outer ellipses, rather than inner and outer cubes in our rounding procedure. This difference is not fundamental, it merely allows us to use a rounding routine that already exists in the literature [22].

The aim of “rounding” a convex body $K \subseteq \mathbb{R}^d$ is to find an invertible affine linear map $L : x \mapsto x_0 + Tx$, such that $B_d \subseteq L(K) \subseteq R'B_d$ with $R' \geq 1$ as small as possible. Intuitively, one can think of rounding as finding a linear transformation that compresses the convex body as much as possible, so that it does not have parts that stick out far from the origin. Most of the existing literature focuses on randomized algorithms to compute this affine linear map [28, 25], with the current state-of-the-art obtaining $R' \in \tilde{O}(\sqrt{d})$ and requiring $\tilde{O}(d^3)$ queries. In our setting, though, we require a deterministic procedure, and since we take d to be a constant we don’t necessarily need R' to scale optimally in the dimension. Below, we sketch how to obtain a deterministic rounding procedure that obtains $R' \in O(d^3)$ and runs in time $O(\text{poly}(d, \log(R)))$.

► **Theorem 2** ([22]). *Let $d \in \mathbb{N}$, $R \geq 1$, and let $K \subseteq \mathbb{R}^d$ be convex such that $B_d \subseteq K \subseteq RB_d$. Then, there is a deterministic algorithm that makes $O(\text{poly}(d, \log(R)))$ queries, and finds an invertible affine linear map L such that $B_d \subseteq L(K) \subseteq R'B_d$, with $R' = d(d+1)^2 \in O(d^3)$. When d is fixed, the algorithm makes $\tilde{O}(1)$ queries, and $R' \in O(1)$.*

Proof. Since $K \subseteq \mathbb{R}^d$ satisfies $B_d \subseteq K \subseteq RB_d$, we observe from [22, Figure 4.1] that we can deterministically turn a membership oracle into a weak separation oracle. The total multiplicative factor incurred from this conversion is $O(\text{poly}(d, \log(R)))$, and hence is $\tilde{O}(1)$ in the case where d is fixed.

Subsequently, from [22, Theorem 4.6.1], we find that with $O(\text{poly}(d, \log(R)))$ calls to a weak separation oracle to K , we can find an ellipsoid E , such that $E \subseteq K \subseteq d(d+1)^2E$. Thus, by letting L be the affine linear transformation that maps E to the unit ball, we obtain that

$$B_d \subseteq L(K) \subseteq R'B_d, \quad \text{where} \quad R' = d(d+1)^2 \in O(d^3).$$

For fixed d , we obtain $R' \in O(1)$, and the query complexity of finding L is in $\tilde{O}(1)$. ◀

2.4 Geometry

We proceed by recalling some theoretical background in geometry. We start by formally introducing the concept of a net.

► **Definition 3** (δ -net). *Let $d \in \mathbb{N}$, $\delta > 0$, and let $S \subseteq \mathbb{R}^d$ be any set. We say that $N \subseteq S$ is a δ -net of S if*

1. *For any two distinct points $v, w \in N$, we have $\|v - w\| \geq \delta$.*
2. *For any $v \in S$, there exists a $w \in N$ such that $\|v - w\| \leq \delta$.*

Intuitively, one can think of a δ -net N as the centers of a set of balls of radius δ , that together cover the set S . Moreover, since these centers are at least separated by distance δ , the balls with radius $\delta/2$ centered at the points in N must be disjoint up to a measure-zero

set. By comparing the surface area of the sphere to that of $\{u \in S_{d-1} : \|u - v\| \leq \delta\}$ (for some fixed $v \in S_{d-1}$), the following bound follows on the number of points in a δ -net on the sphere. We stress that the assumption that d is fixed is crucial in the above proposition. Indeed, the Θ -notation hides a prefactor that might depend exponentially on d .

► **Proposition 4.** *Let $d \in \mathbb{N}$ be fixed. Let $\delta \in (0, 1)$, and let N_δ be a δ -net of S_{d-1} . Then, $|N_\delta| \in \Theta(\delta^{-(d-1)})$.*

Proof. For any $v \in V$ and $r > 0$, let $R_{v,r} = \{w \in S_{d-1} : \|v - w\| \leq r\}$. The surface areas of $R_{v,r}$ is $A_{v,r} \in \Theta(r^{d-1})$ and the surface area of S_{d-1} is $A_{d-1} \in \Theta(1)$. Since N_δ is a δ -net, all $R_{v,\delta/2}$'s with $v \in N_\delta$ are disjoint up to possibly some measure-zero set, and all $R_{v,\delta}$'s cover S_{d-1} . Thus, we find $|N_\delta|A_{v,\delta/2} \leq A_{d-1} \leq |N_\delta|A_{v,\delta}$, from which we conclude that $|N_\delta| \in \Theta(\delta^{-(d-1)})$. ◀

We proceed with formally defining an ε -kernel.

► **Definition 5** (ε -kernel [2]). *Let $d \in \mathbb{N}$, and $\varepsilon \in (0, 1)$. Let $K \subseteq \mathbb{R}^d$ be a convex body. Then a convex body $\tilde{K} \subseteq K$ is an ε -kernel of K if*

$$\forall u \in S_{d-1}, \quad \max_{x \in \tilde{K}} u^T x - \min_{x \in K} u^T x \geq (1 - \varepsilon) \left[\max_{x \in K} u^T x - \min_{x \in K} u^T x \right].$$

One very desirable property of ε -kernels is its invariance under affine linear transformations.

► **Lemma 6** ([2],[1, Lemma 2.5]). *Let $d \in \mathbb{N}$, $\varepsilon \in (0, 1)$, and let L be an invertible affine linear map on \mathbb{R}^d . Let $\tilde{K}, K \subseteq \mathbb{R}^d$ be convex sets. Then, \tilde{K} is an ε -kernel of K if and only if $L(\tilde{K})$ is an ε -kernel of $L(K)$.*

Next, we make a connection between ε -kernels and approximations w.r.t. the Hausdorff metric.

► **Lemma 7.** *Let $d \in \mathbb{N}$, $R \geq 1$, $\varepsilon > 0$, and let $K, \tilde{K} \subseteq \mathbb{R}^d$ be convex sets.*

1. *If $B_d \subseteq K$ and $\tilde{K} \subseteq K$ is an ε -precise Hausdorff approximation of K , then \tilde{K} is also an ε -kernel of K .*
2. *If $K \subseteq RB_d$ and \tilde{K} is an ε -kernel of K , it is a $2\varepsilon R$ -precise Hausdorff approximation of K .*

Proof. Observe that if $\tilde{K} \subseteq K$, then the Hausdorff distance d between \tilde{K} and K is

$$d = \max_{u \in S_{d-1}} \left[\max_{x \in K} u^T x - \max_{x \in \tilde{K}} u^T x \right] = \max_{u \in S_{d-1}} \left[\min_{x \in \tilde{K}} u^T x - \min_{x \in K} u^T x \right]. \quad (1)$$

For the first claim, we take $u \in S_{d-1}$ arbitrarily, and observe from Equation (1) that

$$\max_{x \in \tilde{K}} u^T x - \min_{x \in \tilde{K}} u^T x \geq \max_{x \in K} u^T x - \min_{x \in K} u^T x - 2\varepsilon \geq (1 - \varepsilon) \left[\max_{x \in K} u^T x - \min_{x \in K} u^T x \right].$$

For the second claim, let d be the Hausdorff distance between K and \tilde{K} , and u the vector that maximizes the middle expression in Equation (1). Then, we have

$$\begin{aligned} d &= \max_{x \in K} u^T x - \max_{x \in \tilde{K}} u^T x \leq \left[\max_{x \in K} u^T x - \min_{x \in K} u^T x \right] - \left[\max_{x \in \tilde{K}} u^T x - \min_{x \in \tilde{K}} u^T x \right] \\ &\leq \varepsilon \left[\max_{x \in K} u^T x - \min_{x \in K} u^T x \right] \leq \varepsilon \cdot \text{diam}(K) \leq 2\varepsilon R. \end{aligned} \quad \blacktriangleleft$$

Finally, we observe that an ε -kernel naturally provides a relative Nikodym approximation of K , and can be used to construct an outer approximation of K , in the following proposition.

► **Proposition 8.** *Let $d \in \mathbb{N}$, $\varepsilon \in (0, 1/(4d(d+1)^2))$, and let \tilde{K} be an ε -kernel of a convex set $K \subseteq \mathbb{R}^d$. Then, given a full description of \tilde{K} , we can construct a convex set $\bar{K} \supseteq K$ such that $\text{Vol}(\bar{K} \setminus \tilde{K}) / \text{Vol}(K) \leq (1 + 4\varepsilon d(d+1)^2)^d - 1 \in O(\varepsilon)$, with $\tilde{O}(1)$ membership oracles to K . In particular, this implies that $\text{Vol}(\tilde{K} \Delta K) / \text{Vol}(K) \in O(\varepsilon)$.*

Proof. Since being an ε -kernel and being a relative Nikodym-distance approximation are both invariant under affine linear transformations, we can without loss of generality first round the convex body K using Theorem 2, using just $\tilde{O}(1)$ membership queries. It then suffices to consider the case where $B_d \subseteq K \subseteq RB_d$, with $R = d(d+1)^2 \in O(1)$. Since $\varepsilon < 1/(4R)$, we find by Lemma 7 that $B_d \subseteq K \subseteq \tilde{K} + 2\varepsilon RB_d \subseteq \tilde{K} + B_d/2$, and so $B_d/2 \subseteq \tilde{K}$. Thus,

$$B_d/2 \subseteq \tilde{K} \subseteq K \subseteq \tilde{K} + 2\varepsilon RB_d \subseteq (1 + 4\varepsilon R)\tilde{K} =: \bar{K},$$

from which we find that

$$\frac{\text{Vol}(\bar{K} \setminus \tilde{K})}{\text{Vol}(K)} = \frac{\text{Vol}(\bar{K}) - \text{Vol}(\tilde{K})}{\text{Vol}(K)} = \frac{[(1 + 4\varepsilon R)^d - 1] \text{Vol}(\tilde{K})}{\text{Vol}(K)} \leq (1 + 4\varepsilon R)^d - 1 \in O(\varepsilon). \blacktriangleleft$$

2.5 Problem definitions

We consider three computational problems in this paper. We formally introduce them here.

► **Definition 9** (Problem definitions). *Let $d \in \mathbb{N}$, $R > 1$, and $\varepsilon \in (0, 1)$. Let $K \subseteq \mathbb{R}^d$ be a convex body such that $B_d \subseteq K \subseteq RB_d$. We consider three problems:*

1. *The ε -kernel construction problem is the problem of outputting an ε -kernel of K .*
2. *The ε -Nikodym construction problem is the problem of outputting a convex body $\tilde{K} \subseteq \mathbb{R}^d$ such that $\text{Vol}(K \Delta \tilde{K}) \leq \varepsilon \text{Vol}(K)$.*
3. *The volume estimation problem is the problem of outputting a non-negative real $\tilde{V} \geq 0$ such that $|\text{Vol}(K) - \tilde{V}| \leq \varepsilon \text{Vol}(K)$.*

It follows directly that these problems are qualitatively decreasing in terms of their query complexity. That is, solving the ε -kernel construction problem also solves the Nikodym construction problem with precision $O(\varepsilon)$, by virtue of Proposition 8, which in turn solves the volume estimation problem with the same precision, since we can simply output the volume of the approximation \tilde{K} . These relations are less clear if one considers the runtime instead, e.g., since computing the volume might be very costly.

3 Algorithms

3.1 Convex set estimation

In this subsection, we develop a deterministic algorithm that constructs an ε -kernel of a well-rounded convex set $K \subseteq \mathbb{R}^d$, using membership queries to it. The algorithm follows the same general strategy as in [34], and we replace their approximate nearest-neighbor queries by a query-efficient approximate projection onto a convex body. For this, we use the well-known observation that projection onto a convex body can be phrased as a convex optimization problem, which we can deterministically solve approximately with only polylogarithmically many membership oracles calls, using for example the ellipsoid method [22].

61:10 How to Compute the Volume in Low Dimension?

► **Proposition 10.** *Let $d \in \mathbb{N}$, $R \geq 1$, $\varepsilon > 0$, $K \subseteq \mathbb{R}^d$ convex such that $B_d \subseteq K \subseteq RB_d$. Let $x \in \mathbb{R}^d \setminus K$, and y' its projection onto K . There is a deterministic algorithm that obtains an element $\tilde{y} \in K$ such that $\|\tilde{y} - y'\| \leq \varepsilon$. The algorithm makes $O(\text{poly}(d, \log \|x\|, \log R, \log(1/\varepsilon)))$ queries to a membership oracle of K . When d is fixed and $R \in O(1)$, this complexity is $\tilde{O}(1)$.*

Now, we are ready to state the ε -kernel construction algorithm. We start by considering the well-rounded case, in Algorithm 1. We subsequently prove its properties in Theorem 11.

■ **Algorithm 1** Well-rounded deterministic ε -kernel construction [34].

Input:

1. $d \in \mathbb{N}_{\geq 2}$: the dimension.
2. $\varepsilon \in (0, 1)$: the desired precision.
3. $R \geq 1$: the outer radius, with $R \in O(1)$.
4. O_K : a membership oracle that on input $x \in \mathbb{R}^d$ returns whether $x \in K$, where $B_d \subseteq K \subseteq RB_d$.

Derived constant: $\eta = \sqrt{\varepsilon/R}$.

Output: An ε -kernel \tilde{K} of K .

Number of queries: $\tilde{O}(\varepsilon^{-(d-1)/2})$ (in the limit where d is fixed).

Procedure: `RoundedEpsKernConstr`(d, ε, R, O_K):

1. Let $\{x_j\}_{j=1}^n$ be an η -net of $(R+1)S_{d-1}$.
 2. For $j = 1, \dots, n$,
 - a. Project x_j onto K with precision ε , using Proposition 10. Denote the outcome by p_j .
 3. Output $\text{conv}(\{p_j\}_{j=1}^n)$.
-

► **Theorem 11.** *Let $d \in \mathbb{N}_{\geq 2}$, $R \geq 1$ with $R \in O(1)$, and $K \subseteq \mathbb{R}^d$ be a convex body such that $B_d \subseteq K \subseteq RB_d$. Then, Algorithm 1 computes an ε -kernel of K that satisfies $\tilde{K} \subseteq K$, with $\tilde{O}(\varepsilon^{-(d-1)/2})$ membership oracle queries.*

Proof. Correctness is proven in [34, Theorem 1]. For the bound on the number of queries, observe from Proposition 4 that the η -net contains $O(\eta^{-(d-1)}) = O(\varepsilon^{-(d-1)/2})$ points. Furthermore, for every point, we perform one approximate projection, which costs $\tilde{O}(1)$ queries by Proposition 10. As such, we conclude that the total number of membership oracle queries is $\tilde{O}(\varepsilon^{-(d-1)/2})$. ◀

In the case where K is not well-rounded, we combine Algorithm 1 with the rounding procedure of Theorem 2, and conversion to Nikodym distance in Proposition 8, to obtain the following corollary.

► **Corollary 12** (Deterministic ε -kernel construction). *Let $d \in \mathbb{N}_{\geq 2}$, $R \geq 1$, $\varepsilon \in (0, 1/(4d(d+1)^2))$, and $K \subseteq \mathbb{R}^d$ be convex such that $B_d \subseteq K \subseteq RB_d$. Then, we can deterministically compute an ε -kernel \tilde{K} of K with $\tilde{O}(\varepsilon^{-(d-1)/2})$ membership oracle queries. \tilde{K} is also an $O(\varepsilon)$ -precise approximation of K in the relative Nikodym distance, and $\text{Vol}(\tilde{K})$ is an $O(\varepsilon)$ -precise relative estimate of $\text{Vol}(K)$.*

This corollary proves all the $\tilde{O}(\varepsilon^{-(d-1)/2})$ upper bounds in Table 1. In Section 4.2, we prove that this approach is essentially optimal for the reconstruction problem in all computational models, i.e., even in the randomized and quantum settings we cannot improve significantly over the approach taken in Corollary 12.

3.2 Volume estimation

In this subsection, we switch to the volume estimation problem. To start off, we remark that Corollary 12 solves it in the deterministic setting. In the randomized and quantum settings, however, we can obtain a significant improvement over this approach.

The core idea is to run the (deterministic) ε' -kernel construction algorithm up to some worse precision $\varepsilon' > \varepsilon$, to obtain an inner approximation $\underline{K} \subseteq K$. We can use Proposition 8 to obtain an outer approximation $\overline{K} \supseteq K$, such that $\text{Vol}(\overline{K} \setminus \underline{K}) / \text{Vol}(K) \in O(\varepsilon')$.

Subsequently, in the randomized setting, we sample uniformly from $\overline{K} \setminus \underline{K}$, and compute the fraction of points ξ that is inside K . We then compute $\text{Vol}(\underline{K}) + \xi \text{Vol}(\overline{K} \setminus \underline{K})$ as a refined estimate of the volume of K , eventually resulting in the following theorem.

► **Theorem 13** (Randomized volume estimation). *Let $d \in \mathbb{N}_{\geq 2}$, $R \geq 1$, $\varepsilon \in (0, 1)$, $K \subseteq \mathbb{R}^d$ convex such that $B_d \subseteq K \subseteq RB_d$. We can compute $\tilde{V} \geq 0$ such that $|\tilde{V} - \text{Vol}(K)| \leq \varepsilon \text{Vol}(K)$ with probability at least $2/3$, using a randomized algorithm that makes $\tilde{O}(\varepsilon^{-2(d-1)/(d+3)})$ queries to a membership oracle of K .*

Proof. We first use Corollary 12 to obtain an ε' -kernel $\underline{K} \subseteq K$ with

$$\varepsilon' = \frac{(1 + \varepsilon^{\frac{4}{d+3}})^{\frac{1}{d}} - 1}{4d(d+1)^2} \in \Theta\left(\varepsilon^{\frac{4}{d+3}}\right).$$

We make $\tilde{O}((\varepsilon')^{-(d-1)/2}) = \tilde{O}(\varepsilon^{-2(d-1)/(d+3)})$ membership oracles in this step of the algorithm. Then, we use Proposition 8 to generate an outer approximation $\overline{K} \supseteq K$, using just $\tilde{O}(1)$ membership oracle queries. We find that $\text{Vol}(\overline{K} \setminus \underline{K}) / \text{Vol}(K) \leq \varepsilon'$.

Next, we take $n := \lceil 3(\varepsilon'/\varepsilon)^2 \rceil$ random samples from $\overline{K} \setminus \underline{K}$, and we check for each of them whether they are in K . We denote the fraction of them that is in K by ξ . This step requires $n = \tilde{O}((\varepsilon'/\varepsilon)^2) = \tilde{O}(\varepsilon^{-2(d-1)/(d+3)})$ membership oracle queries too.

Finally, we output $\tilde{V} = \text{Vol}(\underline{K}) + \xi \text{Vol}(\overline{K} \setminus \underline{K})$. We observe that

$$\mathbb{E}[\tilde{V}] = \text{Vol}(\underline{K}) + \mathbb{E}[\xi] \text{Vol}(\overline{K} \setminus \underline{K}) = \text{Vol}(\underline{K}) + \frac{\text{Vol}(K \setminus \underline{K})}{\text{Vol}(\overline{K} \setminus \underline{K})} \text{Vol}(\overline{K} \setminus \underline{K}) = \text{Vol}(K),$$

and

$$\text{Var}[\tilde{V}] = \text{Vol}(\overline{K} \setminus \underline{K})^2 \text{Var}[\xi] \leq \frac{(\varepsilon' \cdot \text{Vol}(K))^2}{n} \leq \frac{(\varepsilon \cdot \text{Vol}(K))^2}{3}.$$

We conclude with Chebyshev's inequality that

$$\mathbb{P}\left[\left|\tilde{V} - \text{Vol}(K)\right| > \varepsilon \text{Vol}(K)\right] \leq \frac{\text{Var}[\tilde{V}]}{(\varepsilon \cdot \text{Vol}(K))^2} \leq \frac{1}{3}. \quad \blacktriangleleft$$

In the quantum case, we speed up the sampling phase from the randomized algorithm by replacing it with a quantum primitive known as amplitude estimation [8]. We attain a quadratic speed-up of this step, and the claimed complexity follows after rebalancing the costs of the different steps.

► **Theorem 14** (Quantum volume estimation). *Let $d \in \mathbb{N}_{\geq 2}$, $R \geq 1$, $\varepsilon \in (0, 1)$, and $K \subseteq \mathbb{R}^d$ convex such that $B_d \subseteq K \subseteq RB_d$. We can compute $\tilde{V} \geq 0$ such that $|\tilde{V} - \text{Vol}(K)| \leq \varepsilon \text{Vol}(K)$ with probability at least $2/3$, using a quantum algorithm that makes $\tilde{O}(\varepsilon^{-(d-1)/(d+1)})$ queries to O_K .*

61:12 How to Compute the Volume in Low Dimension?

Proof. Similarly to the proof of Theorem 13, we start by constructing an ε' -kernel $\underline{K} \subseteq K$ using Corollary 12, and then use Proposition 8 to obtain $\overline{K} \supseteq K$ such that $\text{Vol}(\overline{K} \setminus \underline{K}) \leq \varepsilon' \text{Vol}(K)$. In contrast to Theorem 13, though, we choose

$$\varepsilon' = \frac{(1 + \varepsilon^{\frac{2}{d+1}})^{\frac{1}{d}} - 1}{4d(d+1)^2} \in \Theta\left(\varepsilon^{\frac{2}{d+1}}\right),$$

which brings the total number of membership oracle queries in this part to $\tilde{O}((\varepsilon')^{-(d-1)/2}) = \tilde{O}(\varepsilon^{-(d-1)/(d+1)})$.

Next, we use amplitude estimation to find an estimate ξ of $\text{Vol}(K \setminus \underline{K}) / \text{Vol}(\overline{K} \setminus \underline{K})$ up to precision ε/ε' . To that end, we compute the overlap between the uniform superposition on $\overline{K} \setminus \underline{K}$, and the subspace spanned all the quantum states representing vectors in K . From [8], we find that this can be done with success probability at least $8/\pi^2 > 2/3$, and with a total number of queries that satisfies $\tilde{O}(\varepsilon'/\varepsilon) = \tilde{O}(\varepsilon^{-(d-1)/(d+1)})$. Finally, we output $\tilde{V} = \text{Vol}(K) + \xi \text{Vol}(\overline{K} \setminus \underline{K})$, and observe that

$$|\tilde{V} - \text{Vol}(K)| = \text{Vol}(\overline{K} \setminus \underline{K}) \cdot \left| \xi - \frac{\text{Vol}(K \setminus \underline{K})}{\text{Vol}(\overline{K} \setminus \underline{K})} \right| \leq \varepsilon' \cdot \text{Vol}(K) \cdot \frac{\varepsilon}{\varepsilon'} = \varepsilon \cdot \text{Vol}(K). \quad \blacktriangleleft$$

This concludes our discussion of randomized and quantum algorithms estimating the volume of a convex body. In Section 4.2, we prove that these algorithms are essentially optimal.

4 Lower bounds

In this section, we prove matching lower bounds for the computational problems from Definition 9. We can always rescale our problems, and so we can freely change the assumption that $B_d \subseteq K \subseteq RB_d$ to $(1/R)B_d \subseteq K \subseteq B_d$. For ease of notation, we will phrase everything with the latter assumption in mind.

The lower bounds crucially rely on embedding a bitstring on the boundary of the unit ball in d dimensions. To that end, we let $\{v_j\}_{j=1}^n$ be a net of S_{d-1} , and for every $j \in [n]$, we define a *spherical cap* P_j around v_j , i.e., a small region cut off from the unit ball by a hyperplane that is orthogonal to v_j . We take all the spherical caps to be disjoint and with equal volume, and we let K_0 be the remaining part of the unit ball, as shown in Figure 2.

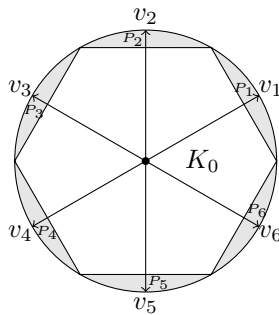


Figure 2 The lower bound construction in two dimensions with $n = 6$. For any $x \in \{0, 1\}^6$, the convex body K_x is formed by taking the union of K_0 and all spherical caps P_j if and only if the corresponding bit x_j is 1.

The core idea is to associate a convex body K_x to every bitstring $x \in \{0, 1\}^n$, defined as the union of K_0 , and all the spherical caps P_j if and only if $x_j = 1$. We then immediately observe that $\text{Vol}(K_x) = \text{Vol}(K_0) + |x| \text{Vol}(P_1)$, and $\text{Vol}(K_x \Delta K_y) = |x \oplus y| \text{Vol}(P_1)$.

Intuitively, now, if we find an approximation to the convex body K_x , then we also find an approximation to the bitstring x . Similarly, if we find a sufficiently precise estimate of K_x 's volume, we also obtain a good approximation of x 's Hamming weight. In short, we can reduce the bitstring recovery problem to convex set estimation, and the Hamming weight estimation problem to volume estimation, and hence lower bounds on the former imply lower bounds on the latter.

4.1 Query complexity of bitstring problems

We first recall the bitstring recovery and the Hamming weight estimation problems. These are well-studied in the existing literature, albeit a bit scattered, and we gather the existing results in two concise theorem statements. The proofs can be found in Section A.

► **Theorem 15** (Bitstring recovery problem). *Let $n \in \mathbb{N}$, and let $x \in \{0, 1\}^n$ be a bitstring that we can access through bit queries. Suppose we wish to output a bitstring $y \in \{0, 1\}^n$ such that $|x \oplus y| \leq n/4$. The query complexities for this problem are $\Theta(n)$ in the deterministic, randomized and quantum setting.*

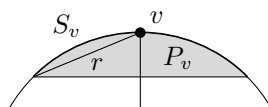
Next, we consider the Hamming weight estimation problem. This problem is sometimes also referred to as the *approximate counting problem*. In short, given a bitstring of length n , we wish to estimate the number of 1's it contains up to some additive precision k . Its hardness is also well-understood – we gather the previous known results in Theorem 16.

► **Theorem 16** (Hamming weight estimation problem). *Let $n \in \mathbb{N}$ and let $x \in \{0, 1\}^n$ be a bitstring that we can access through bit queries. Let $k \in \mathbb{N}$ be such that $1 \leq k \leq n/4$. Suppose we wish to output $w \in [n]_0$ such that $||x| - w| \leq k$. The query complexities for this problem are $\Theta(n)$ in the deterministic setting, $\Theta(\min\{(n/k)^2, n\})$ in the randomized setting, and $\Theta(n/k)$ in the quantum setting.*

4.2 Reduction to convex set estimation and volume estimation

We start by formalizing the concept of a *spherical cap*, and provide a visualization in Figure 3.

► **Definition 17** (Spherical cap). *Let $d \in \mathbb{N}$, $v \in S_{d-1}$ and $r > 0$. Let $S_v = \{u \in S_{d-1} : \|u - v\| < r\}$, and let P_v be its convex hull. Then P_v is the spherical cap around v with radius r .*



■ **Figure 3** The shaded region is a spherical cap around v with radius r .

Next, we prove some properties of spherical caps.

► **Lemma 18** (Properties of spherical caps). *Let $d \in \mathbb{N}$, $v, w \in S_{d-1}$, $r > 0$, and P_v and P_w the spherical caps with radius r around v and w , respectively. Then,*

1. $P_v = \{x \in B_d : x^T v > 1 - r^2/2\}$.
2. If $\|v - w\| \geq 2r$, then P_v and P_w are disjoint.
3. $\text{Vol}(P_v) \in \Theta(r^{d+1})$, in the limit where d is fixed and $r \downarrow 0$.

61:14 How to Compute the Volume in Low Dimension?

Proof. For the first claim, observe that $x \in S_v$ if and only if $x^T v = (\|x\|^2 + \|v\|^2 - \|x - v\|^2)/2 > 1 - r^2/2$. Consequently, $x \in P_v$ if and only if it is inside the unit ball and satisfies this inequality.

For the second claim, we prove the contrapositive, i.e., we prove that $z \in P_v \cap P_w \Rightarrow \|v - w\| < 2r$. To that end, let $z \in P_v \cap P_w$. Then, using that $\|z\| \leq 1$ and Cauchy-Schwarz's inequality, we obtain

$$\begin{aligned} \|v - w\|^2 &= 2\|v\|^2 + 2\|w\|^2 - \|v + w\|^2 \leq 4 - \|z\|^2 \|v + w\|^2 \\ &\leq 4 - (z^T(v + w))^2 < 4 - (2 - r^2)^2 \leq 4r^2. \end{aligned}$$

Finally, computing the volume yields

$$\text{Vol}(P_v) = \Gamma_{d-1} \int_{1-r^2/2}^1 (1-x^2)^{\frac{d-1}{2}} dx = \Gamma_{d-1} \int_0^{r^2/2} (x(2-x))^{\frac{d-1}{2}} dx \in \Theta(r^{d+1}). \quad \blacktriangleleft$$

We now use the above properties to compute how many spherical caps we can pack on the surface of a unit ball.

► **Lemma 19** (Packing of spherical caps). *Let $d \in \mathbb{N}_{\geq 2}$ and let $\delta > 0$. Let $\{v_j\}_{j=1}^n$ be a δ -net, for each $j \in [n]$ let $P_j := P_{v_j}$ be a spherical cap with radius $\delta/2$, and let $P = \{P_j\}_{j=1}^n$. Then, $n = \Theta(\delta^{-(d-1)})$, all the spherical caps are disjoint, and the total volume of the spherical caps is $\text{Vol}(P) := n \text{Vol}(P_1) = \Theta(\delta^2)$.*

Proof. The claims follow immediately from Proposition 4 and Lemma 18. ◀

Finally, we show how embedding a bitstring in these spherical caps naturally leads to a lower bound on the query complexity of the volume estimation and the convex set estimation problems.

► **Theorem 20.** *Let $d \in \mathbb{N}_{\geq 2}$ and $\varepsilon > 0$. Then, any algorithm \mathcal{A} that solves the convex set estimation problem or the volume estimation problem, must make a number of queries that satisfies the complexities listed in Table 1.*

Proof. We start with the lower bounds on the convex set estimation problems. To that end, suppose that an algorithm \mathcal{A} solves it with relative Nikodym distance ε .

For every $\delta > 0$, we choose a particular δ -net $\{v_j\}_{j=1}^n$, and then we let $\{P_j\}_{j=1}^n$ and P be as in Lemma 19. Observe that $\text{Vol}(P) \in \Theta(\delta^2)$. We now choose the smallest $\delta > 0$ such that $\text{Vol}(P) \geq 8\varepsilon\Gamma_d$, which for small enough $\varepsilon > 0$ is guaranteed to be well-defined. Then $\delta \in \Theta(\sqrt{\varepsilon})$, which implies that $n \in \Theta(\varepsilon^{-(d-1)/2})$.

Now, for any bitstring $x \in \{0, 1\}^n$, we define a convex body K_x by

$$K_x = K_0 \cup \bigcup_{\substack{j=1 \\ x_j=1}}^n P_j, \quad (2)$$

Then, given any bitstring $x \in \{0, 1\}^n$, \mathcal{A} outputs \tilde{K} , an approximation of K_x with relative precision ε . Next, we can define a bitstring $z \in \{0, 1\}^n$ such that

$$\text{Vol}(\tilde{K} \Delta K_z) = \min\{\text{Vol}(\tilde{K} \Delta K_y) : y \in \{0, 1\}^n\} \leq \text{Vol}(\tilde{K} \Delta K_x) \leq \varepsilon \text{Vol}(K_x) \leq \varepsilon\Gamma_d.$$

Then, we obtain that

$$|x \oplus z| = \frac{\text{Vol}(K_z \Delta K_x)}{\text{Vol}(P_1)} \leq \frac{n(\text{Vol}(\tilde{K} \Delta K_z) + \text{Vol}(\tilde{K} \Delta K_x))}{\text{Vol}(P)} \leq \frac{n(\varepsilon\Gamma_d + \varepsilon\Gamma_d)}{8\varepsilon\Gamma_d} = \frac{n}{4}.$$

Finally, we observe that any call to the membership oracle of K_x can be simulated with at most one query to the bitstring x . Thus, according to Theorem 15, \mathcal{A} must make at least $\Omega(n) = \Omega(\varepsilon^{-(d-1)/2})$ queries to the membership oracle of K_x , in the deterministic, randomized and quantum settings. Finally, since the ε -kernel construction problem is harder than the ε -Nikodym construction problem, the lower bounds from the first two columns in Table 1 follow.

It remains to prove the lower bounds for volume estimation. To that end, suppose that \mathcal{A} solves the volume estimation problem up to precision ε . Again for every $\delta > 0$ we choose a particular δ -net $\{v_j\}_{j=1}^n$, and we let $\{P_j\}_{j=1}^n$ and P as in Lemma 19, which means that $\text{Vol}(P) \in \Theta(\delta^2)$. Next, we let $\varepsilon' > 0$ arbitrarily and choose the smallest $\delta > 0$ such that $\text{Vol}(P) \geq 8\varepsilon'\Gamma_d$. For small enough ε' , this is well-defined, and we find that $\delta \in \Theta(\sqrt{\varepsilon'})$, from which we find that $n \in \Theta((\varepsilon')^{-(d-1)/2})$.

We let $x \in \{0, 1\}^n$ and K_x as in Equation (2). Then,

$$\text{Vol}(K_x) = \text{Vol}(K_0) + |x| \text{Vol}(P_1).$$

\mathcal{A} outputs an estimate \tilde{V} of $\text{Vol}(K_x)$, such that $|\tilde{V} - \text{Vol}(K_x)| \leq \varepsilon \text{Vol}(K_x) \leq \varepsilon\Gamma_d$. Hence, we can compute $w = (\tilde{V} - \text{Vol}(K_0)) / \text{Vol}(P_1)$, and we observe that

$$|w - |x|| = \left| \frac{\tilde{V} - \text{Vol}(K_0)}{\text{Vol}(P_1)} - \frac{\text{Vol}(K_x) - \text{Vol}(K_0)}{\text{Vol}(P_1)} \right| = \frac{n|\tilde{V} - \text{Vol}(K_x)|}{\text{Vol}(P)} \leq \frac{n\varepsilon\Gamma_d}{8\varepsilon'\Gamma_d} = \frac{n\varepsilon}{8\varepsilon'} =: k.$$

Finally, similarly as before, since we can simulate any query to a membership oracle to K_x by one query to the bitstring x , we can use the lower bounds from Theorem 16 to lower bound the query complexity of \mathcal{A} .

For the lower bounds of Theorem 16 to apply, we have to choose $\varepsilon' > 0$ as a function of ε such that $1 \leq k \leq n/4$. To that end, we let $\ell = \varepsilon/2$ and $u = n\varepsilon/8$, and observe that we must choose $\varepsilon' \in [\ell, u]$. We choose ε' as ℓ , $\sqrt{\ell u}$ and u , in the deterministic, randomized, and quantum settings, respectively. We remark in the randomized setting that $n/k = 8\varepsilon'/\varepsilon \in \Theta(\sqrt{n})$, and so the lower bound from Theorem 16 becomes $\Omega(n)$. Similarly, in the quantum case, we have $k \in \Theta(1)$, and so the lower bound also becomes $\Omega(n)$. It remains to plug in the asymptotic scaling of n , from which we obtain the following lower bounds for the volume estimation problem:

$$\text{Deterministic: } \Omega(n) = \Omega((\varepsilon')^{-\frac{d-1}{2}}) = \Omega(\varepsilon^{-\frac{d-1}{2}})$$

$$\text{Randomized: } \Omega(n) = \Omega(n^{\frac{4}{d+3}} \cdot n^{\frac{d-1}{d+3}}) = \Omega((\varepsilon')^{-\frac{2(d-1)}{d+3}} \cdot (\varepsilon'/\varepsilon)^{\frac{2(d-1)}{d+3}}) = \Omega(\varepsilon^{-\frac{2(d-1)}{d+3}}).$$

$$\text{Quantum: } \Omega(n) = \Omega(n^{\frac{2}{d+1}} \cdot n^{\frac{d-1}{d+1}}) = \Omega((\varepsilon')^{-\frac{d-1}{d+1}} \cdot (\varepsilon'/\varepsilon)^{\frac{d-1}{d+1}}) = \Omega(\varepsilon^{-\frac{d-1}{d+1}}). \quad \blacktriangleleft$$

References

- 1 Pankaj K Agarwal and Sariel Har-Peled. Computing instance-optimal kernels in two dimensions. *Discrete & Computational Geometry*, pages 1–28, 2024.
- 2 Pankaj K Agarwal, Sariel Har-Peled, and Kasturi R Varadarajan. Approximating extent measures of points. *Journal of the ACM (JACM)*, 51(4):606–635, 2004. doi:10.1145/1008731.1008736.
- 3 Andris Ambainis. Quantum lower bounds by quantum arguments. In *Proceedings of the 32nd annual ACM symposium on Theory of computing*, pages 636–643, 2000. doi:10.1145/335305.335394.
- 4 Nikolay Baldin and Markus Reiß. Unbiased estimation of the volume of a convex body. *Stochastic Processes and their Applications*, 126(12):3716–3732, 2016.

- 5 Imre Bárány and Zoltán Füredi. Computing the volume is difficult. *Discret. Comput. Geom.*, 2:319–326, 1987. doi:10.1007/BF02187886.
- 6 Robert Beals, Harry Buhrman, Richard Cleve, Michele Mosca, and Ronald de Wolf. Quantum lower bounds by polynomials. *Journal of the ACM (JACM)*, 48(4):778–797, 2001. doi:10.1145/502090.502097.
- 7 Shalev Ben-David and Eric Blais. A tight composition theorem for the randomized query complexity of partial functions. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*, pages 240–246. IEEE, 2020.
- 8 Gilles Brassard, Peter Hoyer, Michele Mosca, and Alain Tapp. Quantum amplitude amplification and estimation. *Contemporary Mathematics*, 305:53–74, 2002.
- 9 E.M. Bronshteyn and L.D. Ivanov. The approximation of convex sets by polyhedra. *Siberian Mathematical Journal*, 16(5):852–853, 1975.
- 10 Victor-Emmanuel Brunel. Adaptive estimation of convex and polytopal density support. *Probability Theory and Related Fields*, 164(1):1–16, 2016.
- 11 Victor-Emmanuel Brunel. Methods for estimation of convex sets. *Statistical Science*, 33(4):615–632, 2018.
- 12 Shouvanik Chakrabarti, Andrew M. Childs, Shih-Han Hung, Tongyang Li, Chunhao Wang, and Xiaodi Wu. Quantum algorithm for estimating volumes of convex bodies. *ACM Transactions on Quantum Computing*, 4(3):1–60, 2023.
- 13 Timothy M Chan. Faster core-set constructions and data-stream algorithms in fixed dimensions. *Computational Geometry*, 35(1-2):20–35, 2006. doi:10.1016/J.COMGEO.2005.10.002.
- 14 Yuansi Chen. An almost constant lower bound of the isoperimetric coefficient in the KLS conjecture. *Geometric and Functional Analysis*, 31:34–61, 2021.
- 15 Arjan Cornelissen and Yassine Hamoudi. A sublinear-time quantum algorithm for approximating partition functions. In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1245–1264. SIAM, 2023. doi:10.1137/1.9781611977554.CH46.
- 16 Arjan Cornelissen, Yassine Hamoudi, and Sofiene Jerbi. Near-optimal quantum algorithms for multivariate mean estimation. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, pages 33–43, 2022.
- 17 Ben Cousins and Santosh Vempala. Gaussian cooling and $O^*(n^3)$ algorithms for volume and Gaussian volume. *SIAM Journal on Computing*, 47(3):1237–1273, 2018.
- 18 R.M. Dudley. Metric entropy of some classes of sets with differentiable boundaries. *Journal of Approximation Theory*, 10(3):227–236, 1974.
- 19 Martin Dyer, Alan Frieze, and Ravi Kannan. A random polynomial-time algorithm for approximating the volume of convex bodies. *Journal of the ACM (JACM)*, 38(1):1–17, 1991. doi:10.1145/102782.102783.
- 20 György Elekes. A geometric inequality and the complexity of computing volume. *Discrete & Computational Geometry*, 1:289–292, 1986. doi:10.1007/BF02187701.
- 21 Edward Farhi, Jeffrey Goldstone, Sam Gutmann, and Michael Sipser. Bound on the number of functions that can be distinguished with k quantum queries. *Physical Review A*, 60(6):4331, 1999.
- 22 Martin Grötschel, László Lovász, and Alexander Schrijver. *Geometric algorithms and combinatorial optimization*. Springer-Verlag Berlin Heidelberg, 1 edition, 1988.
- 23 Sariel Har-Peled and Mitchell Jones. Proof of Dudley’s convex approximation. *arXiv preprint arXiv:1912.01977*, 2019. arXiv:1912.01977.
- 24 Arun Jambulapati, Yin Tat Lee, and Santosh S. Vempala. A slightly improved bound for the KLS constant. *arXiv preprint arXiv:2208.11644*, 2022. doi:10.48550/arXiv.2208.11644.
- 25 He Jia, Aditi Laddha, Yin Tat Lee, and Santosh Vempala. Reducing isotropy and volume to KLS: an $O^*(n^3\psi^2)$ volume algorithm. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, pages 961–974, 2021.
- 26 Bo’az Klartag. Logarithmic bounds for isoperimetry and slices of convex sets. *arXiv preprint arXiv:2303.14938*, 2023.

- 27 Bo'az Klartag and Joseph Lehec. Bourgain's slicing problem and KLS isoperimetry up to polylog. *Geometric and Functional Analysis*, 32(5):1134–1159, 2022.
- 28 László Lovász and Santosh Vempala. Simulated annealing in convex bodies and an $O^*(n^4)$ volume algorithm. *Journal of Computer and System Sciences*, 72(2):392–417, 2006. doi:10.1016/J.JCSS.2005.08.004.
- 29 Ashwin Nayak and Felix Wu. The quantum query complexity of approximating the median and related statistics. In *Proceedings of the thirty-first annual ACM symposium on Theory of computing*, pages 384–393, 1999. doi:10.1145/301250.301349.
- 30 Michael A. Nielsen and Isaac L. Chuang. *Quantum computation and quantum information*. Cambridge University Press, 2010.
- 31 Luis Rademacher and Santosh Vempala. Dispersion of mass and the complexity of randomized geometric algorithms. *Advances in Mathematics*, 219(3):1037–1069, 2008.
- 32 Carsten Schütt. Random polytopes and affine surface area. *Mathematische Nachrichten*, 170(1):227–249, 1994.
- 33 Miklós Simonovits. How to compute the volume in high dimension? *Mathematical programming*, 97:337–374, 2003. doi:10.1007/S10107-003-0447-X.
- 34 Hai Yu, Pankaj K Agarwal, Raghunath Poreddy, and Kasturi R Varadarajan. Practical methods for shape fitting and kinetic data structures using coresets. *Algorithmica*, 52:378–402, 2008. doi:10.1007/S00453-007-9067-9.

A Query complexities of bitstring problems

In this appendix, we provide the proofs to the query complexities of the bitstring problems that we use in the lower bound constructions in Section 4. We restate the theorems for convenience.

► **Theorem 15** (Bitstring recovery problem). *Let $n \in \mathbb{N}$, and let $x \in \{0, 1\}^n$ be a bitstring that we can access through bit queries. Suppose we wish to output a bitstring $y \in \{0, 1\}^n$ such that $|x \oplus y| \leq n/4$. The query complexities for this problem are $\Theta(n)$ in the deterministic, randomized and quantum setting.*

Proof. The (deterministic) algorithm is trivial – simply query all the entries of x and output the bitstring exactly. Since the randomized and quantum settings can simulate the deterministic setting, this provides the upper bounds on the query complexities.

It remains to prove that this algorithm is optimal in all three models up to constant multiplicative factors. Since we can simulate any deterministic and randomized algorithm in the quantum setting, proving hardness in the latter suffices. An information-theoretic argument tells us that quantumly we need at least $\Omega(n)$ queries to output a bitstring that differs from x in at most $n/4$ positions. The core technique stems from [21], and all remaining details can for instance be found in [16, Lemma 4.6]. ◀

► **Theorem 16** (Hamming weight estimation problem). *Let $n \in \mathbb{N}$ and let $x \in \{0, 1\}^n$ be a bitstring that we can access through bit queries. Let $k \in \mathbb{N}$ be such that $1 \leq k \leq n/4$. Suppose we wish to output $w \in [n]_0$ such that $||x| - w| \leq k$. The query complexities for this problem are $\Theta(n)$ in the deterministic setting, $\Theta(\min\{(n/k)^2, n\})$ in the randomized setting, and $\Theta(n/k)$ in the quantum setting.*

Proof. We first prove the upper bounds. The deterministic upper bound is trivial. In the randomized setting, we need to show that the query complexity is $O((n/k)^2)$. To that end, suppose that we sample $3(n/k)^2$ bits at random, with replacement, and output n times the fraction of 1's observed. Let X_j be the random variable describing the output of the j th

61:18 How to Compute the Volume in Low Dimension?

sample, and we write the output of the algorithm as $w := n/3 \cdot (k/n)^2 \cdot \sum_{j=1}^{3(n/k)^2} X_j$. Then, we trivially find that $\mathbb{E}[w] = |x|$, and its variance can be bounded as

$$\text{Var}[w] = \text{Var} \left[\frac{n}{3} \left(\frac{k}{n} \right)^2 \cdot \sum_{j=1}^{3(n/k)^2} X_j \right] = \frac{n^2}{9} \left(\frac{k}{n} \right)^4 \cdot 3 \left(\frac{n}{k} \right)^2 \text{Var}[X_1] = \frac{1}{3} k^2 \text{Var}[X_1] \leq \frac{1}{3} k^2,$$

and so by Chebyshev's inequality, we have

$$\mathbb{P}[|w - |x|| \geq k] \leq \frac{\text{Var}[w]}{k^2} \leq \frac{1}{3}.$$

Finally, quantumly, the algorithm is known as the quantum approximate counting algorithm [8], and the upper bound follows directly.

Then, it remains to prove the lower bounds. Deterministically, the problem is easiest when k is largest, so it suffices to prove the lower bound when $k = n/4$. Once we have queried $n/2 - 1$ bits of the bitstring – let's say we observed w ones – there are still $n/2 + 1$ bits left to query. Then, we know that the true Hamming weight must be between w and $w + n/2 + 1$. Since this interval is larger than $2k$, we cannot guarantee that whatever number we output is at most k away from the true Hamming weight. Thus, we need to make at least $n/2 \in \Omega(n)$ queries.

In the randomized setting, we use [7, Lemma 26]. If $k = \sqrt{n}$, we need at least $\Omega(n)$ queries. Since the problem becomes more difficult when k decreases, this proves the lower bound for all $k \leq \sqrt{n}$. On the other hand, if $k > \sqrt{n}$, then we define $n' \in \Theta((n/k)^2)$. For any bitstring $x' \in \{0, 1\}^{n'}$, we define a bitstring x of length $\Theta(n)$ by duplicating all the bits of x' a total of $n/n' \in \Theta(k^2/n)$ times. Then, a Hamming weight estimation algorithm with precision k will find an approximation of the Hamming weight of x with precision k , which means that it outputs an approximation of the Hamming weight of x' with precision $\Theta(k/(k^2/n)) = \Theta(n/k) = \Theta(\sqrt{n'})$. Thus, this algorithm must query x at least $\Omega(n') = \Omega((n/k)^2)$ times.

Finally, in the quantum setting, the result follows easily from [3]. For $k = 1$, the hardness already follows from the majority function, whose hardness was proved in [6], and the hardness of the general case was basically proved in [29], even though not phrased as such. ◀