Quantum algorithms on convex bodies

Tongyang Li

Department of Computer Science, University of Maryland

arXiv:1809.01731, arXiv:1908.03903

Quantum Algorithms Workshop, Simons Institute for the Theory of Computing
February 25, 2020

Overview

Talks by Richard and Anupam: fast quantum algorithms for solving SDP, cone programming, etc., with *explicit* matrix inputs stored in quantum data structures.

What if we don't have explicit matrix inputs, but only **implicit oracle access?**

In particular, LP, SDP, and many others have convex feasible region $\mathrm{K} \subseteq \mathbb{R}^n$.
A common oracle access is membership oracle:

$$\mathrm{MEM}_{\mathrm{K}}(x) = 1 \textit{ if } x \in \mathrm{K} \textit{ and } \mathrm{MEM}_{\mathrm{K}}(x) = 0 \textit{ if } x \notin \mathrm{K}.$$

Main topic today: What we can do on a convex body K with a quantum oracle

$$O_{\mathrm{K}}|x, 0\rangle = |x, \mathrm{MEM}_{\mathrm{K}}(x)\rangle \qquad \forall\, x \in \mathbb{R}^n.$$

First question about a convex body

Given a convex body $\mathrm{K} \subset \mathbb{R}^n$, a very first question may be to ask its volume:

$$\mathrm{Vol}(\mathrm{K}) := \int_{x \in \mathrm{K}} \mathrm{d}x.$$

Assume $0 < \epsilon < 1/2$, the goal is to return a value $\in [(1-\epsilon)\,\mathrm{Vol}(\mathrm{K}), (1+\epsilon)\,\mathrm{Vol}(\mathrm{K})]$ with success probability at least $2/3$.

Why do we study this problem?

- One of the most basic problems in geometry;
- Can be viewed as a *continuous version of counting*, and quantum counting is widely applied in Grover-type algorithms.

Formulation

Input model: Quantum membership oracle.

$$O_{\mathrm{K}}|x,0\rangle = |x,\delta[x\in\mathrm{K}]\rangle \qquad \forall\, x\in\mathbb{R}^n.$$

Constraints on K: $0\in\mathrm{K}$, and known inner and outer bounds on K:

$$\mathrm{B}_2(0,r)\subseteq\mathrm{K}\subseteq\mathrm{B}_2(0,R),$$

where $\mathrm{B}_2(x,l)$ is the ball of radius l in ℓ_2-norm centered at $x\in\mathbb{R}^n$.

Classically, the best-known query complexity is $\tilde{O}(n^4)$ (Lovász and Vempala, 2003).

Main result

Main Theorem

There is a quantum algorithm that returns a value $\widetilde{\mathrm{Vol}(\mathrm{K})}$ *satisfying*

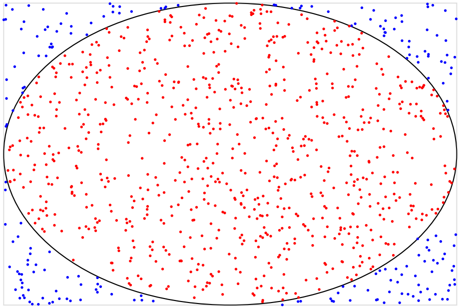
$$(1-\epsilon)\,\mathrm{Vol}(\mathrm{K}) \leq \widetilde{\mathrm{Vol}(\mathrm{K})} \leq (1+\epsilon)\,\mathrm{Vol}(\mathrm{K})$$

using $\tilde{O}(n^3 + n^{2.5}/\epsilon)$ *quantum queries to* O_{K}*, where* $\tilde{O}$ *omits poly-logarithm factors.*

	Classical bounds	Quantum bounds (our result)
Query complexity	$\tilde{O}(n^4 + n^3/\epsilon^2)$, $\tilde{\Omega}(n^2)$	$\tilde{O}(n^3 + n^{2.5}/\epsilon)$, $\Omega(\sqrt{n})$
Arithmetic complexity	$\tilde{O}(n^6 + n^5/\epsilon^2)$	$\tilde{O}(n^5 + n^{4.5}/\epsilon)$

Classical volume estimation algorithm

Naive algorithm: Monte Carlo, brute-force counting.



Because Vol(K) takes an exponentially large range: $[r^n, R^n]$ (up to a constant), the brute-force counting argument takes exponential cost.

Classical volume estimation algorithm

Instead, consider the parameterized value

$$Z(a) := \int_{\mathrm{K}} e^{-a\|x\|_2} \, \mathrm{d}x.$$

- On the one hand, $Z(0) = \mathrm{Vol}(\mathrm{K})$.
- On the other hand, because $e^{-\|x\|_2}$ decays exponentially fast with $\|x\|_2$, a large enough a ensures that $Z(a)$ is closed to a fixed value:

$$Z(a) \approx \int_{\mathrm{B}_2(0,r)} e^{-a\|x\|_2} \, \mathrm{d}x.$$

Classical volume estimation algorithm

Therefore, a natural strategy is **simulated annealing (SA)**: consider a sequence $a_0 > a_1 > \cdots > a_m$ with a_0 sufficiently large and a_m close to 0. Iteratively changes a_i to a_{i+1} and estimates $\mathrm{Vol}(\mathrm{K})$ by the telescoping product

$$\mathrm{Vol}(\mathrm{K}) \approx Z(a_m) = Z(a_0) \prod_{i=0}^{m-1} \frac{Z(a_{i+1})}{Z(a_i)}.$$

How to estimate each telescoping ratio? **Take i.i.d. samples and average.**

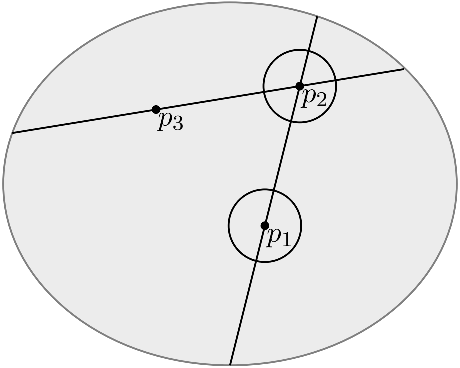
Sampling: in the i^{th} step, use a **random walk** to sample over K with density $\propto e^{-a_i \|x\|_2}$. Denote the output as X_i, and let $V_i := e^{(a_i - a_{i+1})}\|X_i\|_2$. Then

$$\mathbb{E}[V_i] = \int_{\mathrm{K}} e^{(a_i - a_{i+1})\|x\|_2} \frac{e^{-a_i\|x\|_2}}{Z(a_i)} \,\mathrm{d}x = \int_{\mathrm{K}} \frac{e^{-a_{i+1}\|x\|_2}}{Z(a_i)} \,\mathrm{d}x = \frac{Z(a_{i+1})}{Z(a_i)}.$$

Hit-and-run walk

One particular random walk in a convex body: hit-and-run.

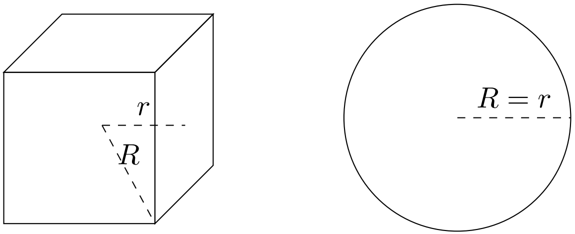
1. Pick a uniformly distributed random line ℓ through the current point p;
2. Move to a random point along the chord $\ell \cap \mathrm{K}$ with density $\propto e^{-a\|x\|_2}$.



Summary

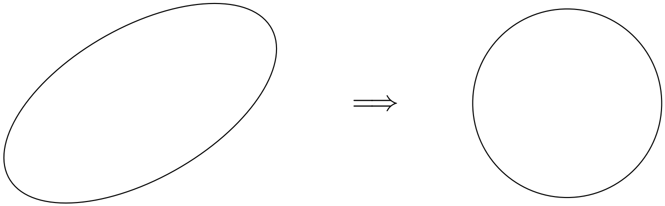
The classical algorithm by Lovász and Vempala [LV06]:

- The simulated annealing algorithm goes through $\tilde{O}(\sqrt{n})$ iterations;
- Each iteration takes $\tilde{O}(\sqrt{n}/\epsilon^2)$ i.i.d. samples;
- Each sample takes $\tilde{O}(n^3)$ steps of the hit-and-run walk if the convex body is *well-rounded (i.e., $R/r = O(\sqrt{n})$).*



Summary

For any convex body, they construct an affine transformation $\mathrm{K} \to A\mathrm{K} + b$ that makes the convex body well-rounded, using $\tilde{O}(n^4)$ queries.



In total, the query complexity is

$$\tilde{O}(n^4) + \tilde{O}(\sqrt{n}) \cdot \tilde{O}(\sqrt{n}/\epsilon^2) \cdot \tilde{O}(n^3) = \tilde{O}(n^4/\epsilon^2).$$

Classical volume estimation algorithm

Three key factors of a successful volume estimation algorithm:

1) High level: The algorithm follows an SA framework, where the volume is estimated by a telescoping product.

2) Middle level: The number of i.i.d. samples used to estimate each telescoping ratio is small, in other words $\mathrm{Var}[V_i]$ is small.

3) Low level: The random walk converges fast so that we can take each i.i.d. sample of V_i efficiently.

Quantum volume estimation algorithm

Three key factors of a successful volume estimation algorithm:

1) **High level**: The algorithm follows an SA framework, where the volume is estimated by a telescoping product. A quantum SA framework?

2) **Middle level**: The number of i.i.d. samples used to estimate each telescoping ratio is small, in other words $\mathrm{Var}[V_i]$ is small. Speedup by quantum counting?

3) **Low level**: The random walk converges fast so that we can take each i.i.d. sample of V_i efficiently. Speedup by quantum walk?

Quantum volume estimation algorithm: low level

Previous quantum walks are mainly in discrete spaces. How about continuous-space quantum walks?

Theorem. Szegedy's theory can be generalized to the continuous case. Furthermore, there is Cheeger's inequality for continuous-space Markov chains:

$$\Phi_p \leq \sqrt{2\Delta_p}.$$

Error analysis: We also introduce a discretized hit-and-run walk with conductance lower bound and implementation details.

Classical hit-and-run walk converges in time $\tilde{O}(n^3)$ $\Longrightarrow$ Quantum hit-and-run walk has spectral gap $\tilde{\Omega}(n^{-1.5})$

Next question: How to use the quantum hit-and-run walk in an SA procedure?

Quantum volume estimation algorithm: high level

We follow a quantum SA framework by Wocjan and Abeyesinghe:

Given quantum walk operators $W_1, \ldots, W_r$ with stationary states $|\pi_1\rangle, \ldots, |\pi_r\rangle$. Assume $|\langle\pi_i|\pi_{i+1}\rangle|^2 \geq p$ for all $i \in [r-1]$, and Δ lower bounds the spectral gaps of $W_1, \ldots, W_r$. Question: Given $|\pi_1\rangle$, how to prepare $|\pi_r\rangle$?

Let $\Pi_i := |\pi_i\rangle\langle\pi_i|$, $\Pi_i^\perp := I - \Pi_i$. $\pi/3$-amplitude amplification (which gives fixed-point search):

$$R_i := \omega\Pi_i + \Pi_i^\perp \qquad \text{where} \quad \omega = e^{i\frac{\pi}{3}}.$$

It can be shown that $|\langle\pi_2|R_1R_2|\pi_1\rangle|^2 \geq 1 - (1-p)^3$.

Quantum volume estimation algorithm: high level

Recursively: $U_0 = I,\ U_{m+1} = U_m\, R_i\, U_m^\dagger\, R_{i+1}\, U_m$. Then we have

$$|\langle \pi_{i+1}|U_m|\pi_i\rangle|^2 \geq 1 - (1-p)^{3^m},$$

and the unitaries in $\{R_i, R_i^\dagger, R_{i+1}, R_{i+1}^\dagger\}$ are used at most 3^m times in U_m.

Quantum SA: Apply $\pi/3$-am-am throughout: $|\pi_1\rangle \to |\pi_2\rangle \to \cdots \to |\pi_r\rangle$. Take the SA slow enough so that $p = \Omega(1)$, making $m = \tilde{O}(1)$ in each iteration.

Total complexity: $\tilde{O}(r/p\Delta)$ calls to the quantum walk operators $W_1, \ldots, W_r$.

Quantum volume estimation algorithm: high level

In our quantum algorithm for volume estimation:

$$\begin{aligned} r &: \quad \tilde{O}(\sqrt{n}) \text{ iterations} \\ W_i &: \quad \text{Quantum hit-and-run walk with density} \propto e^{-a_i\|x\|_2} \\ \pi_0 &: \quad \text{Easy to prepare, } \approx \text{density} \propto e^{-2n\|x\|_2} \text{ on the small ball} \\ \Delta &: \quad \text{Spectral gaps at least } \tilde{\Omega}(n^{-1.5}) \end{aligned}$$

Remain question: How many copies of $|\pi_i\rangle$ should we prepare, such that we give a good estimate of each telescoping ratio by **quantum counting**?

Quantum volume estimation algorithm: middle level

Quantum counting: Estimate p within ϵ using $O(1/\epsilon)$ calls to U and $U^\dagger$ where $U|0\rangle|0\rangle = \sqrt{p}|0\rangle|\phi\rangle + |0^\perp\rangle$, quadratic speedup compared to classical $O(1/\epsilon^2)$.

Quantum Chebyshev inequality: Assume that U is a unitary such that $U|0\rangle|0\rangle = \sum_{x\in\Omega}\sqrt{p_x}|\psi_x\rangle|x\rangle$, where $\{|\psi_x\rangle : x \in \Omega\}$ are unit vectors. Denote

$$\mu_U := \sum_{x\in\Omega} p_x x, \qquad \sigma_U^2 := \sum_{x\in\Omega} p_x (x-\mu_U)^2.$$

Then Hamoudi and Magniez can output an estimate $\tilde{\mu}_U$ such that $|\tilde{\mu}_U - \mu_U| \leq \epsilon\mu_U$ w.h.p. using $\tilde{O}(\sigma_U/\epsilon\mu_U)$ calls to U and $U^\dagger$.

For volume estimation: $\sigma_U^2 = O(1)\cdot\mu_U^2$ (such process is called Chebyshev cooling).

$\tilde{O}(\sqrt{n}/\epsilon^2)$ classical samples from π_i by Chebyshev inequality	$\Longrightarrow$	$\tilde{O}(\sqrt{n}/\epsilon)$ quantum copies of states $\|\pi_i\rangle$ by quantum Chebyshev inequality

Summary

- **High level**: We adopt a quantum SA framework proposed by [WA08], where the volume is estimated by a telescoping product. # of iterations: $\tilde{O}(\sqrt{n})$.
- **Middle level**: We estimate each ratio in the telescoping product using the nondestructive version of the quantum Chebyshev inequality. # of calls to the quantum hit-and-run walk operator: $\tilde{O}(\sqrt{n}/\epsilon)$.
- **Low level**: # of queries to implement one step of quantum walk: $\tilde{O}(n^{1.5})$.

We also give a rounding algorithm using $\tilde{O}(n^3)$ quantum queries.

Total query complexity: $\tilde{O}(n^3) + \tilde{O}(\sqrt{n}) \cdot \tilde{O}(\sqrt{n}/\epsilon) \cdot \tilde{O}(n^{1.5}) = \tilde{O}(n^3 + n^{2.5}/\epsilon)$.

# of additional arithmetic operations: $\tilde{O}(n^3 + n^{2.5}/\epsilon) \cdot O(n^2) = \tilde{O}(n^5 + n^{4.5}/\epsilon)$ due to the affine transformations for rounding (n-dim matrix-vector products).

Next question: convex optimization

Next, what if we are given a convex function $f\colon \mathbb{R}^n \to \mathbb{R}$ and we want to optimize f on K? The goal is to return an $\tilde{x} \in \mathrm{K}$ such that

$$f(\tilde{x}) \leq \min_{x \in \mathrm{K}} f(x) + \epsilon.$$

The oracle for f is the **evaluation oracle:**

$$O_f|x\rangle|0\rangle = |x\rangle|f(x)\rangle \qquad \forall\, x \in \mathbb{R}^n.$$

Classically, it is well-known that such an $\tilde{x}$ can be found in polynomial time. Currently, the state-of-the-art result by Lee, Sidford, and Vempala uses $\tilde{O}(n^2)$ queries and runs in time $\tilde{O}(n^3)$.

Convex optimization: main result

Main result. Convex optimization takes

- $\tilde{O}(n)$ and $\Omega(\sqrt{n})$ quantum queries to O_{K};
- $\tilde{O}(n)$ and $\tilde{\Omega}(\sqrt{n})$ quantum queries to O_f.

Furthermore, the quantum algorithm runs in time $\tilde{O}(n^3)$.[1]

Outcome:

- The first nontrivial quantum upper bound on general convex optimization.
- Impossibility of generic exponential quantum speedup of convex optimization! The speedup is at most polynomial.

[1] Similar results are independently obtained by van Apeldoorn, Gilyén, Gribling, and de Wolf.

Convex optimization: reduction to a linear evaluation function

A common trick: reduce to a problem with a linear objective function

$$\min_{x' \in \mathbb{R},\, x \in \mathrm{K}} x' \quad \text{such that} \quad f(x) \leq x' \leq M.$$

Observe that a membership query to the new convex set

$$\mathrm{K}' := \{(x', x) \in \mathbb{R} \times \mathrm{K} \mid f(x) \leq x' \leq M\}$$

can be implemented with one query to the membership oracle O_{K} and one query to the evaluation oracle O_f.

Convex optimization: reduction to a linear evaluation function

As a result, it suffices to optimize a linear function

$$\min_{x \in \mathrm{K}} \; c^T x$$

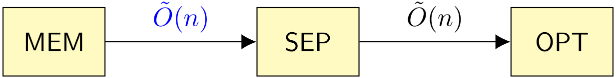
for $c \in \mathbb{R}^n$ and $\mathrm{K} \subseteq \mathbb{R}^n$, given membership oracle O_{K}.

Two other oracles:

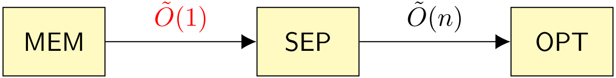
- *Optimization oracle*: Input c, solve the problem.
- *Separation oracle:* Input $x \in \mathbb{R}^n$, return a hyperplane separating x from K if $x \notin \mathrm{K}$, and return $x \in \mathrm{K}$ otherwise.

Convex optimization: oracle reductions

Lee-Sidford-Vempala gives classical oracle reductions:



We give corresponding quantum oracle reductions:



Both papers use the same cutting plane based reduction from OPT to SEP. We show an improved upper bound by reducing the query complexity of the reduction from SEP to MEM.

Separation to Membership

The reduction from SEP to MEM relies on the following geometric idea:

- For some $y \notin \mathrm{K}$, consider a line segment L in the direction y from 0.
- Let L intersect the boundary of K at p.
- Return a supporting hyperplane of K at p.

Such a hyperplane can be found from a subgradient at 0 of a *height function*:

$$h_y(x) = -\max\left\{t \mid x + \frac{ty}{\|y\|_2} \in \mathrm{K}\right\}.$$

The problem thus reduces to one of finding approximate subgradients of an arbitrary L-Lipschitz, convex function f.

Separation to Membership

Classically gradient computation takes $\tilde{\Theta}(n)$ evaluation queries. For the quantum speedup, Jordan's algorithm computes gradient using $\tilde{O}(1)$ quantum queries:

- Given the oracle $O_f|x\rangle|0\rangle = |x\rangle|f(x)\rangle$, the state $e^{if(x)}|x\rangle$ can be prepared by one query to O_f via phase kickback. WLOG we take the gradient at 0.
- Smooth functions have Taylor approximation $f(x) \approx \sum_{k=1}^{n} \frac{\partial f}{\partial x_k} x_k$; hence

$$\sum_x e^{if(x)}|x\rangle \approx \sum_x \bigotimes_{k=1}^{n} e^{i\frac{\partial f}{\partial x_k} x_k}|x_k\rangle.$$

Applying the QFT on all n coordinates reveals $\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}$.

Technical contribution: generalize Jordan's algorithm to subgradient computation using mollification of the convex function + a randomness trick.

Conclusions

Main result 1: an algorithm that gives ϵ-approx. to Vol(K) using $\tilde{O}(n^3 + n^{2.5}/\epsilon)$ quantum queries to O_{K} and $\tilde{O}(n^5 + n^{4.5}/\epsilon)$ additional arithmetic operations.

Main result 2: an algorithm that gives ϵ-approx. to $\min_{x \in \mathrm{K}} c^T x$ using $\tilde{O}(n)$ quantum queries to O_{K} and $\tilde{O}(n^3)$ additional arithmetic operations.

Open questions:

- Can we apply our simulated annealing framework to solve other problems, such as partition function estimation?
- Can we give fast quantum algorithms for other types of optimization problems, in particular some kind of nonconvex optimization problems?
- Can we prove better quantum lower bounds? For both problems, the current bound $\Omega(\sqrt{n})$ is due to Grover search.