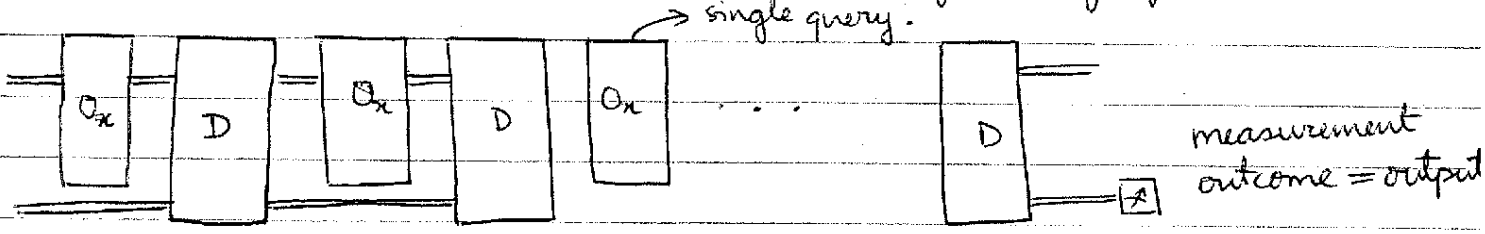
May 12, 2011 SDP in Quantum Information

Ashwin Nayak

Quantum query complexity

One of the first quantum algorithms to be discovered was the Grover algorithm for unordered search. This algorithm may be viewed as searching for a '1' in a list of n bits. Equivalently, it may be viewed as an algorithm that computes the logical OR of n bits. It has the following form:



alternately

where we query an "oracle" for the value x_i of the ith bit like so:

$|i\rangle|w\rangle \mapsto (-1)^{x_i}|i\rangle|w\rangle$ (where w is the content of the "workspace")

and apply a unitary D to the algorithm qubits. The algorithm may query in superposition and this leads to a quadratic speed-up over classical algorithms. How many queries do we need to evaluate a general function $f:\{0,1\}^n \to \{0,1\}$ using such an algorithm? A strong lower bound was proposed by Hoyer, Lee, Spalek, building on work that started with one by Ambainis:

Quantum Adversary bound

Let $Q_\varepsilon(f)$ denote the minimum number of queries required to determine $f(x)$ for an arbitrary $x \in \{0,1\}^n$, with error $\leq \varepsilon < \frac{1}{2}$.

Theorem $Q_\varepsilon(f) \geq \left(\frac{1-2\sqrt{\varepsilon(1-\varepsilon)}}{2}\right) \cdot Adv^{\pm}(f)$,

where $Adv^{\pm}(f) = \max \|\Gamma\|$ (standard operator norm $= \max_{\|\beta\|=1} |\langle\beta|\Gamma|\beta\rangle|$)

such that $\Gamma \in \mathbb{R}^{X\times X}$ symmetric, where $X \subseteq \{0,1\}^n$,

$\Gamma \circ F = 0 \iff \Gamma_{x,y} = 0$ if $f(x) = f(y)$, $\forall x,y \in X$,

$\|\Gamma \circ F_j\| \leq 1$, $\forall j = 1,\dots,n$

where $F_j = \sum_{x,y:\, x_j \neq y_j} |x\rangle\langle y|$.

$F = \sum_{\substack{x,y: \\ f(x)=f(y)}} |x\rangle\langle y|$

& '$\circ$' denotes the entrywise / Hadamard product of matrices

Although this is not obvious at first glance, this is an SDP. Here are a few observations that help us see this.

By rearranging the inputs (rows & columns of Γ), we see that it has an anti-diagonal block structure (when feasible):

$$\Gamma = \begin{array}{c} \\ f(x)=0 \\ f(x)=1 \\ \\ f(x)=0 \quad f(x)=1 \end{array} \left[\begin{array}{c|c} 0 & \Gamma_{\neq} \\ \hline \Gamma_{\neq}^T & 0 \end{array}\right]$$

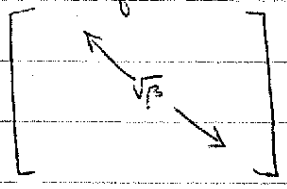
$\Rightarrow \Gamma \circ F_j$ is also of this form. (F_j is symmetric)

Proposition For any symmetric matrix M of the above block anti-diagonal form $\|M\| \leq 1 \iff M \leq I$.

So the constraint $\|\Gamma \circ F_j\| \leq 1$ is a P.S.D. inequality:

$$\Gamma \circ F_j \leq I .$$

The objective function $\|\Gamma\|$: if this were a minimization problem we could replace $\min \|\Gamma\|$ by $\min \gamma : \Gamma \leq \gamma I, -\Gamma \leq \gamma I$. There is an indirect way to express the max norm, as a linear objective function.

Consider $\tilde{\Gamma} = \Delta(\sqrt{\beta}) \cdot \Gamma \cdot \Delta(\sqrt{\beta})$, where $\sqrt{\beta} = (\sqrt{\beta_x})$ is the principal eigenvector of Γ — the one achieving its norm, and $\Delta(\sqrt{\beta}) =$  is the operator with $\sqrt{\beta}$ on its diagonal, 0 elsewhere.

$\tilde{\Gamma}$ is also symmetric, $\tilde{\Gamma} \circ F = 0$, $\tilde{\Gamma} \circ F_j = \Delta(\sqrt{\beta})\,(\Gamma \circ F_j)\,\Delta(\sqrt{\beta})$

So that $\Gamma \circ F_j \leq I \iff \tilde{\Gamma} \circ F_j \leq \Delta(\beta)$

Further $\|\Gamma\| = \mathrm{Tr}\left[J \Delta(\sqrt{\beta})\, \Gamma\, \Delta(\sqrt{\beta})\right] = \mathrm{Tr}(J\tilde{\Gamma})$, where J is the all-1 matrix.

W.l.o.g., no entry of β is 0, and its phase may be absorbed into $\tilde{\Gamma}$. Then $\Delta(\beta) \geq 0$ & $\mathrm{Tr}(\Delta(\beta)) = 1$.

With this correspondence, the adversary bound is seen to be equivalent to:

$$\max \langle J, \tilde{\Gamma} \rangle$$

subject to

$$\tilde{\Gamma} \circ F = 0$$

$$\tilde{\Gamma} \circ F_j \leq \Delta(\beta) \qquad \forall\, j = 1, 2, \cdots, n.$$

$$\Delta(\beta) \geq 0 \qquad \forall\, x \in X$$

$$\sum_x \beta_x = \mathrm{Tr}(\Delta(\beta)) \leq 1$$

$$\beta \in \mathbb{R}^X ; \quad \tilde{\Gamma} \in \mathbb{R}^{X \times X}, \text{ symmetric.}$$

The significance of this SDP formulation was recognized by Reichardt, who pointed out that the dual SDP corresponds to a "span program", a construct studied earlier in a different context in classical complexity theory. He also showed how we could design a quantum query algorithm with complexity equal to the "witness size" of the span program. The witness size equals the objective function in the dual. Along with strong duality — which holds in this case, this shows

(1) That the adversary method is optimal for establishing quantum lower bounds, and

(2) That it suffices to restrict ourselves to designing efficient span programs for a function, in order to design optimal quantum query algorithms. This has the advantage of simplifying the search for efficient algorithms, as span programs have a mathematically simple form.

Span Program

A span program for a Boolean function $f: \{0,1\}^n \to \{0,1\}$ is a vector system $(u_x)_{x \in \{0,1\}^n}$, where each vector $u_x \in \bigoplus_{i=1}^{n} \mathbb{R}^d$, for some d, so that we may view the vector as $u_x = (u_{x,1}, \ldots, u_{x,n})$, where each $u_{x,i} \in \mathbb{R}^d$. Furthermore, for each pair x, y such that $f(x) \neq f(y)$, we have:

$$\sum_{\substack{j=1 \\ \text{s.t } x_j \neq y_j}}^{n} \langle u_{x,j}, u_{y,j} \rangle = 1 \quad .$$

Example $f : \{0,1\}^n \to \{0,1\}$

given as $f(x) = \bigvee_{i=1}^{n} x_i$.

Define $u_{x,i} := \begin{cases} 1 & \text{if } x_i = 1 \text{ \& } i \text{ is the smallest } j : x_j = 1 \\ 0 & \text{otherwise} \end{cases}$ (*), for $x : f(x) = 1$

(* any unique choice of i will suffice)

Also, $u_{x,i} = 1$ for all i, if $x = 0^n$.

We see that $u_x \in \mathbb{R}^n$ for all x, and the constraint is satisfied:

$$\sum_{j : x_j \neq y_j} \langle u_x, u_{y_j} \rangle = 1 \qquad \forall\ x, y \ s.t.\ f(x) \neq f(y).$$

Complexity of a span program.

This is defined as the maximum norm squared of any vector $= \max_x \|u_x\|^2$

The complexity of the above example program is $n = \|u_{0^n}\|^2$.

However, by rescaling the vectors, we may bring this down to $\sqrt{n}$:

$$v_x = u_x / n^{1/4} \quad \text{for } x = 0^n$$

$$\& \quad v_x = u_x \cdot n^{1/4} \quad \text{for } x \neq 0^n.$$

Note that the problem of finding a span program of minimum witness size is an SDP:

$$\inf \ t$$

$$\text{such that} \quad \langle u_x, u_x \rangle = \sum_i \langle u_{x,i}, u_{x,i} \rangle \leq t \quad \forall x$$

$$\& \ \sum_{j : x_j \neq y_j} \langle u_{x,j}, u_{y,j} \rangle = 1 \quad \forall\ x, y : f(x) \neq f(y)$$

This is because the objective function and the constraints are linear over elements of the Gram matrix $U^\dagger U$ of the vectors $(u_{x,i})$:

$$U = \sum_{x,i} |u_{xi}\rangle\langle x,i| \quad .$$

(Reichardt)

Theorem The SDPs for the adversary bound and the span program are dual to each other, and strong duality holds.

Proof: Exercise.

As mentioned before, we can design a quantum algorithm for f using a span program. The query complexity of the algorithm is proportional to the witness size of the span program. The following presentation is due to Lee, Mittal, Reichardt, Spalek & Szegedy. (The original algorithm & its earlier simplifications & improvements were due to Reichardt.)

Algorithms from Span programs

Given a span program (u_x) for $f: \{0,1\}^n \to \{0,1\}$.
Let $A = \max_x \|u_x\|^2$ be its "witness size".

Define two orthogonal projection operations

$$\Delta = \text{projection onto Span} \left\{ |0\rangle + \frac{1}{10\sqrt{A}} \sum_{i=1}^{n} |i\rangle |u_{yi}\rangle |y_i\rangle \right.$$
$$\left. \text{for all } y \text{ such that } f(y) = 1 \right\},$$

where $|0\rangle$ is a vector orthogonal to all the remaining terms in the sum.

Let $\Pi_x = |0\rangle\langle 0| + \sum_{i=1}^{n} |i\rangle\langle i| \otimes I \otimes |x_i\rangle\langle x_i|$.

The algorithm is derived from the product of reflections through the corresponding spaces: $R(\Pi_x) \cdot R(\Delta)$, where $R(\Pi) = 2\Pi - I$ for any orthogonal projector Π.

We show the following about the state $|0\rangle$, which is defined as above:

1) When $f(x) = 1$, $|0\rangle$ is "close" to the intersection of the subspace $\text{Img}(\Delta) \cap \text{Img}(\Pi_x)$. This implies that $|0\rangle$ is close to a 1-eigenvector of $R(\Pi_x) \cdot R(\Delta)$ (. the eigenphase of this eigenvector is 0)

2) When $f(x) = 0$, the projection of $|0\rangle$ onto the eigenspaces of $R(\Pi_x) \cdot R(\Delta)$ with eigenphase $\leq \frac{1}{100A}$ (in magnitude) is "small".

Therefore, if we perform phase estimation on the state $|0\rangle$ with operator $R(\Pi_x)\,R(\Delta)$, the phase estimate will be 0 with "high" probability when $f(x) = 1$, and will be $> \frac{1}{100A}$ with high probability when $f(x) = 0$. The number of repetitions of $R(\Pi_x)\,R(\Delta)$ required to distinguish between phases 0 & those $> \frac{1}{100A}$ in magnitude is $O(A)$.

Note also that the operator $R(\Delta)$ is independent of the input and that Π_x may be implemented with 1 query.

Proof of (1) above: Note that since $f(x)=1$, the following vector belongs to $\mathrm{Img}(\Delta)$:

$$|v_x\rangle = |0\rangle + \frac{1}{10\sqrt{A}} \sum_i |i\rangle\langle u_{xi}|x_i\rangle .$$

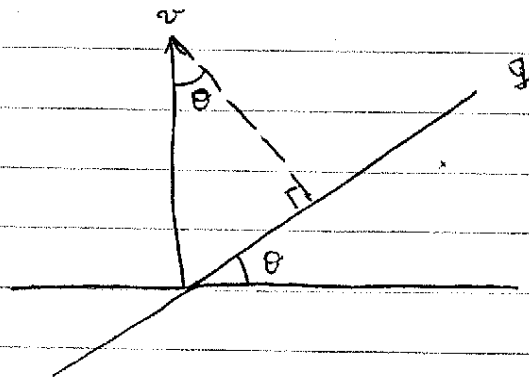
Its norm squared is $1 + \frac{1}{100A}\|u_x\|^2 \le 1 + \frac{1}{100} \approx 1$.

Therefore, the inner product of $|0\rangle$ with the normalized vector $|v_x\rangle$ is ≈ 1. Moreover, $|v_x\rangle$ is in $\mathrm{Img}(\Pi_x)$. ∎

For the proof of (2) above, we need the following property of a product of reflections, which is a consequence of an 1875 lemma due to Jordan:

Proposition Let v be perpendicular to $\mathrm{Img}(\Delta)$. Then, if P_θ denotes the orthogonal projection onto the subspace corresponding to eigenphases $\le \theta$ of the operator $R(\Pi_x)R(\Delta)$, we have:

$$\| P_\theta(\Pi_x v) \| \le \theta \cdot \|v\| .$$



$$\frac{\|\Pi_x v\|}{\|v\|} = \sin\theta \le \theta .$$

The eigenspaces of $R(\Pi_x)\cdot R(\Delta)$ are 1 or 2-dimensional, and the corresponding eigenphases are 0 or π, or θ, where θ is the angle between the restriction of the projectors onto the eigenspace. The above figure shows v as belonging

to one of these 2-dim subspaces with eigenphase $= \theta$ so that $\|P_\theta \Pi_x v\| = \|\Pi_x v\| \leq \theta \cdot \|v\|$

Proof of (2)

We have $x : f(x) = 0$.

Consider the vector $|0\rangle$:

Since $\Pi_x \sum_j |j\rangle|u_{xj}\rangle|\bar{x}_j\rangle = 0$, & $\Pi_x|0\rangle = |0\rangle$,

$\bar{x}_j$ ↳ bits of x negated.

We have:

$$|0\rangle = \Pi_x \Big(\overbrace{|0\rangle - 10\sqrt{A} \sum_j |j\rangle|u_{xj}\rangle|\bar{x}_j\rangle}^{|v\rangle} \Big) = \Pi_x |v\rangle$$

& $\langle v| \left(|0\rangle + \frac{1}{10\sqrt{A}} \sum_i |i\rangle|u_{yi}\rangle|y_i\rangle \right)$, for any $y \in f^{-1}(1)$

equals $\langle 0|0\rangle - \sum_{j : x_j \neq y_j} \langle u_{xj}, u_{yj}\rangle$

$= 1 - 1 = 0$ (by the property of the vector system, for $x, y : f(x) \neq f(y)$).

$\therefore$ $|0\rangle \in \text{Img}(\Delta)^\perp$, and $|0\rangle = \Pi_x|0\rangle$

By the proposition above $\|P_0|0\rangle\| = \|\Pi_x v\|$

$\leq \theta \cdot \|v\|$.

$\|v\|^2 = 1 + .100A^2$, so that choosing $\theta = \frac{1}{100A}$,

we get $\|P_0|0\rangle\| \leq \frac{1}{10}\left(1 + \frac{1}{100A^2}\right)^{1/2} \leq \frac{1}{4}$, as $A \geq 1$

for any function that is not constant. (The constraint implies $1 = \sum_{j : x_j \neq y_j} \langle u_{xj}, u_{yj}\rangle \leq \|u_x\| \cdot \|u_y\|$, so at least one of them is ≥ 1.)

This completes the proof.