Quantum Computation **(CMU 18-859BB, Fall 2015)**

Lecture 15: Reichardt's Theorem II: Evaluation of Span Programs

October 28, 2015

Lecturer: Ryan O'Donnell *Scribe: Vikesh Siddhu*

1 Short Summary

We have discussed in the past, general methods to lower bound the quantum query complexity. Now we discuss a way to convert the lower bounds given by the general adversary method [HLS07] into an upper bound [Rei09]. In this lecture we will cover the proof that leads to this result. Two key ingredients are Span Programs and how one defines their complexity.

2 Recall Span Programs

Let $w \in \{0,1\}^N$ be a string and $F : \{0,1\}^N \mapsto \{0,1\}$ be a function. A span program computes P for given w. Let $\{|v_i\rangle\}_{i=1}^m$ be vectors in $\mathbb{R}^d$ that are columns of a matrix V and let $|\tau\rangle \in \mathbb{R}^d$ be another vector called *target*. V is split into $2N$ blocks, the $2k^{th}$ and $2k^{th}+1$ block each have vectors corresponding to $w_k = 0$ and $w_k = 1$ resp.

Given w the span program makes available some N blocks, call the these set of vectors in the block avail(w) and the rest unavail(w). For example $w_1 = 0, w_2 = 1, \ldots, w_N = 1$ makes available the blocks $1, 4, \ldots 2N$ and the rest $2, 3, \ldots 2N-1$ become unavailable. Given a span program the function $P(w) = 1$ iff $|\tau\rangle \in \text{span}\{|v_i\rangle : |v_i\rangle \in \text{avail}(w)\}$

Suppose P computes $F : \mathcal{D}^{\subseteq\{0,1\}^N} \mapsto \{0,1\}$, then

- For $y \in F^{-1}(1)$, a *positive witness* is $|\alpha\rangle \in \mathbb{R}^m$ s.t.

$$\alpha_i = 0 \quad \forall i \in \text{unavail}(y) \tag{1}$$
$$V|\alpha\rangle = |\tau\rangle \tag{2}$$

  we define it's *size* to be $||\,|\alpha\rangle\,||^2$.

- For $z \in F^{-1}(0)$, a *negative witness* is $\langle\beta| \in \mathbb{R}^d$ s.t.

$$\langle\beta\,|\,v_i\rangle = 0 \quad \forall i \in \text{avail}(y) \tag{3}$$
$$\langle\beta\,|\,\tau\rangle = 1 \tag{4}$$

  we define it's *size* to be $||\,\langle\beta|\,V||^2$.

An *extended span program* is a span program along with positive and negative witnesses for all possible inputs w. We can define the complexity of an extended span program as follows

- Let $|\alpha_y\rangle$ be a positive witness for $y \in F^{-1}(1)$ then the YES complexity is defined as $T_1 \equiv \max_{y \in F^{-1}(1)} \{\text{size}(|\alpha_y\rangle)\}$.

- Let $|\beta_y\rangle$ be a negative witness for $y \in F^{-1}(0)$ then the NO complexity is defined as $T_0 \equiv \max_{y \in F^{-1}(0)} \{\text{size}(|\beta_y\rangle)\}$

- The overall complexity of the span program is $T = \sqrt{T_0 T_1}$

3 Reichardt's Theorem II

Theorem 3.1. *If a span program P with complexity T computes F, then there exists a quantum query algorithm for F making $O(T)$ queries of the oracle $O_f^{\pm}$.*

Fact 3.2. *The complexity T for the span program P to compute the function F is equal to the adversary lower bound $Adv^{\pm}(F)$*

Example 3.3. *Let $F = OR_N$ i.e. OR on N bits. Define a span program with vectors such that, for $w_i = 1$, the block in V has one vector $|v_i\rangle = [1]$ and the for $w_i = 0$, the block in V a null vector $|v_i\rangle = \phi$. Then*

1. *YES complexity $T_1 = 1$*
2. *NO complexity $T_0 = N$*
3. *The overall complexity is $T = \sqrt{T_0 T_1} = \sqrt{N}$*

We now come to the proof of Theorem 3.1.

Proof. Let $|\tilde{\tau}\rangle = \frac{|\tau\rangle}{c\sqrt{T_1}}$ and define $\tilde{V} \in \mathbb{R}^{d \times m+1}$ as

$$\tilde{V} = \left[\ |\tilde{\tau}\rangle \ \middle|\ V\ \right] \tag{5}$$

For the Grover case $\tilde{V} = [\frac{1}{c}\ 1\ 1\ \ldots 1]$.

For now the algorithm will work in the $\mathbb{R}^{m+1}$ space and any intermediate state is given by a vector $|s\rangle = \sum_{i=0}^{m} \alpha_i |i\rangle$ where each $\langle i \,|\, j\rangle = \delta_{ij}$. Define

$$K = \ker(\tilde{V}) = \{|u\rangle \in \mathbb{R}^{m+1} \mid \tilde{V}\,|u\rangle = 0\} \tag{6}$$

For the Grover case K consists of all vectors of mean 0.

Define R_K to be the reflection through K. Then R_K is a unitary operator on reals, i.e. an orthogonal matrix. For the Grover case R_K flips a vector in $\mathbb{R}^{m+1}$ across its mean.

Given $w \in \{0,1\}^N$, let

$$A_w = \text{span}\{|i\rangle\ \ 0 \le i \le m\ i \in \text{avail}(w)\} \tag{7}$$

by definition $|\tilde{\tau}\rangle \in A_w$ and is always available. Let R_{A_w} be the reflection through A_w, which mean we negate all the entries of a vector in $\mathbb{R}^{m+1}$ that are at the unavailable coordinates, so $R_{A_w} = -O_w^{\pm}$.

Let $U = R_{A_w} R_K$, computing R_K is a 0 query step and computing R_{A_w} takes 1 query (well 2 query if you un-compute the garbage).

We now describe a fake algorithm that to give some intuition behind how computes F on w using $O(T)$ queries.

1. Initialize the state $|\psi\rangle = |0\rangle$
2. For $t = 1, 2, \ldots CT$ apply U to $|\psi\rangle$
3. Measure $|\psi\rangle$ in standard basis, output 1 iff you observe $|0\rangle$

The basic idea is that

(i) If w is a YES instance, then U fixes $|0\rangle$

(ii) If w is a NO instance, then $U^{CT} |\psi\rangle$ is far from $|0\rangle$

The **first idea** can also be stated as, if $y \in F^{-1}(1)$ then $|0\rangle$ is 99% in K and A_y, hence U fixes 99% of $|0\rangle$.

Fact 3.4. *The accurate fact is $\exists$ $|\eta\rangle$ of length $\leq .01$ s.t. $|0\rangle - |\eta\rangle$ is an eigen vector of U.*

Proof. Let $|\alpha_y\rangle$ be a YES instance, let $|\eta\rangle = \frac{|\alpha_y\rangle}{c\sqrt{T_1}}$, we know $||\,|\alpha_y\rangle\,||^2 \leq T_1$ which implies, for $c \geq 100$

$$\sqrt{\langle \eta \,|\, \eta\rangle} \leq \frac{1}{c} \leq .01 \tag{8}$$

$U = R_{A_y} R_K$ where R_K is the reflection through $K = \ker(\tilde{V})$ and R_{A_y} is the reflection through A_w. Notice

1. $(|0\rangle - |\eta\rangle)$ is in the kernel of $\tilde{V}$ so $R_K(|0\rangle - |\eta\rangle) = (|0\rangle - |\eta\rangle)$

$$\tilde{V}(|0\rangle - |\eta\rangle) = |\tilde{\tau}\rangle - \frac{1}{c\sqrt{T_1}} \tilde{V} |\alpha_y\rangle \tag{9}$$

$$= |\tilde{\tau}\rangle - \frac{1}{c\sqrt{T_1}} |\tau\rangle \tag{10}$$

$$= |\tilde{\tau}\rangle - |\tilde{\tau}\rangle \tag{11}$$

$$= 0 \tag{12}$$

2. $(|0\rangle - |\eta\rangle)$ is in A_y because by definition $|0\rangle \in A_y$ and $|\eta\rangle \propto |\alpha_y\rangle$ and $|\alpha_y\rangle$ is in A_y, so $R_{A_y}(|0\rangle - |\eta\rangle) = (|0\rangle - |\eta\rangle)$

Hence U fixes $|0\rangle - |\eta\rangle$ $\square$

The **second idea** states, if $z \in F^{-1}(0)$ then $|0\rangle$ is far from states fixed by U.

Fact 3.5. *If $w \in F^{-1}(0)$ then $\exists$ $|u\rangle$ s.t. $Proj_{A_w}(|u\rangle) = |0\rangle$ and $||\,|u\rangle\,|| \leq 2cT$ and $|u\rangle \perp K$.*

Proof. Let $\langle\beta_w|$ be a NO witness for w, define

$$\langle u| \equiv c\sqrt{T_1}\,\langle\beta_w|\,\tilde{V} \tag{13}$$

Clearly $\tilde{V}\,|u\rangle \neq 0$, hence $|u\rangle \perp \ker(\tilde{V}) \implies |u\rangle \perp K$. Rewrite $|u\rangle$ as follows

$$\langle u| = c\sqrt{T_1}\{\langle 0|\,\langle\beta_w\,|\,\tilde{\tau}\rangle + \langle\beta_w|\,V\} \tag{14}$$

$$= \langle 0| + c\sqrt{T_1}\,\langle\beta_w|\,V \tag{15}$$

where the second equality follows from eq. (4) which states $\langle\tau\,|\,\beta_w\rangle = 1$ and $|\tilde{\tau}\rangle = \frac{|\tau\rangle}{c\sqrt{T_1}}$. Notice

(a) $\langle\beta_w|\,V$, the second term in the rhs of eq. (15) is a linear combination of unavailable vectors (since $|\beta_w\rangle$ is orthogonal to all available vectors)

(b) $||\,\langle\beta_w|\,V||^2 \leq T_0$ (since *size* of $\langle\beta_w|$ is at most T_0)

Lets switch back to kets $|u\rangle = [\langle u|]^\dagger$ where † is the conjugate-transpose (since everything is real here, it is just the transpose). Using (a) we conclude $\text{Proj}_{A_w}\,|u\rangle = |0\rangle$, using (b) we conclude

$$||\,|u\rangle\,|| \leq \sqrt{1 + c^2 T_0 T_1} \leq 1 + c\sqrt{T_0 T_1} \leq 2cT \tag{16}$$

□

Another key idea is the Kitaev Phase Estimation, which we shall delve into a little later. Before going further we review a few facts about orthogonal matrices from the 19^{th} century. Let $U \in \mathbb{R}^{m\times m}$ then

- U offers a decomposition of $\mathbb{R}^m$ s.t

$$R_m = \underbrace{H_1 \oplus H_2 \ldots \quad \ldots}_{\text{1 dim spaces where } U \text{ is } \mathbb{I}} \quad \underbrace{H_k \oplus H_{k+1} \ldots \quad \ldots}_{\text{1 dim spaces where } U \text{ is } -\mathbb{I}} \quad \underbrace{H_r \oplus H_{r+1} \oplus \ldots}_{\text{2 dim spaces where } U = R(\theta)} \tag{17}$$

  where $R(\theta)$ is a $2-D$ rotation by $\theta \in (-\pi, \pi]$. In other words, there are eigen spaces of U with eigen value $+1$ (the identity spaces), -1 (the reflection space) and $e^{i\theta}$ (the 2-d rotation space)

- Let A, B be subspaces of $\mathbb{R}^m$ and R_A, R_B be reflection through these spaces, construct $U = R_A R_B$. Let H be a 2-d θ rotation subspaces of U, then it is true, that $H \cap A$ and $H \cap B$ are 1 dimensional subspaces of $\mathbb{R}^m$ and the angle between $H \cap A$ and $H \cap B$ is $\theta/2$

  Lemma 3.6. *Suppose $|u\rangle \in H$, $u \perp (H \cap B)$ then $||proj_{A\cap H}(|u\rangle)|| \leq \frac{|\theta|}{2}||\,|u\rangle\,||$*

  Proof. Using Figure (1) we see that $||\text{proj}_{A\cap H}(|u\rangle)|| = \sin\frac{\theta}{2}||\,|u\rangle\,|| \leq \frac{\theta}{2}||\,|u\rangle\,||$ □

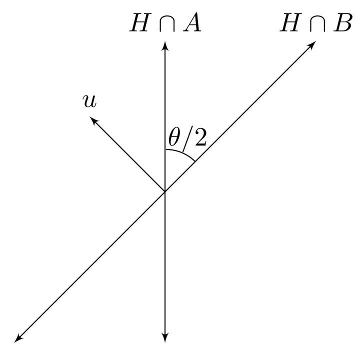

Figure 1: Intersecting subspaces $H \cap A$, $H \cap B$

Corollary 3.7. *Let P_δ be the projection onto all 2-d rotation subspaces of U with angle $\theta \leq \delta$, then*

$$||P_\delta[Proj_A(|u\rangle)]|| \leq \frac{\delta}{2}||\,|u\rangle\,|| \tag{18}$$

Proof. Apply lemma (3.6) subspace by subspace to $\text{Proj}_A(|u\rangle) = |v\rangle$ where it is given that $|u\rangle \perp B$. □

We now make the **second idea** precise. If $w \in F^{-1}(0)$ then $|0\rangle$ is far from states fixed by U. Recall $U = R_{A_w}R_K$ and $w \in F^{-1}(0) \implies \text{Proj}_A(|u\rangle) = |0\rangle$. Since $|0\rangle \perp K$ we use Corollary 3.7 and write

$$||P_\delta\,|0\rangle\,|| \leq \frac{\delta}{2}||\,|u\rangle\,|| \leq \delta c T \tag{19}$$

where the final inequality follows from eq. (16). By setting $\delta = \frac{1}{CT}$ and $\frac{c}{C} \leq 100$ we get

$$||P_\delta\,|0\rangle\,|| \leq .01 \tag{20}$$

In essence we have shown

- When $w \in F^{-1}(0)$ then $||P_\delta\,|0\rangle\,|| \leq .01$
- When $w \in F^{-1}(1)$ then $||P_0\,|0\rangle\,|| \geq .99$, where P_0 is a projection onto the +1 eigen space of U.

In order to distinguish whether $w \in F^{-1}(0)$ or $w \in F^{-1}(1)$, we must be able to tell whether $|0\rangle$ is 99% in U's rotation 0 eigen space or $|0\rangle$ is only $\leq 1\%$ in U's rotation $\leq \delta$ subspace. This can be achieved by Kitaev's phase estimation algorithm. □

4 Phase Estimation

Phase Detection is actually a special case of a more general algorithm called **Phase Estimation**, due to Kitaev[Kit97]. Here it the theorem:

Theorem 4.1. *Let U be an unitary operation on $\mathbb{R}^M$, given to a quantum algorithm as a "black box". Let $|\psi\rangle$ be an eigenvector of U, also given (in a sense) as a "black box". Say the eigenvalue of $|\psi\rangle$ is $e^{i\theta}$, where $\theta \in (-\pi, \pi]$ Then with only $O(1/\delta)$ applications of U, it is possible to distinguish the case $\theta = 0$ from $\theta \geq \delta$ with high probability.*

Let's be a bit more precise. Our algorithm (quantum circuit) will work with two registers; an M-dimensional register, and a "workspace" register of dimension $\Theta(1/\delta)$. (You can think of the workspace register as roughly $\log(1/\delta)$ additional qubits.) The circuit is allowed to use U gates on the first register, although it doesn't "know" what U is. (Actually, it will use controlled-U gates; there is a basic quantum circuit theorem, which we skipped, showing that one can construct controlled-U gates from U gates.) It is also assumed that the input to the circuit will be $|\psi\rangle \otimes |0\rangle$, where again, $|\psi\rangle$ is some ("unknown") eigenvector of U with eigenvalue $e^{i\theta}$. Then the Phase Detection circuit has the following properties:

- it contains at most $O(1/\delta)$ controlled-U operations;
- if $\theta = 0$, i.e. $|\psi\rangle$ is fixed by U, then the final state of the circuit will always be exactly $|\psi\rangle \otimes |0\rangle$, the same as the initial state;
- it $\theta \geq \delta$, then the final state will be of the form $|\psi\rangle \otimes |\phi\rangle$, where $|\langle \phi \,|\, 0 \rangle| \leq 1/4$

Then, in a typical use of Phase Detection, you just measure at the end, and look at the second (workspace) register. If $\theta = 0$ then you will see $|0\rangle$ with probability 1, and if $\theta \geq \delta$ you will see $|0\rangle$ with probability at most $1/4$.

Now actually, it's not 100% immediate to finish Reichardt's theorem with Phase Detection, because the summary that we ended suggested applying it with $|\psi\rangle$ equal to this "$|0\rangle$" vector, and $|0\rangle$ wasn't necessarily an eigenvalue of U, even in the YES case (in that case, it was only 99% equal to a 1-eigenvalue of U). Still, we're *almost* finished; I leave it as an exercise for the reader to complete the proof of Reichardt's theorem using the two facts we ended the summary one, plus the Phase Detection algorithm. (Hint: every input to Phase Detection can be written as a linear combination of U's orthogonal eigenvalues; so apply Phase Detection's guarantee to each, and use the fact that the Phase Detection algorithm is, like every quantum algorithm, a unitary linear transformation.)

We now give the proof of the Phase Detection theorem.

Proof. Let D be $8/\delta$ rounded up to the nearest integer power of 2, so $D = O(1/\delta)$. The workspace register will consist of exactly $d = \log 2D$ qubits, thought of as encoding an integer between 0 and $D1$. Now here is the algorithm:

- UniformSuperposition(workspace register) (this is just d little Hadamard gates, one on each workspace wire).

- for $t = 1, \ldots, D-1$
  do "controlled-U" on the first register, where the control condition is that the integer in the second register is at least t.

- UniformSuperposition^{-1}(workspace register).

That's it. You see it's indeed $O(1/\delta)$ applications of U. Let us now track the state of the registers throughout the algorithm.

(a) Initial state: $|\psi\rangle \otimes |0\rangle$

(b) After step 1:

$$|\psi\rangle \otimes \left(\frac{1}{\sqrt{D}} \sum_{j=0}^{D-1} |j\rangle \right) = \frac{1}{\sqrt{D}} \sum_{j=0}^{D-1} |\psi\rangle \otimes |j\rangle$$

(c) After step 2:

$$\frac{1}{\sqrt{D}} \sum_{j=0}^{D-1} U^{j+1} |\psi\rangle \otimes |j\rangle = \frac{1}{\sqrt{D}} \sum_{j=0}^{D-1} e^{i\theta(j+1)} |\psi\rangle \otimes |j\rangle \text{ since } |\psi\rangle \text{ is an } e^{i\theta} \text{ eigen vector of } U$$
$$= |\psi\rangle \otimes |\phi\rangle$$

where $|\phi\rangle = \frac{1}{\sqrt{D}} \sum_{j=0}^{D-1} e^{i\theta(j+1)} |j\rangle$ and the final eq

Let us now consider the two cases we need to analyze for the theorem.

- **Case 1**: $\theta = 0$ In this case we simply have $|\phi\rangle = \frac{1}{\sqrt{D}} \sum_{j=0}^{D-1} |j\rangle$ i.e., $|\phi\rangle$ is the uniform superposition in the second register. hus after step 3, it will turn back into $|0\rangle$. Hence the final state is indeed $|\psi\rangle \otimes |0\rangle$

- **Case 2**: $|\theta| \geq \delta$ since UniformSuperposition is a unitary transformation, it (and its inverse) preserve angles. It follows that the exact statement we must show is that $\langle \phi \,|\, \text{uniform superposition} \rangle^2 \leq 1/4$. The unsquared quantity on the left (which we must bound by $1/2$) is

$$\left| \left(\frac{1}{\sqrt{D}} \sum_{j=0}^{D-1} e^{-i\theta(j+1)} \langle j| \right) \left(\frac{1}{\sqrt{D}} \sum_{j=0}^{D-1} |j\rangle \right) \right| = \left| \frac{1}{D} \sum_{j=0}^{D-1} e^{-i\theta(j+1)} \right|$$

  You should be able to see how this will work out; we have a unit complex number with angle $-\theta$ where $|\theta| \geq \delta$. We're averaging it over D rotations of itself, where $D \gg \frac{1}{\delta}$. It should come out close to 0. To be completely precise, the above quantity is exactly (by the formula for the sum of a geometric series)

$$\frac{1}{D} \frac{|1 - e^{-i\theta D}|}{|1 - e^{-i\theta}|}$$

We have $\frac{1}{D} \leq \frac{\delta}{8}$, the numerator above is trivially at most 2, and the denominator is at least $|\theta|/2$ (simple trig), which is at least $\delta/2$. So the above expression is indeed at most $1/2$, as desired

□

References

[HLS07] Peter Hoyer, Troy Lee, and Robert Spalek. Negative weights make adversaries stronger. In *Proceedings of the Thirty-ninth Annual ACM Symposium on Theory of Computing*, STOC '07, pages 526–535, New York, NY, USA, 2007. ACM.

[Kit97] A Yu Kitaev. Quantum computations: algorithms and error correction. *Russian Mathematical Surveys*, 52(6):1191, 1997.

[Rei09] B. W. Reichardt. Span programs and quantum query complexity: The general adversary bound is nearly tight for every boolean function. In *Foundations of Computer Science, 2009. FOCS '09. 50th Annual IEEE Symposium*, pages 544–551, October 2009.