Quantum algorithms (CO 781/CS 867/QIC 823, Winter 2013)
Andrew Childs, University of Waterloo

LECTURE 15: The adversary method

We now discuss a second approach to proving quantum query lower bounds, the *quantum adversary method*. In fact, we'll see later that the generalized version of the adversary method we consider here (allowing negative weights) turns out to be an *upper bound* on quantum query complexity, up to constant factors.

Quantum adversaries

Motivation for the quantum adversary method comes from the following construction. Suppose the oracle is operated by an adversarial party who holds a quantum state determining the oracle string, which is in some superposition $\sum_{x\in S} a_x|x\rangle$ over the possible oracles. To implement each query, the adversary performs the "super-oracle"

$$O := \sum_{x\in S} |x\rangle\langle x| \otimes O_x. \tag{1}$$

An algorithm does not have direct access to the oracle string, and hence can only perform unitary operations that act as the identity on the adversary's superposition. After t steps, an algorithm maps the overall state to

$$\begin{aligned} |\psi^t\rangle &:= (I\otimes U_t)O\dots(I\otimes U_2)O(I\otimes U_1)O\left(\sum_{x\in S} a_x|x\rangle\otimes|\psi\rangle\right) & (2)\\ &= \sum_{x\in S} a_x|x\rangle\otimes|\psi_x^t\rangle. & (3)\end{aligned}$$

The main idea of the approach is that for the algorithm to learn x, this state must become very entangled. To measure the entanglement of the pure state $|\psi^t\rangle$, we can consider the reduced density matrix of the oracle,

$$\rho^t := \sum_{x,y\in S} a_x^* a_y \langle\psi_x^t|\psi_y^t\rangle\, |x\rangle\langle y|. \tag{4}$$

Initially, the state ρ^0 is pure. Our goal is to quantify how mixed it must become (i.e., how entangled the overall state must be) before we can compute f with error at most ϵ. To do this we could consider, for example, the entropy of ρ^t. However, it turns out that other measures are easier to deal with.

In particular, we have the following basic fact about the distinguishability of quantum states (for a proof, see for example section A.9 of KLM):

Fact. *Given one of two pure states $|\psi\rangle, |\phi\rangle$, we can make a measurement that determines which state we have with error probability at most $\epsilon\in[0,1/2]$ if and only if $|\langle\psi|\phi\rangle| \le 2\sqrt{\epsilon(1-\epsilon)}$.*

Thus it is convenient to consider measures that are linear in the inner products $\langle\psi_x^t|\psi_y^t\rangle$.

The adversary method

To obtain an adversary lower bound, we choose a matrix $\Gamma \in \mathbb{R}^{|S|\times|S|}$ with rows and columns indexed by the possible black-box inputs. The entry $\Gamma_{x,y}$ is meant to characterize how hard it is to distinguish between x and y. We say Γ is an *adversary matrix* if

1. $\Gamma_{xy} = \Gamma_{yx}$ and
2. if $f(x) = f(y)$ then $\Gamma_{xy} = 0$.

The second condition reflects that we do not need to distinguish between x and y if $f(x) = f(y)$.

The original adversary method made the additional assumption that $\Gamma_{xy} \geq 0$, but it turns out that this condition is not actually necessary. Sometimes we refer to the *negative* or *generalized* adversary method to distinguish it from the original, positive-weighted method. While it may not be intuitively obvious what it would mean to give a negative weight to the entry characterizing distinguishability of two inputs, it turns out that this flexibility can lead to significantly improved lower bounds for some functions.

Given an adversary matrix Γ, we can define a weight function

$$W^j := \sum_{x,y\in S} \Gamma_{xy} a_x^* a_y \langle \psi_x^j | \psi_y^j \rangle. \tag{5}$$

Note that this is a simple function of the entries of ρ^j. The idea of the lower bound is to show that W^j starts out large, must become small in order to compute f, and cannot change by much if we make a query.

The initial value of the weight function is

$$\begin{aligned} W^0 &= \sum_{x,y\in S} \Gamma_{xy} a_x^* a_y \langle \psi_x^0 | \psi_y^0 \rangle & (6) \\ &= \sum_{x,y\in S} a_x^* \Gamma_{xy} a_y & (7) \end{aligned}$$

since $|\psi_x^0\rangle$ cannot depend on x. To make this as large as possible, we take a to be a principal eigenvector of Γ, an eigenvector with eigenvalue $\pm\|\Gamma\|$. Then $|W^0| = \|\Gamma\|$.

The final value of the weight function is easier to bound if we assume a nonnegative adversary matrix. The final value is constrained by the fact that we must distinguish x from y with error probability at most ϵ whenever $f(x) \neq f(y)$. For this to hold after t queries, we need $|\langle \psi_x^t | \psi_y^t \rangle| \leq 2\sqrt{\epsilon(1-\epsilon)}$ for all pairs $x, y \in S$ with $f(x) \neq f(y)$ (by the above Fact). Thus we have

$$\begin{aligned} |W^t| &\leq \sum_{x,y\in S} \Gamma_{xy} a_x^* a_y 2\sqrt{\epsilon(1-\epsilon)} & (8) \\ &= 2\sqrt{\epsilon(1-\epsilon)}\|\Gamma\|. & (9) \end{aligned}$$

Here we can include the terms where $f(x) = f(y)$ in the sum since $\Gamma_{xy} = 0$ for such pairs. We also used the fact that the principal eigenevector of a nonnegative matrix can be taken to have nonnegative entries (by the Perron-Frobenius theorem).

A similar bound holds if Γ has negative entries, but we need a different argument. In general, one can only show that $|W^t| \leq (2\sqrt{\epsilon(1-\epsilon)} + 2\epsilon)\|\Gamma\|$. But if we assume that $f\colon S \to \{0,1\}$ has Boolean output, then we can prove the same bound as in the non-negative case, and the proof is simpler than for a general output space. We use the following simple result, stated in terms of the Frobenius norm $\|X\|_F := \sum_{a,b} |X_{ab}|^2$:

Proposition. *For any* $X \in \mathbb{C}^{m\times n}, Y \in \mathbb{C}^{n\times n}, Z \in \mathbb{C}^{n\times m}$*, we have* $|\mathrm{tr}(XYZ)| \le \|X\|_F\|Y\|\|Z\|_F$.

Proof. We have

$$\mathrm{tr}(XYZ) = \sum_{a,b,c} X_{ab}Y_{bc}Z_{ca} \tag{10}$$

$$= \sum_a (x^a)^\dagger Y z^a \tag{11}$$

where $(x^a)_b = X^*_{ab}$ and $(z^a)_c = Z_{ca}$. Thus

$$|\mathrm{tr}(XYZ)| \le \sum_a \|x^a\|\|Yz^a\| \tag{12}$$

$$\le \|Y\| \sum_a \|x^a\|\|z^a\| \tag{13}$$

$$\le \|Y\| \sqrt{\sum_a \|x^a\|^2 \sum_{a'} \|z^{a'}\|^2} \tag{14}$$

$$= \|Y\|\|X\|_F\|Z\|_F \tag{15}$$

as claimed, where we used the Cauchy-Schwarz inequality in the second and third steps. □

To upper bound $|W^t|$ for the negative adversary with Boolean output, write $W^t = \mathrm{tr}(\Gamma V)$ where $V_{xy} := a^*_x a_y \langle\psi^t_x|\psi^t_y\rangle\delta[f(x) \ne f(y)]$. Define

$$C := \sum_{x\in S} a_x \Pi_{f(x)}|\psi^t_x\rangle\langle x| \tag{16}$$

$$\bar{C} := \sum_{x\in S} a_x \Pi_{1-f(x)}|\psi^t_x\rangle\langle x| \tag{17}$$

with Π_0, Π_1 denoting the projectors onto the subspaces indicating $f(x) = 0, 1$, respectively. Then

$$(C^\dagger\bar{C})_{xy} = a^*_x a_y \langle\psi^t_x|\Pi_{f(x)}\Pi_{1-f(y)}|\psi^t_y\rangle, \tag{18}$$

so

$$(C^\dagger\bar{C} + \bar{C}^\dagger C)_{xy} = a^*_x a_y \langle\psi^t_x|(\Pi_{f(x)}\Pi_{1-f(y)} + \Pi_{1-f(x)}\Pi_{f(y)})|\psi^t_y\rangle \tag{19}$$

$$= a^*_x a_y \langle\psi^t_x|\psi^t_y\rangle\delta[f(x) \ne f(y)], \tag{20}$$

i.e., $V = C^\dagger\bar{C} + \bar{C}^\dagger C$. Thus we have

$$W^t = \mathrm{tr}(\Gamma(C^\dagger\bar{C} + \bar{C}^\dagger C)) \tag{21}$$

$$= \mathrm{tr}(\bar{C}\Gamma C^\dagger) + \mathrm{tr}(C\Gamma\bar{C}^\dagger). \tag{22}$$

By the Proposition, $|W^t| \le 2\|\Gamma\|\|C\|_F\|\bar{C}\|_F$. Finally, we upper bound $\|C\|_F$ and $\|\bar{C}\|_F$. We have

$$\|C\|^2_F + \|\bar{C}\|^2_F = \sum_{x,y\in S} |a_x|^2(|\langle y|\Pi_{f(x)}|\psi^t_x\rangle|^2 + |\langle y|\Pi_{1-f(x)}|\psi^t_x\rangle|^2) = 1 \tag{23}$$

$$\|\bar{C}\|^2_F = \sum_{x\in S} |a_x|^2\|\Pi_{1-f(x)}|\psi^t_x\rangle\|^2 \le \epsilon. \tag{24}$$

Therefore $\|C\|_F\|\bar{C}\|_F \le \max_{x\in[0,\epsilon]} \sqrt{x(1-x)} = \sqrt{\epsilon(1-\epsilon)}$ (assuming $\epsilon \in [0,1/2]$), and we find that $|W^t| \le 2\sqrt{\epsilon(1-\epsilon)}\|\Gamma\|$, as claimed.

It remains to understand how much the weight function can decrease at each step of the algorithm. We have

$$W^{j+1} - W^j = \sum_{x,y\in S} \Gamma_{xy} a_x^* a_y (\langle\psi_x^{j+1}|\psi_y^{j+1}\rangle - \langle\psi_x^j|\psi_y^j\rangle). \tag{25}$$

Consider how the state changes when we make a query. We have $|\psi_x^{j+1}\rangle = U^{j+1}O_x|\psi_x^j\rangle$. Thus the elements of the Gram matrix of the states $\{|\psi_x^{j+1}\rangle : x \in S\}$ are

$$\begin{aligned} \langle\psi_x^{j+1}|\psi_y^{j+1}\rangle &= \langle\psi_x^j|O_x^\dagger(U^{j+1})^\dagger U^{j+1}O_y|\psi_y^j\rangle && (26)\\ &= \langle\psi_x^j|O_xO_y|\psi_y^j\rangle && (27) \end{aligned}$$

since U^{j+1} is unitary and $O_x^\dagger = O_x$. Therefore

$$W^{j+1} - W^j = \sum_{x,y\in S} \Gamma_{xy} a_x^* a_y \langle\psi_x^j|(O_xO_y - I)|\psi_y^j\rangle. \tag{28}$$

Observe that $O_xO_y|i,b\rangle = (-1)^{b(x_i\oplus y_i)}|i,b\rangle$. Let $P_0 = I \otimes |0\rangle\langle 0|$ denote the projection onto the $b=0$ states, and let P_i denote the projection $|i,1\rangle\langle i,1|$. (As with O_x, the projections P_i implicitly act as the identity on any ancilla registers, so $\sum_{i=0}^n P_i = I$.) Then $O_xO_y = P_0 + \sum_{i=1}^n (-1)^{x_i\oplus y_i}P_i$, so $O_xO_y - I = -2\sum_{i\colon x_i\ne y_i} P_i$. Thus we have

$$W^{j+1} - W^j = 2\sum_{x,y\in S}\sum_{i\colon x_i\ne y_i} \Gamma_{xy} a_x^* a_y \langle\psi_x^j|P_i|\psi_y^j\rangle. \tag{29}$$

Now for each $i \in \{1,\ldots,n\}$, let Γ_i be a matrix with

$$(\Gamma_i)_{xy} := \begin{cases} \Gamma_{xy} & \text{if } x_i \ne y_i \\ 0 & \text{if } x_i = y_i \end{cases} \tag{30}$$

Then we have

$$\begin{aligned} W^{j+1} - W^j &= 2\sum_{x,y\in S}\sum_{i=1}^n (\Gamma_i)_{xy} a_x^* a_y \langle\psi_x^j|P_i|\psi_y^j\rangle && (31)\\ &= 2\sum_{i=1}^n \operatorname{tr}(Q_i\Gamma_iQ_i^\dagger) && (32) \end{aligned}$$

where $Q_i := \sum_x a_x P_i|\psi_x^j\rangle\langle x|$.

Using the triangle inequality and the above Proposition, we have

$$\begin{aligned} |W^{j+1} - W^j| &\le 2\sum_{i=1}^n |\operatorname{tr}(Q_i\Gamma_iQ_i^\dagger)| && (33)\\ &\le 2\sum_{i=1}^n \|\Gamma_i\|\|Q_i\|_F^2. && (34) \end{aligned}$$

Since

$$\begin{aligned}
\sum_{i=1}^n \|Q_i\|_F^2 &= \sum_{i=1}^n \sum_{x\in S} |a_x|^2 \|P_i|\psi_x^j\rangle\|^2 && (35)\\
&\le \sum_{x\in S} |a_x|^2 && (36)\\
&= 1, && (37)
\end{aligned}$$

we have

$$|W^{j+1} - W^j| \le 2 \max_{i\in\{1,\ldots,n\}} \|\Gamma_i\|. \tag{38}$$

Combining these three facts gives the adversary lower bound. Since $|W^0| = \|\Gamma\|$, we have

$$|W^t| \ge \|\Gamma\| - 2t \max_{i\in\{1,\ldots,n\}} \|\Gamma_i\|. \tag{39}$$

Thus, to have $|W^t| \le 2\sqrt{\epsilon(1-\epsilon)}\|\Gamma\|$, we require

$$t \ge \frac{1-2\sqrt{\epsilon(1-\epsilon)}}{2}\,\mathrm{Adv}(f). \tag{40}$$

where

$$\mathrm{Adv}(f) := \max_\Gamma \frac{\|\Gamma\|}{\max_{i\in\{1,\ldots,n\}}\|\Gamma_i\|} \tag{41}$$

with the maximum taken over all adversary matrices Γ for the function f. (Often the notation $\mathrm{Adv}(f)$ is reserved for the maximization over nonnegative adversary matrices, with the notation $\mathrm{Adv}^{\pm}(f)$ for the generalized adversary method allowing negative weights.)

Example: Unstructured search

As a simple application of this method, we prove the optimality of Grover's algorithm. It suffices to consider the problem of distinguishing between the case of no marked items and the case of a unique marked item (in an unknown location). Thus, consider the partial function where S consists of the strings of Hamming weight 0 or 1, and f is the logical OR of the input bits. (Equivalently, we consider the total function OR but only consider adversary matrices with zero weight on strings of Hamming weight more than 1.)

For this problem, adversary matrices have the form

$$\Gamma = \begin{pmatrix} 0 & \gamma_1 & \cdots & \gamma_n \\ \gamma_1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_n & 0 & \cdots & 0 \end{pmatrix} \tag{42}$$

for some nonnegative coefficients $\gamma_1, \ldots, \gamma_n$. Symmetry suggests that we should take $\gamma_1 = \cdots = \gamma_n$. This can be formalized, but for the present purposes we can take this as an ansatz.

Setting $\gamma_1 = \cdots = \gamma_n = 1$ (since an overall scale factor does not affect the bound), we have

$$\Gamma^2 = \begin{pmatrix} n & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & \cdots & 1 \end{pmatrix} \tag{43}$$

which has norm $\|\Gamma^2\| = n$, and hence $\|\Gamma\| = \sqrt{n}$. We also have

$$\Gamma_1 = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \tag{44}$$

and similarly for the other Γ_i, so $\|\Gamma_i\| = 1$. Thus we find $\mathrm{Adv}(\textsc{or}) \geq \sqrt{n}$, and it follows that $Q_\epsilon(\textsc{or}) \geq \frac{1-2\sqrt{\epsilon(1-\epsilon)}}{2}\sqrt{n}$. This shows that Grover's algorithm is optimal up to a constant factor (recall that Grover's algorithm finds a unique marked item with probability $1 - o(1)$ in $\frac{\pi}{4}\sqrt{n} + o(1)$ queries).