Orthogonal transformations (Raz '99)

Suppose Alice has a unit vector $\boldsymbol{x} \in \mathbb{R}^n$, and two orthogonal subspaces $M_0, M_1 \leq \mathbb{R}^n$, and Bob has an orthogonal (i.e. real unitary) $n \times n$ matrix U. Let Π_0, Π_1 be the orthogonal projection operators onto M_0 and M_1. Alice and Bob want to determine whether

$$\left\|\Pi_0 U \boldsymbol{x}\right\|^2 \geq 1 - \varepsilon \qquad \text{or} \qquad \left\|\Pi_1 U \boldsymbol{x}\right\|^2 \geq 1 - \varepsilon,$$

given that one of them holds, for some small $\varepsilon > 0$. That is, given that $U\boldsymbol{x}$ is close either to the space M_0 or the space M_1, which is the case?

Convention: For the purposes of this problem, we assume that various the real numbers in this problem are approximated using $O(\log n)$ bits each. (The input then takes up $O(n^2 \log n)$ bits: $O(n \log n)$ for the representation of $\boldsymbol{x}$, and $O(n^2 \log n)$ for each of U, M_0, and M_1.)

The quantum strategy for solving this problem is fairly straightforward:

1. Alice performs the encoding

$$\boldsymbol{x} \longmapsto |\phi_{\boldsymbol{x}}\rangle = \sum_{j=1}^{n} x_j |j\rangle :$$

   that is, she represents the vector $\boldsymbol{x}$ by the quantum state similar to $\boldsymbol{x}$. The size of this representation requires n different basis states, and so requires $\log n$ qubits. Alice then transmits $|\phi_{\boldsymbol{x}}\rangle$ to Bob.

2. Bob applies the U to the state he recieves as a quantum operation, and transmits the result $U|\phi_{\boldsymbol{x}}\rangle$ back to Alice.

3. Alice performs a measurement with respect to M_0 and M_1 using $\{\Pi_0, \Pi_1\}$ as a partial measurement.

The state will collapse into the subspace similar to M_0 or M_1 when Alice performs the measurement. Then, she will get the correct result with high probability (greater than $1-\varepsilon$), with the result depending on which space $U|\phi_{\boldsymbol{x}}\rangle$ has the largest projection onto.

Altogether, this strategy requires $\lceil \log n \rceil$ qubits of communication. This is actually independent of the number of bits used to represent the real numbers involved in the input (although the size of the representation may affect the amount of work Alice and Bob may need to do to prepare, transform, and measure vectors precisely).

Claim 1 *Every classical protocol for this problem requires $\Omega(\sqrt{n})$ bits to solve this problem with probability greater than $\frac{1}{2}$ in the worst case.*

Claim 2 *This problem is* complete *for two-round quantum communication protocols, in the sense that any two-round quantum problem can be reduced to this problem.*

Sampling Disjoint Subsets

Suppose Alice and Bob wish to generate subsets $S, T \subset [n]$, such that $|S| = |T| = k$ for some $1 \le k \le n$ and $S \cap T = \varnothing$, in a uniformly random manner, so that Alice ends up with S and Bob with T. That is, they wish to sample from the probability distribution

$$p(S,T) \;=\; \left\{ \begin{array}{ccl} \frac{1}{\binom{n}{k}\binom{n-k}{k}} & , & \text{if } \;|S| = |T| = k\; \text{ and }\; S \cap T = \varnothing \\ 0 & , & \text{otherwise} \end{array} \right\},$$

or a distribution close to p in the ℓ_1 norm, with no inputs. (By a distribution f "close to p in the ℓ_1 norm", we mean that $\|f - p\|_1 \le \varepsilon$ should hold for some chosen constant $\varepsilon > 0$.)

Theorem 3 (Ambainis, Schulman, Ta-Shma, Vazirani, Widgerson '98) *Classically, for some $k = \Theta(\sqrt{n})$, this requires $\Omega(\sqrt{n})$ bits of communication to achieve.*

[ASTVW'98] showed that it requires only $O(\log n)$ qubits of quantum communication. The basic strategy is to construct a state $|\psi\rangle$ which is very close to

$$|\phi\rangle \;=\; \frac{1}{\sqrt{\binom{n}{k}\binom{n-k}{k}}} \sum_{\substack{S \cap T = \varnothing \\ |S|=|T|=k}} |S\rangle|T\rangle$$

with respect to the ℓ_2 norm; that is, such that $\Big\| \, |\phi\rangle - |\psi\rangle \, \Big\|_2 \le \varepsilon$, and then apply the following result:

Lemma 4 (Bernstein,Vazirani'93) *If $\Big\| \, |\phi\rangle - |\psi\rangle \, \Big\|_2 \le \delta$, and p, q are probability distributions obtains by measuring $|\phi\rangle$ and $|\psi\rangle$ with respect to some fixed measurement, then $\| \, p - q \, \|_1 \le 2\delta$.*

The proof of this Lemma is left to the reader.

Consider the Schmidt decomposition of $|\phi\rangle$: for some orthogonal bases $\{ \, |e_j\rangle \, \}_{j=0}^N$ of Alice's space and $\{ \, |f_j\rangle \, \}_{j=0}^N$ of Bob's space, we have

$$|\phi\rangle \;=\; \sum_{j=0}^{N} \sqrt{\lambda_j} |e_j\rangle |f_j\rangle \, .$$

Without loss of generality, we may order these bases so that the coefficients λ form a descending sequence (i.e. so that $\lambda_0 \ge \lambda_1 \ge \lambda_2 \ge \cdots$). We can then approximate $|\phi\rangle$ by truncating the Schmidt decomposition and renormalizing. For some constant $\nu > 0$, let $M le N$ be such that

$$A_M \;=\; \sum_{j=0}^{M} \lambda_j \;\ge\; 1 - \nu \, .$$

Then, we define

$$|\psi'\rangle \;=\; \sum_{j=0}^{M} \sqrt{\lambda_j} \, |e_j\rangle|f_j\rangle \qquad\qquad \text{(having norm } \sqrt{A_M} \ge \sqrt{1-\nu}\,)$$

$$|\psi\rangle \;=\; \frac{|\psi'\rangle}{\| \, |\psi'\rangle \, \|_2} \;=\; \frac{|\psi'\rangle}{\sqrt{A_M}} \qquad\qquad \text{(a unit vector, i.e. a quantum state vector, parallel to } |\psi'\rangle\,)$$

From the triangle inequality, we then have the following:

$$\begin{aligned} \Big\| \, |\psi\rangle - |\phi\rangle \, \Big\|_2 &\leq \Big\| \, |\psi\rangle - |\psi'\rangle \, \Big\|_2 + \Big\| \, |\psi'\rangle - |\phi\rangle \, \Big\|_2 \\ &\leq (1 - \sqrt{1-\nu}) + \sqrt{\nu} \\ &\leq 2\sqrt{\nu} \end{aligned}$$

We can describe a simple quantum strategy for Alice and Bob to create the state $|\psi\rangle$ in a distributed environment:

1. Alice prepares the state

$$|\Psi\rangle = \frac{\sum_{j=0}^{M} \sqrt{\lambda_j} \, |j\rangle|j\rangle}{\sqrt{A_M}}$$

   and sends the latter part to Bob.

2. Alice and Bob perform the unitary embeddings of basis U and V, where

$$\begin{aligned} U|j\rangle|\bar{0}\rangle &= |e_j\rangle \\ V|j\rangle|\bar{0}\rangle &= |f_j\rangle \end{aligned}$$

This allows Alice and Bob to prepare the state $|\psi\rangle$ in a distributed manner. If they then measure their states in the standard basis, they will obtain a probability distribution over pairs of subsets of $[n]$ which is different from p in the ℓ_1 norm by at most $8\nu^2$, by Lemma 4. The communication cost of this strategy is $\lceil \log M \rceil$ qubits; it remains only to investigate how small M can be made and still achieve a reasonable probability of success.

Fact: *There are exactly $\binom{n}{k}$ non-zero Schmidt coefficients λ_j. In particular, for each $0 \leq h \leq k$, the value*

$$\frac{\binom{n-k-h}{k-h}^2}{N}$$

occurs with multiplicity $\binom{n}{h} - \binom{n}{h-1}$, for $h \geq 1$, and the value $\binom{n-k}{k}^2/N$ occurs exactly once, among the Schmidt coefficients.

In our case, we are interested in values of $k \in \theta(\sqrt{n})$. Define the sequence $\{q_t\}_{t=1}^{k}$ as follows:

$$q_t = \frac{\binom{n-k-t}{k-t}^2 \left[\binom{n}{t} - \binom{n}{t-1} \right]}{N};$$

then, q_t is the sum of the coefficients of similar type described in the fact above. We can now apply a second fact:

Fact: $\frac{q_{t+1}}{q_t} \in O(\frac{1}{t+1})$.

This means that the sequence $\{q_t\}_{t=1}^{k}$ converges more quickly than a geometric series, so an infinite series built from it will be absolutely convergent. To approximate such an infinite series (or a partial sum of many terms) with constant error ν, it suffices to use a constant number of terms T.

Now, each term q_t is a sum of $\binom{n}{t} - \binom{n}{t-1}$ Schmidt coefficients. Then, in order to have enough terms for $A_M \geq 1 - \nu$, we require

$$M = \sum_{t=0}^{T} \left[\binom{n}{t} - \binom{n}{t-1} \right] = \binom{n}{T}$$

So, $M \in \mathrm{poly}(n)$, whence the communication cost of the strategy for sampling disjoint sets given above has a cost of $O(\log n)$.

Complexity of Set Disjointness (Razborov '03)

Continuing from last lecture, we will start the demonstration of the quantum lower bound of $\Omega(\sqrt{n})$ for the problem of set disjointness, which is the problem of deciding whether for two binary strings $\boldsymbol{x}, \boldsymbol{y} \in \{0,1\}^n$, there is any $j \in [n]$ such that $x_j y_j = 1$.

Step 1. Consider the matrix $P = p(\boldsymbol{x}, \boldsymbol{y})$, given by the probability p of a protocol accepting with m qubits of communication with inputs $\boldsymbol{x}, \boldsymbol{y} \in \{0,1\}^n$.

If this protocol is correct, then P will be close in the ℓ_∞ norm to the matrix M_f obtainable from the probability distribution f which perfectly decides DISJ_n. That is, the elements $p(\boldsymbol{x}, \boldsymbol{y})$ of the matrix P will be such that $|p(x,y) - f(x,y)| \leq \varepsilon$ for all $\boldsymbol{x}, \boldsymbol{y} \in \{0,1\}^n$, for some reasonably small constant $\varepsilon > 0$.

Claim 5 *rank*$(P) \leq 2^{2m}$.

Step 2. For any matrix Q approximating M_f to within $\frac{1}{3}$ in the ℓ_∞ norm (i.e. which differs element-wise from M_f by at most $\frac{1}{3}$), we have $\log(\text{rank}(Q)) \in \Omega(\sqrt{n})$.

We will explore these ideas more in the next lecture.