Topics in Quantum Information	**U. Waterloo CO 781, Spring 2004**
Ashwin Nayak (Instructor)	**Marcus Silva (Scribe)**
Lower bound for set disjointness	**May 18, 2004**

1 Razborov's Set Disjointness Lower Bound

There are two major steps in the proof:

Step 1: [Yao93] If $P(x, y)$ is the probability of acceptance on input x, y, then

$$\mathrm{rank}\,(P) \leq 2^{2m},$$

where m is the communication cost of the protocol. Note that this is not restricted to the set disjointness problem: it is applicable to any communication protocol to compute some function with communication cost m.

Step 2: [AR03]

For the disjointness function $f(x, y) = 1 \leftrightarrow x \cap y = \emptyset$, and any matrix Q such that

$$\|Q - M_f\|_\infty \;\; \leq \;\; \frac{1}{3},$$

where M_f is the communication matrix for the function f, we have

$$\mathrm{rank}\,(Q) \geq 2^{\Omega(\sqrt{n})}.$$

Since the probability matrix P approximates M_f in the ℓ_∞ norm, we conclude that $m \in \Omega(\sqrt{n})$.

Proof. **(Step 1)** We show that the probability of acceptance on the possible inputs form a matrix of "low" rank.

From Yao's lemma for quantum communication protocols, the final joint state of the two parties may be written as

$$\sum_{c \in \{0,1\}^m} |a_c(x)\rangle |b_c(y)\rangle.$$

Assume that the last bit of c is the output of the protocol. Then

$$\begin{aligned} P(x, y) &= \Pr(\text{protocol accepts on input } x, y) \\ &= \left\| \sum_{c' \in \{0,1\}^{m-1}} |a_{c'1}(x)\rangle |b_{c'1}(y)\rangle \right\|^2 \\ &= \sum_{d_1, d_2} \langle a_{d_1}(x) | a_{d_2}(x)\rangle \langle b_{d_1}(y) | b_{d_2}(y)\rangle. \end{aligned}$$

In the above expression, d_1, d_2 are strings of length m that end in 1. Define a $2^n \times 2^{2m}$ matrix A with entries defined as

$$[A]_{x, d_1 d_2} = \langle a_{d_1}(x) | a_{d_2}(x)\rangle,$$

and a $2^{2m} \times 2^n$ matrix B with entries

$$[B]_{d_1 d_2, y} = \langle b_{d_1}(y) | b_{d_2}(y)\rangle,$$

so that

$$P = AB.$$

Thus

$$\mathrm{rank}\,(P) \quad \leq \quad \min\left\{\mathrm{rank}\,(A)\,, \mathrm{rank}\,(B)\right\}.$$

However, for any $i \times j$ matrix M

$$\begin{aligned} \mathrm{rank}\,(M) &\leq i & (1)\\ \mathrm{rank}\,(M) &\leq j, & (2) \end{aligned}$$

so we have

$$\mathrm{rank}\,(P) \quad \leq \quad 2^{2m}.$$

This concludes the first step in the proof. □

Proof. **(Step 2)** We show that any matrix that approximates the communication matrix for the set dijointness function has rank at least $2^{\Omega(\sqrt{n})}$.

This step is a classic application of the *Polynomial Method*. The method consists of the following parts:

a) We construct a polynomial that approximates the logical NOR function $\mathrm{NOR}(z_1, \cdots, z_n) = \neg \bigvee_{i=1}^{n} z_i$. (Here, we think of z_i eas $x_i \wedge y_i$.) The polynomial is a function of the hamming weight of z and has degree $O(n)$.

b) The polynomial constructed above has too high a degree to be of use in our proof; we go through a degree reduction procedure and derive a polynomial of degree $O(m)$ that continues to approximates the NOR function.

c) Finally, we show that any such polynomial has high degree $(\Omega(\sqrt{n}))$.

We start with the construction of the polynomial (part a). Let Q be a matrix such that

$$||Q - M_f||_\infty \quad \leq \quad \frac{1}{3},$$

i.e., $Q(x, y) \in \left[\frac{2}{3}, \frac{4}{3}\right]$ for $x \cap y = \emptyset$ and $Q(x, y) \in \left[-\frac{1}{3}, \frac{1}{3}\right]$ for $|x \cap y| > 0$.

For technical reasons, we will restrict our attention to input subsets that are of size $k = n/8$. Let N_i be the number of such input pairs that have intersection size equal to i, and let $p(i)$ be the average acceptance probability on such inputs:

$$\begin{aligned} N_i &= \binom{n}{i}\binom{n-i}{k-i}\binom{n-k}{k-i} & (3)\\ p(i) &= \sum_{x,y:|x \cap y| = i} \frac{Q(x, y)}{N_i}. & (4) \end{aligned}$$

Define

$$J_i(x, y) \equiv J_{n,k,i}(x, y) \quad = \quad \begin{cases} \frac{1}{N_i} & \text{if } |x| = |y| = k,\, |x \cap y| = i \\ 0 & \text{otherwise,} \end{cases}$$

We can rewrite

$$p(i) \quad = \quad \mathrm{Tr}(QJ_i).$$

The matrices J_i have been studied extensively, and are very well understood. We will invoke the following properties of $\{J_i\}$:

i – $[J_i, J_l] = 0$, i.e., the matrices commute.

ii – The $\{J_i\}$ are symmetric and therefore are simultaneously diagonalizable (due to property **i**). Each matrix J_i has $k+1$ real eigenvalues corresponding to $k+1$ orthogonal eigenspaces $E_0, E_1, \cdots, E_k$. The space $E_\ell, \ell > 0$ has dimension $\binom{n}{\ell} - binomn\ell - 1$. E_0 is one dimensional.

iii – Let the eigenvalue of J_i corresponding to the eigenspace E_k be $\lambda_{i,t}$. Then, $\lambda_{i,t}$ is a polynomial of degree t in i.

iv – For $k = \frac{n}{8}$, we have

$$|\lambda_{i,t}| \;\leq\; \frac{1}{\binom{n}{k}} \mathrm{e}^{-\alpha t},$$

for some absolute constant $\alpha > 0$.

Let U be the unitary matrix that diagonalizes all J_i simultaneously, and let

$$\Lambda_i = U J_i U^\dagger,$$

so that

$$\Lambda_i \;=\; \mathrm{diag}(\lambda_{i,0}, \underbrace{\lambda_{i,1}, \ldots, \lambda_{i,1}}_{\binom{n}{1}-\binom{n}{0}}, \underbrace{\lambda_{i,2}, \ldots, \lambda_{i,2}}_{\binom{n}{2}-\binom{n}{1}}, \cdots).$$

Now,

$$\begin{aligned} p(i) &= \mathrm{Tr}(QJ_i) && (5)\\ &= \mathrm{Tr}(QU^\dagger \Lambda_i U) && (6)\\ &= \mathrm{Tr}((U^\dagger Q U)\Lambda_i) && (7)\\ &= \mathrm{Tr}(\tilde{Q}\Lambda_i). && (8) \end{aligned}$$

Therefore we can rewrite $p(i)$ as

$$p(i) \;=\; \sum_{t=0}^{k} a_t \lambda_{i,t},$$

where a_t is the sum of the diagonal entries of $\tilde{Q}$ corresponding to the eigenspace E_t. Along with property **iii** of the $\{J_i\}$ matrices, this shows that $p(i)$ is a polynomial in i of degree k, thus concluding part (a) of Step 2.

For the proof of part (b), the degree reduction step, we will need some basic facts from matrix algebra.

Not all matrices are unitarily diagonalizable. However, every matrix may be cast in a form reminiscent of such a diagonalization.

Theorem 1 (Singular value decomposition). *Any $n \times n$ complex matrix R may be written as*

$$R \;=\; \sum_{l=1}^{n} \sigma_l |u_l\rangle\langle v_l|,$$

where $\sigma_l \geq 0$, and $\{|u_l\rangle\}, \{|v_l\rangle\}$ are orthonormal bases for $\mathbb{C}^n$. The σ_l are called the singular values *of R.*

Do you notice the similarity with the Schmidt decomposition theorem? These two are in fact, equivalent.

The SVD theorem immediately implies

Corollary 2. *a)* rank $(()\ R)$ *equals the number of non-zero singular values of R.*

b) For all l, $\sigma_l(R) \;\leq\; \max_{|\alpha\rangle,|\beta\rangle} |\langle\alpha|R|\beta\rangle|$.

Definition 3. *The* trace norm *of a matrix R is given by*

$$\| R \|_{\mathrm{tr}} \quad = \quad \sum_{l=1}^{n} \sigma_l(R),$$

where $\sigma_l(R)$ are the singular values of R.

(We leave it to the reader to verify that this is indeed a norm.)

The following is another useful consequence of the SVD theorem.

Lemma 4. $\| R \|_{\mathrm{tr}} \quad = \quad \max_{\textit{unitary } U} |\mathrm{Tr}(UR)|$.

It follows that

Corollary 5. $\sum_i |R_{ii}| \quad \leq \quad \| R \|_{\mathrm{tr}}$.

We urge the reader to verify all these properties.

In the next lecture, we will see how these properties help us conclude Razborov's lower bound. □

References

[AR03] Alexander Razborov. Quantum Communication Complexity of Symmetric Predicates. *Izvestiya: Mathematics*, 67(1):145–159, 2003.

[Yao93] A.C.-C. Yao. Quantum circuit complexity. In *Proceedings of the 34th Annual Symposium on Foundations of Computer Science*, pages 352–361, 1993.