10/29/2003

Notes for Lecture 17

17 A Lower Bound for two query locally decodable codes

In today's lecture we will prove a lower bound for 2-query locally decodable codes. Consider the case of a locally decodable code, $C : \{0,1\}^k \to \{0,1\}^n$, which is not necessarily linear, with a non-adaptive decoding algorithms that uses two queries, is $\frac{1}{\delta}$-smooth, and outputs the value of any message bit with probability at least $\frac{1}{2} + \epsilon$ on all inputs with at most δn corrupt entries.

Last time we saw some quantum information theory, and using that, we will prove a $n = 2^{\Omega(k)}$ lower bound for that code (which is also polynomial in ϵ, δ).

17.1 Quantum Information Theory

For our purposes, we consider a quantum state to be some $q \in \mathbb{R}^S, \sum_{i \in S} q_i^2 = 1$. Observation of q will yield any given $i \in S$ with probability q_i^2. A matrix $U \in \mathbb{R}^{S \times S}$ is unitary if $\forall_{q | \|q\|_2 = 1} \|Uq\|_2 = 1$. This is equivalent to applying a probabilistic function to a random variable.

A mixed quantum state is a set $(\{p_j\}, \{q^j\}), p_j \geq 0, \sum_j p_j = 1, \sum_{i \in S} (q_i^j)^2 = 1$. Here the probability of measuring i is $\sum_j p_j (q_i^j)^2$. Applying U to a mixed quantum state yields $(\{p_j\}, \{Uq^j\})$. Any mixed quantum state can be entirely represented with an $S \times S$ density matrix $\rho = \sum_j p_j q^j (q^j)^T$. The probability of measuring any i is just the ith entry along the diagonal of ρ. For $(\{p_j\}, \{Uq^j\})$, the density matrix is conveniently $\sum_j p_j U q^j (q^j)^T U^T = U \rho U^T$.

Let $|S| = n$. Given that ρ has eigenvalues $\lambda_1, \ldots, \lambda_n, \lambda_i \geq 0, \sum_i \lambda_i = 1$. The entropy of ρ is $S(\rho) = \sum_i \lambda_i \log \frac{1}{\lambda_1}$ and $S(\rho) \leq \log n$.

17.2 Quantum Argument

The first step is the same as in last lecture, we must construct recovery graphs that show which pairs of queries give a success probability bounded away from $\frac{1}{2}$. As before, each graph has a matching of size $\Omega(\epsilon \delta n)$. In the linear case, each input bit is uniquely determined by each edge.

Let $C_{j_1}(x), C_{j_2}(x)$ be the values of point j_1, j_2 in the recovery graph for x_1. If there is an edge between j_1, j_2 then there is a decoding function $D(C_{j_1}(x), C_{j_2}(x))$ that gives the correct value of x_1 with probability at least $\frac{1}{2} + \frac{\epsilon}{2}$. It can be shown that if this is the case then one of the following holds:

- $C_{j_1}(x)$ is correlated to x_1
- $C_{j_2}(x)$ is correlated to x_2

- $C_{j_1}(x) \oplus C_{j_2}(x)$ is correlated to x_1

The random variables $C_1(x), x_1$ are correlated if $\mathbf{Pr}_x[C_1(x) = x_1 \vee C_1(x) \neq x_1] \geq \frac{1}{2} + \frac{\epsilon}{\sigma}$. We can assume that the third case holds by letting $C'(x) = C(x) \cdot \underbrace{0 \cdots 0}_{n}$ and attaching the j that fall into the first two cases to a zero in the graph. Hence in the matching for graph G_i, an edge (j_1, j_2) implies that $\mathbf{Pr}[c_{j_1}(x) \oplus c_{j_2}(x) = x_i] \geq \frac{1}{2} + \frac{\epsilon}{\sigma}$. Hence we can still assume we xor two bits.

Consider the code:

$$C' : \{0,1\}^k \to \{0,1\}^{2n}$$

This can be reinterpreted as unit vectors:

$$C'' : \{0,1\}^k \to \mathbb{R}^{2n}$$

We replace zeros with ones and ones with negative ones so that if $C'(x) = (0, 1, \ldots, 1, 0)$ then $C''(x) = \frac{1}{\sqrt{2n}}(1, -1, \ldots, -1, 1)$. There is a quantum decoding algorithm for this code. Knowing a pair from the reconstruction graph yields an advantage in decoding. A small example of the decoding algorithm, involving only one pair is:

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} a \\ b \end{pmatrix} = \frac{1}{2} \begin{pmatrix} a+b \\ a-b \end{pmatrix}$$

If $a = b$, then the result is $\binom{1}{0}$ or $\binom{-1}{0}$. If $a \neq b$ then the result is $\binom{0}{1}$ or $\binom{0}{-1}$. In the more general case, with the whole vector the transformation would resemble (if there is an edge between 1,2 and 4,5):

$$\frac{1}{\sqrt{2n}} \begin{pmatrix} 1 & 1 & & & & \\ 1 & -1 & & & & \\ & & 1 & & & \\ & & & 1 & 1 & \\ & & & 1 & -1 & \\ & & & & & \ddots \end{pmatrix}$$

The result can then be measured. If we measure a row not corresponding to a pair (e.g., 3), we learn nothing. Otherwise we learn whether a pair is equal or unequal. The probability that the decoding is correct is the probability of hitting a good pair multiplied by the advantage in that case. The probability of hitting a good pair is some constant, $\Omega(\epsilon^2 \delta)$. Hence the probability of a correct decoding is at least $\frac{1}{2} + \Omega(\epsilon^2 \delta)$. We will show that the entropy is at least k bits, so that $k \leq \log n$.

It is not clear how to directly analyze this quantum system, so we must consider a more complex model. Consider the quantum state $q \in \mathbb{R}^{\{0,1\}^k \times [n]}$. This system has an entry for every encoding of every string of length k. We can define the following mixed quantum state, $\forall x \in \{0,1\}^k$, $\left(\frac{1}{2^k}, q^x\right)$, where $q^x_{x',j} = 0$ if $x' \neq x$ and:

$$q^x_{x,j} = \begin{cases} \frac{1}{\sqrt{2n}} & \text{if the } j\text{th bit of } C(x) \text{ is } 0 \\ -\frac{1}{\sqrt{2n}} & \text{if the } j\text{th bit of } C(x) \text{ is } 1 \end{cases}$$

The density matrix of this mixed state is $\rho = \sum_x \frac{1}{2^k} q^x (q^x)^T$.

The classical equivalent of this system is two random variables taken from the same sample space. A classical dependence between those variables is equivalent to quantum entanglement. Here the two components are X, M where X is a random string of k bits and M is the encoding of X. The entropy of the random vector (or quantum state) is $S(XM) \geq S(X) = k$. Classically, the mutual information between X, M is $I(X, M)$. This obeys the relationship:

$$\begin{aligned} I(X,Y) &= H(X) - H(X|Y) \\ &= -H(XY) + H(X) + H(Y) \\ &= H(Y) - H(Y|X) \end{aligned}$$

The quantum equivalent of this is $S(X;M) = S(X) - S(X|M)$. An upper bound on this is $S(X;M) \leq S(M) \leq \log 2n$. Also $S(X) = k$ so to show that $S(X;M) = \Omega(k)$ it suffices to show that $S(X|M) \leq (1-\alpha)k$. Consider:

$$\begin{aligned} S(X|M) &= S(X_1|M) + S(X_2|M, X_1) + \ldots + S(X_k|M, X_1, X_2, \ldots, X_{k-1}) \\ &\leq S(X_1|M) + S(X_2|M) + \ldots + S(X_k|M) \\ &\leq k(1-\alpha) \text{ for } \alpha = 1 - H(\frac{1}{2} + \Omega(\epsilon^2 \delta)) \end{aligned}$$

That is, for every bit of the input, there is a measurement that gives that bit with probability greater than a half. Since $\alpha k \leq \log 2n, n \geq 2^{\Omega(k)}$.