CS 294-2 Quantum Communication: Holevo's Bound 3/21/07
Spring 2007 Lecture 17

1 Quantum Communication

Suppose we have a quantum channel over which we can send n entangled qubits. How much information can be sent over this quantum channel?

In the problem of quantum communication, we have an m-bit random variable $X = x_1 x_2 \cdots x_m$ which we want to send across the channel. Our general procedure will be to encode x into the n-dimensional quantum state $|\Phi_x\rangle$ and send this across the channel. The person on the other side adds an ancilla to get $|\Phi_x\rangle |0^{m-n}\rangle$, and then measures in some basis to get an m-bit random variable Y. We would like for Y to be close to X in some way.

One measure of how close X and Y are is the mutual information, denoted by $I(X:Y)$, defined as

$$I(X:Y) = H(X) + H(Y) - H((X,Y)) \tag{1}$$

Example: $X = Y$

$$\begin{aligned} I(X:Y) &= H(X) + H(X) - H((X,X)) && (2)\\ I(X:Y) &= H(X) + H(X) - H(X) && (3)\\ I(X:Y) &= H(X) && (4) \end{aligned}$$

Example: X and Y are independent

$$\begin{aligned} I(X:Y) &= H(X) + H(Y) - H((X,Y)) && (5)\\ I(X:Y) &= H(X) + H(Y) - (H(X) + H(Y)) && (6)\\ I(X:Y) &= 0 && (7) \end{aligned}$$

Let us now see what we can prove about the mutual information between X and Y for the quantum channel above. More generally, suppose that each x appears with probability p_x and is encoded into a mixed quantum state ρ_x. Then, $\rho = \sum_x p_x \rho_x$ is the density matrix of the state sent across the channel. Then we have the following, where S indicates the Von Neumann entropy:

Theorem 17.1: **(Holevo)**

$$I(X:Y) \leq S(\rho) - \sum_x p_x S(\rho_x) \tag{8}$$

Given that the Von Neumann entropy is a nonnegative quantity, it immediately follows that $I(X:Y) \leq S(\rho)$. But $S(\rho) \leq n$ as ρ describes a state on n qubits. Hence, $I(X:Y) \leq n$. So if X is uniform and $Y = X$, then $m \leq n$. Of course, this is no better than we can achieve classically.

The proof of Holevo's bound is slightly technical. The simple intuition is that, by a quantum equivalent of the data processing inequality, the mutual information between X and ρ should be an upper bound to $I(X:Y)$. Then, if we are willing to believe the Von Neumann entropy behaves similarly to its classical counterpart:

$$I(X:Y) \leq S(\rho) - S(\rho|X) = S(\rho) - \sum_x p_x S(\rho_x)$$

A related statement is much simpler to prove fully:

Theorem 17.2: **(Nayek)** *If X is a m bit binary string, we send it using n qubits, and decode it via some mechanism back to an m bit string Y, then our probability of correct decoding is given by*

$$Pr[X=Y] \leq \frac{2^n}{2^m} \tag{9}$$

This shows that any encoding using a number of qubits much smaller than m will be "really bad".

Proof: Say x gets mapped to $\left|\phi_x\right\rangle$. Consider the message space as being a $C_2^{\otimes n}$ subspace of the full decoding space. So each $\left|\phi_{x_0}\right\rangle$ is mapped to some $\left|e_{x_0}\right\rangle$. Then, since the decoder is measuring in an orthonormal state

$$Pr[X=Y] = \frac{1}{2^m}\sum_x ||P_x\left|\phi_x\right\rangle||^2 \tag{10}$$

$$Pr[X=Y] = \frac{1}{2^m}\sum_{x,j} ||\langle\phi_x|e_{x,j}\rangle||^2 \tag{11}$$

$$Pr[X=Y] \leq \frac{1}{2^m}\sum_{x_j} ||Q|e_{x,j}\rangle||^2 \tag{12}$$

where Q is some projection back into the message space, $e \rightarrow \phi$. The reverse projection will be at least as long.

Now pick your basis $\left|f_1\right\rangle ... \left|f_{2^n}\right\rangle$ orthonormal to the message space.

So the projection equals

$$= \frac{1}{2^m}\sum_{x,j}\sum_{i}^{2^n} ||\langle e_{x,j}||f_i\rangle||^2 \tag{13}$$

$$= \frac{1}{2^m}\sum_{i}^{2^n}\sum_{x,j} ||\langle e_{x,j}||f_i\rangle||^2 \tag{14}$$

But $\sum_{x,j} ||\langle e_{x,j}||f_i\rangle||^2 = 1$ because the f_i were unit vectors. Leaving us with

$$= \frac{1}{2^m}\sum_{i}^{2^n} 1 \tag{15}$$

$$= \frac{2^n}{2^m} \tag{16}$$

□

2 Random Access Codes

We have just seen that quantum does not buy us much in communication as transmitting m bits of information requires a quantum channel on $O(m)$ qubits. Consider however the scenario where the decoding party is not interested in the

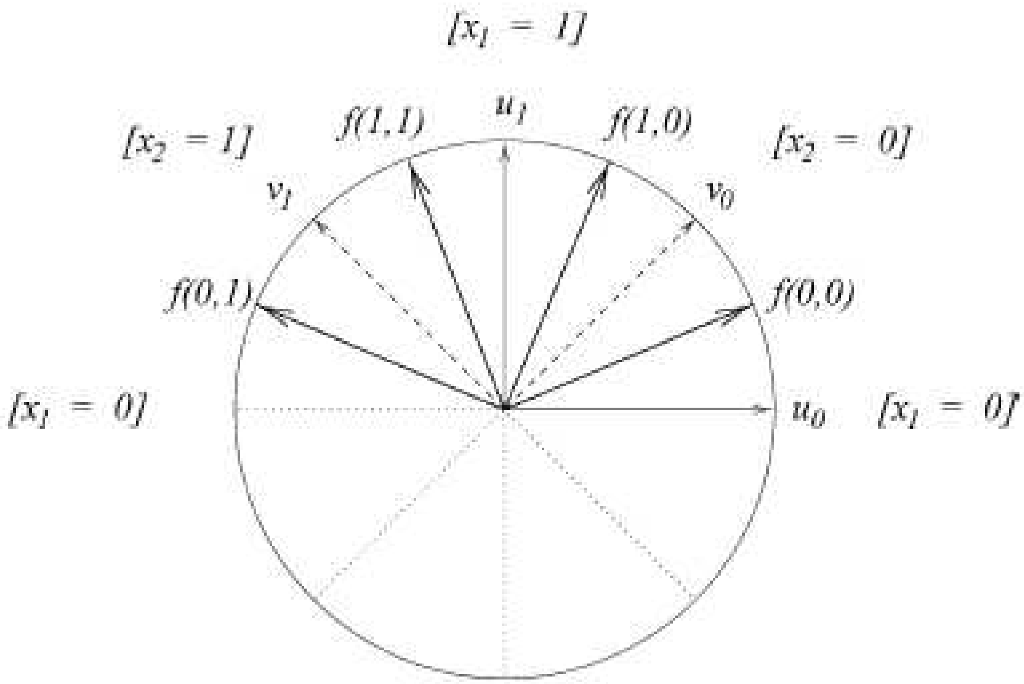

Figure 1: A quantum random access code with m=2, n=1

whole string $X_1, \cdots, X_m$ but in an arbitrary bit X_i, with i unknown to the encoder. The question is: can we recover X_i with probability $p \geq 1/2+\varepsilon$ for any i and still use very few qubits, i.e. $n = o(m)$?

Notice that the previous bound does not apply to this case. In particular, the question above does not pose itself in the classical case: in fact, if we are able to recover every single bit, we must be able to recover the whole string. However, in the quantum case, a single measurement to decode one bit will collapse the state and make it impossible to read off other bits.

2.1 Construction for m=2, n=1

Figure 2.1 shows how to encode each x into a quantum state $f(x)$, such that we can retrieve x_1 by measuring along the basis u and x_2 by measuring along the basis v. Notice that the angle between each vector in the image is $\pi/8$, so that the probability of success is $\cos^2 \pi/8 \approx .85$ for both bits.

2.2 Lower bound for random access codes

Is it possible to extend this construction to higher values of m? Notice that it is possible to show that there exists a family of c^M almost orthogonal states (with inner product $< \varepsilon$) in C^M. For $M = 2^m$, this might suggest that we could get an exponential compression in the quantum encoding. However, it turns out that mutual orthogonality does not suffice to ensure the robust decoding of every bit required by random access codes. Indeed, we are going to prove the following lower bound:

$$n \geq m(1-H(p))$$

The following lemma is a consequence of Holevo's bound:

Lemma 17.1: *Let σ_0 and σ_1 be different mixed states. Suppose that there exists a measurement yielding $b \in 0,1$ on σ_b with probability at least p. Moreover, let $\sigma = \frac{1}{2}(\sigma_0 + \sigma_1)$. Then:*

$$S(\sigma) \geq \frac{1}{2}(S(\sigma_0) + S(\sigma_1)) + (1 - H(p))$$

Proof: Let the random variable Y be the output of the measurement and take b to be uniform over $\{0,1\}$. By Holevo's bound:

$$I(b:Y) \leq S(\sigma) - \frac{1}{2}(S(\sigma_0) + S(\sigma_1))$$

But we know that $Pr[b = Y] \geq p$. This easily implies that $I(b:Y) \geq 1 - H(p)$. By combining the two inequalities and rearranging terms we obtain the theorem. $\Box$

Now we are ready to prove the lower bound. We have:

$$\rho = \frac{1}{2^n} \sum_{x \in \{0,1\}^n} \rho_x$$

For $y \in \{0,1\}^k$, $k < n$, let:

$$\rho_y = \frac{1}{2^{n-k}} \sum_{x \in \{0,1\}^{n-k}} \rho_{xy}$$

We use induction on k to prove that $S(\rho_y) \geq (1 - H(p))(m - k)$. This is trivial for $k = m$, as $S(\rho_y) \geq 0$, by definition of Von Neumann's entropy. Assuming the hypothesis for $k+1$, notice that $\rho_y = \frac{1}{2}(\rho_{0y} + \rho_{1y})$. By the lemma and the inductive hypothesis:

$$S(\rho_y) \geq \frac{1}{2}(S(\rho_{0y}) + S(\rho_{1y})) + (1 - H(p)) \geq$$

$$\frac{1}{2}(1 - H(p))2(m - k - 1) + (1 - H(p)) = (1 - H(p))(m - k)$$

as required. Finally, the fact that $S(\rho) \leq n$ yields the lower bound.