Quantum communication, information theory and entanglement assisted protocols

Sudebkumar Prasant Pal
Department of Computer Science and Engineering and
Centre for Theoretical Studies
IIT Kharagpur 721302, India
email: spp@cse.iitkgp.ernet.in

March 20, 2008

When quantum resources (apparatus) are distributed in two or more geographically separated locations, we may not be able to implement unitary operations without resorting to either quantum or classical communication. So, local operations and communication may be combined to achieve quantum computation. It is therefore necessary to determine the amount of communication necessary in such operations. Such costs may be deterministic worst case costs or even probabilistic costs, such as average communication costs over discrete distributions. We develop the necessary fundamentals and illustrate a few examples of analyses.

1 Shannon's entropy

For a random source of symbols from a certain discrete distribution, we know that (as small as) expected $H(X)$ bits of information (on the average), can be used for coding information coming out of X. If $p(x)$ is the probability of x in the source X, then this (Shannon Entropy) $H(X)$ or $H(p(x))$ is $\sum_x p(x) \log \frac{1}{p(x)}$. Once we have two such generators X and Y on the same set of symbols, we can define $H(X,Y)$ as $\sum_{(x,y)} p(x,y) \log \frac{1}{p(x,y)}$, where $p(x,y)$ is the probability that x comes out of X and y out of Y. If X and Y are independent (that is, $p(x,y) = p(x)p(y)$), then $H(X,Y) = H(X) + H(Y)$. Otherwise, $H(X,Y)$ is less than the sum of $H(X)$ and $H(Y)$. Naturally, joint entropy is less than sum of entropies if the processes are dependent.

We may view the joint entropy $H(X,Y)$ of X and Y as the sum of the entropy $H(Y)$ of Y and the conditional entropy $H(X|Y)$ of X given Y. In other words $H(X|Y)$, called the *conditional entropy of X given Y*, is the difference between the joint entropy and original entropy, i.e., $H(X|Y) = H(X,Y) - H(Y)$.

$H(X|Y)$ defined as $H(X,Y) - H(Y)$ can now be written as $\sum_{(x,y)} p(x,y) \log \frac{p(y)}{p(x,y)} = \sum_{(x,y)} p(x,y) \log \frac{1}{p(x|y)}$. The conditional entropy is like the uncommon information between X and Y, because this information is needed for X conditional over Y. So, subtracting the conditional entropy $H(X|Y)$ from $H(X)$ gives the mutual or common information $H(X:Y)$ between the two sources X and Y. that is, $H(X) - H(X|Y) = H(Y) - H(Y|X)$, usually denoted as $H(X:Y)$ or $I(X:Y)$. We may now view $H(X:Y)$ as $H(X) - H(X|Y) = H(X) - (H(X,Y) - H(Y)) = H(X) + H(Y) - H(X,Y)$, and the symmetry in its definition.

2 Density operators and von Neumann entropy

Given the density operator ρ for a quantum state, determining the von Neumann entropy $S(\rho)$ amounts to determining the (real) eigenvalues λ_x of ρ and computing $\sum_x \lambda_x \log \frac{1}{\lambda_x}$. Indeed, the spectral decomposition of ρ is $\sum_x \lambda_x |\psi_x\rangle\langle\psi_x|$, where $|\psi_x\rangle$ are the eigenvectors defining an orthonormal basis for the Hilbert space.

As operators on the state space

We will now see how these operators can operate on individual states. If ρ operates on an eigenstate $|\psi_x\rangle$ then we get $\lambda_x \rho_x$.

Postulates and traces

We already know that the expectation of a projective measurement with Hermitian observable M of a pure state $|\psi\rangle$ is $\langle\psi|M|\psi\rangle$. Writing the state as a density operator $\rho = |\psi\rangle\langle\psi|$, this expectation is $tr(M\rho) = tr(\rho M) = tr(|\psi\rangle\langle\psi|M) = \langle\psi|M|\psi\rangle$.

For density operators of mixed states and measurements using POVM measurement operators M_m for results m, see section 2.4 in [3]. Here, the measurement elements are $E_m = M_m^+ M_m$, where (by definition, measurement postulate), E_m are positive, $\sum_m E_m = I$. [M_m^+ is the adjoint of M_m.] Further, for a pure state $|\psi\rangle$, $p(m) = \langle\psi|E_m|\psi\rangle$. Such measurements are called POVM and M_m is written as $\sqrt{E_m}$. For a mixed state denoted by a density operator ρ, a unitary operation would take it to state represented by the density operator $\rho' = U\rho U^+$. A measurement yields m with probability $p(m) = tr(M_m^+ M_m \rho)$. The state resulting due to measurement of m is $\frac{M_m \rho M_m^+}{tr(M_m^+ M_m \rho)}$. We use POVM measurements in applications where the Holevo bound is used to estimate upper bounds on the mutual information between a quantum information source at one end and a measured result at the other end.

Logarithms of density operators

A method for finding $\log A$ for a diagonalizable matrix A is as follows. Let V be the matrix of eigenvectors of A (each column of V is an eigenvector of A). Find the inverse V^{-1} of V. Consider AV; observe that $AV = VA'$, where A' is a diagonal matrix whose diagonal elements are eigenvalues of A. We get $\log A'$ by replacing each diagonal element of A' by its logarithm. Now, we can write $\log A$ as $V \log A' V^{-1}$. [It is now easy to check that the operator $e^{\log A}$ is identical to the operator A. In other words, verify that $e^{\log A}|\psi\rangle = A|\psi\rangle$, for all $\psi\rangle$.] So, for a density operator A, we can write $S(A) = -tr(A \log A) = -tr(AV \log A' V^{-1}) = -tr(V^{-1}AV \log A') = -tr(A' \log A') = \sum_x \lambda_x \log \frac{1}{\lambda_x}$.

[Note also that $(\log A)^n = V(\log A')^n V^{-1}$.]

Klein's inequality

This is from [3], Theorem 11.7, page 511. The relative entropy $S(\rho||\sigma)$ is defined as $-S(\rho) - tr(\rho \log \sigma)$. Using the orthonormal decomposition of $\rho = \sum_i p_i |i\rangle\langle i|$, the first term is $\sum_i p_i \log p_i$. Since unitary operators preserve trace, the second term can be written as $-\sum_i \langle i|\rho \log \sigma|i\rangle$. Also, $\langle i|\rho = p_i\langle i|$. Since we have the orthonormal decompostion of $\sigma = \sum_j q_j |j\rangle\langle j|$, we know that $\log \sigma$ is $V \log \sigma' V^{-1}$, where σ' is the diagonal matrix with $\log q_j$ as the jst diagonal element and V is the matrix with columns given by the eigenvectors $|j\rangle$ of σ. So, the second term would be $-\sum_i p_i \langle i|V \log \sigma' V^{-1}|i\rangle = -\sum_i p_i \sum_j P_{ij} \log q_j$, where $P_{ij} = \langle i|j\rangle\langle j|i\rangle$. The rest of the proof that the relative entropy is non-negative is based on the double stochasticity of the matrix representaed by $P'_{ij}s$, and the concavity of the log function.

Projective measurements and entropy change

We know that entropy changes from $S(\rho)$ to $S(\rho')$ where $\rho' = \sum_i P_i \rho P_i$. Here, P_i are elements of the complete set of projectors of the Hermitian observable.

[We need to show that P_i commutes with $\log \rho' = V' \log \rho'' V'^{-1}$, where V is the matrix of eigenvectors of ρ' and ρ'' is the diagonal matrix of eigenvalues of ρ'. It is easy to show that $\rho' P_i = P_i \rho P_i = P_i \rho'$. Also, $P_i \log \rho' = \lambda'_i |v'_i\rangle\langle v'_i| = \log \rho' P_i$. That is, P_i commutes with ρ' as well as with $\log \rho'$. Here, λ'_i and $|v'_i\rangle$ are eigenvalues and eigenvectors of ρ'.]

[We also use the facts (i) $\sum_i P_i = I$, and (ii) $P_i^2 = P_i$.]

By Klein's inequality we know that $S(\rho||\rho')$ is non-negative. We show that $S(\rho') = -tr(\rho \log \rho')$, thereby establishing $S(\rho') \geq S(\rho)$. We have

$$-tr(\rho \log \rho') = -tr((\sum_i P_i)\rho \log \rho')$$

$$= -tr(\sum_i P_i \rho \log \rho') = -tr(\sum_i P_i \rho \log \rho' P_i)$$

$$= -tr(\sum_i P_i \rho P_i \log \rho')$$

$$-tr(\rho' \log \rho') = S(\rho')$$

Holevo's bound

Alice owns a state $\rho = \sum_{i=0}^{n} p_i \rho_i$. She encodes $X = 0, 1, \cdots, n$ as states $\rho_0, \rho_1, \cdots, \rho_n$ with probabilities $p_0, p_1, \cdots, p_n$, respectively. Bob performs a measurement described by POVM elements $\{E_y\} = \{E_0, E_1, \cdots, E_m\}$ on the (mixed) state provided by Alice and gets outcome Y. The Holevo (upper) bound on $H(X : Y)$ is $S(\rho) - \sum_i p_i S(\rho_i)$, often called the *Holevo Chi quantity,* $\chi(\rho_X)$. [The superscipt X for ρ here is simply indicative of the probability distribution over the index set X of messages x (with probability p_x), from the classical generator X.]

We consider the trio of the preparation system P, the quantum system Q, and the measuring device M and observe that initially the entire system may be viewed as represented by

$$\rho^{PQM} = \sum_x p_x |x\rangle\langle x| \otimes \rho_x \otimes |0\rangle\langle 0|$$

This is like the system P with Alice, providing the state ρ_x to Bob for measurement into the system M through the set of POVM measurement elements $\{E_y\}$ in the quantum system Q; the subsystem QM realizes the POVM measurement operation defined by $\epsilon(\sigma \otimes |0\rangle\langle 0|)$ creating the state

$$\sum_y \sqrt{E_y}\sigma\sqrt{E_y} \otimes |y\rangle\langle y|$$

Observe that in the combined system QM, covering all the elements of the POVM measurement, sets the result of the measurement in $M's$ register. Naturally, the mutual information between sources X with Alice and Y with Bob, depend of on the initial state ρ and POVM measurement.

Now note that $S(P : Q) = S(P : Q, M)$ since M is initially isolated and therefore uncorrelated with P and Q. Applying the quantum operation ϵ to subsystem QM cannot increase mutual information between P and Q. So, $S(P : Q, M) \geq S(P' : Q', M')$. Finally, discarding M' does not increase mutual information, i.e., $S(P' : Q', M') \geq S(P' : M')$. So, we have

$$S(P' : M') \leq S(P : Q)$$

The quantity $S(P : Q)$ is easily shown to be the expression of the Holevo chi quantity by using the definition of $S(\rho)$ and the *Joint Entropy Thoerem.* So, all we need to do now is to show that

$$H(X : Y) = S(P' : M')$$

This is done by tracing out Q' from $P'Q'M'$ and showing that

$$\rho^{P'M'} = \sum_{x,y} p(x, y) |x\rangle\langle x| \otimes |y\rangle\langle y|$$

where

$$\rho^{P'Q'M'} = \sum_{xy} p_x |x\rangle\langle x| \otimes \sqrt{E_y}\rho_x\sqrt{E_y} \otimes |y\rangle\langle y|$$

To see this, recall the definition of POVM measurements and the expression for the probability of the result y as $tr(\rho_x E_y)$, so that $p(y|x) = tr(\rho_x E_y) = tr(\sqrt{E_y}\rho_x\sqrt{E_y}))$. So, tracing out Q' results in the above state $\rho^{P'M'}$, whose mutual information comes out directly from the joint entropy $S(\rho^{P'M'}) = S(P', M') = H(X, Y)$ and the two traced out systems' von Neumann entropies $S(P) = H(X)$ and $S(M') = H(Y)$. These von Neumann entropies are identical to the Shannon entropies, yielding $H(X : Y) = H(X, Y) - H(X) - H(Y)$ for $S(P' : M')$.

Qubit communication complexity results: Application of Holevo's bound

We present the following results as in [1]. Alice and Bob run a quantum protocol exchanging qubits. However, they do not exploit any pre-shared quantum entanglement resource. We show that at least $\lceil \frac{n}{2} \rceil$ qubits must be sent from Alice to Bob if Alice wishes to convey n bits of (classical) information to Bob.

Bob wishes to extract n bits of information. What matters is the Holevo chi quantity at the end of the protocol in the quantum system with Bob. Let ρ_i be the density operator representing the state defined by the collection of qubits with Bob at the end of the ist step. Clearly, the information generator provides the state ρ_i^x, from the mixed state $\rho_i = \Sigma_x p_x \rho_i^x$. The upper bound on the mutual information on measurements by Bob is the Holevo chi quantity $\chi(\rho_i^X) = S(\rho_i) - \Sigma_x p_x S(\rho_i^x)$. It is easy to see that Alice's unitary operations on its own qubits do not alter this chi quantity; the qubits in Bob's system are not tampered with in such operations at Alice's end. That is, it does not alter ρ_i^X, and therefore does not alter either $S(\rho_i)$ or $\chi(\rho_i^X)$. Moreover, χ and S are invariant under unitary transformations at Bob's site. So, we consider only two non-trivial cases (i) when Alice sends a qubit to Bob, and (ii) when Bob sends a qubit to Alice. In case (i), let B denote the subsystem of qubits after i steps with Bob and Q the single new qubit obtained from Alice in the $(i+1)$st step. We know that $S(Q) \leq 1$ (a single qubit !). Also, by subadditivity property,

$$S(BQ) \leq S(B) + S(Q) \leq S(B) + 1$$

We can also show (Araki-Lieb inequality [3]) that

$$S(BQ) \geq S(B) - S(Q) \geq S(B) - 1$$

Clearly therefore,

$$S(\rho_{i+1}) \leq S(\rho_i) + 1$$

(due to subadditivity as shown above), and

$$\chi(\rho_{i+1}^X) = S(\rho_{i+1}) - \Sigma_x p_x S(\rho_{i+1}^x)$$

$$\leq (S(\rho_i) + 1) - \Sigma_x p_x (S(\rho_i^x) - 1)$$
$$= \chi(\rho_i^X) + 2$$

(due to the Araki-Lieb inequality as shown above)

In case (ii), χ cannot increase [3]; we are tracing out a single qubit from Bob's site. So,

$$\chi(\rho_{i+1}^X) \leq \chi(\rho_i^X)$$

Further, by the Araki-Lieb inequality, we have

$$S(\rho_{i+1}) \leq S(\rho_i) + 1$$

We therefore conclude that the chi quantity goes up (by 2 units), only when a qubit is sent from Alice to Bob. It is now clear that Alice would have to send at least $\lceil \frac{n}{2} \rceil$ qubits to Bob to raise the chi quantity at Bob's end to at least n, so that Bob may extract n bits of information.

Further, observe that whenever a qubit is communicated (either way), the von Neumann entropy does not decrease at Bob's end. The entropy may rise by at most one unit. So, the total rise in entropy at Bob's end is less that the total number of qubits comunicated either way. Since the entropy was initially zero and second term in the *chi* quantity is also initially zero and finally non-zero, we can say that the rise in entropy exceeds the rise in the *chi* quantity, or equivalently, exceeds the final *chi* quantity. This *chi* quantity clearly must be larger than n, the number of classical bits conveyed from Alice to Bob. Since the total number of qubits communicated exceeds the net rise in entropy, we can say that this also exceeds the total number of classical bits conveyed. So, although at least $\lceil \frac{n}{2} \rceil$ qubits need to be comminicated from Alice to Bob, a total of at least n qubits need to be communicated in order to transfer n bits of classical information from Alice to Bob.

References

[1] R. Cleve, W. van Dam, M, Nielsen and A.Tapp, Qunatum entanglement and the communication complexity of the inner product function, quant-ph 9708019 v3, 1998.

[2] J. Gruska, *Quantum Computing*, McGraw-Hill, 1990.

[3] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, 2002.