Quantum Information Theory III

Sibasish Ghosh

The Institute of Mathematical Sciences
CIT Campus, Taramani, Chennai 600 113, India.

Recalling Schumacher compression

- **Signal $x \in X$, which occurs with probability p_x, is encoded as a pure state $|\phi_x\rangle \in H_S$: Initial density matrix, corresponding to the non-spectral ensemble $\mathcal{E}_{NS} = \{p_x; |\phi_x\rangle | x \in X\}$, is $\rho = \sum_{x \in X} p_x |\phi_x\rangle\langle\phi_x|$. $|\phi_x\rangle$'s are non-orthogonal, in general. The spectral ensemble $\mathcal{E}_S = \{w_l; |\psi_l\rangle | l \in L\}$ corresponds to the spectral decomposition $\rho = \sum_{l \in L} w_l |\psi_l\rangle\langle\psi_l|$ of ρ.**

Recalling Schumacher compression (continued)

- **Consider a string $\rho^{\otimes n} = \rho \otimes \rho \otimes \ldots (n \text{ times})$ of length n. So we have now the non-spectral ensemble $\mathcal{E}_{NS}^{(n)} = \{P_{NS}(x_1, x_2, \ldots, x_n) \equiv p_{x_1} p_{x_2} \ldots p_{x_n}; |\Phi_{NS}(x_1, x_2, \ldots, x_n)\rangle \equiv |\phi_{x_1}\rangle \otimes |\phi_{x_2}\rangle \otimes |\phi_{x_n}\rangle | (x_1, x_2, \ldots, x_n) \in X^n\}$ as well as the spectral ensemble**
$\mathcal{E}_{S}^{(n)} = \{P_{S}(l_1, l_2, \ldots, l_n) \equiv w_{l_1} w_{l_2} \ldots w_{l_n}; |\Phi_{S}(l_1, l_2, \ldots, l_n)\rangle \equiv |\psi_{l_1}\rangle \otimes |\psi_{l_2}\rangle \otimes |\psi_{l_n}\rangle | (l_1, l_2, \ldots, l_n) \in L^n\}$**.**

Recalling Schumacher compression (continued)

- **Consider the random variable** $L \equiv \{l; w_l | l \in L\}$**. So** $H(L) = S(\rho)$**. Consider the classical messages** $l_1 l_2 \ldots l_n$ **with associated probabilities** $P_S(l_1, l_2, \ldots, l_n)$**. Given** $\epsilon, \delta > 0$**, for sufficiently large** n**,** $l_1 l_2 \ldots l_n$ **will be a typical sequence if** $2^{-n(H(L)-\delta)} \geq P_S(l_1, l_2, \ldots, l_n) \geq 2^{-n(H(L)+\delta)}$ **and the total probability of all such typical messages exceeds** $1 - \epsilon$**. So the total no.** $N(\epsilon, \delta; n)$ **of such typical sequences will satisfy:** $2^{n(H(L)+\delta)} \geq N(\epsilon, \delta; n) \geq (1 - \epsilon) 2^{n(H(L)-\delta)}$**.**

Recalling Schumacher compression (continued)

- **The states $|\Phi_S(l_1, l_2, \ldots, l_n)\rangle$ (whose total no. is $N(\epsilon, \delta; n)$), corresponding to the typical sequences $l_1 l_2 \ldots l_n$, are pairwise orthogonal and so they will span a $N(\epsilon, \delta; n)$-dimensional subspace (called as the 'typical' subspace, and is denoted by Λ) of $H_S^{\otimes n}$.**
- **Consider now the projective measurement $\{E, I - E\}$ on $\rho^{\otimes n}$, where $E : H_S^{\otimes n} \rightarrow \Lambda$ is the projector on Λ and $(I - E)$ is the projector on $\Lambda^{\perp}$.**

Recalling Schumacher compression (continued)

- **If E clicks, encode the states $|\Phi_{typical}\rangle$ (of Λ) in the range of $E\rho^{\otimes n}E$ as: $\mathcal{A}(|\Phi_{typical}\rangle\langle\Phi_{typical}|) = |\Psi_{comp}\rangle$, where $|\Psi_{comp}\rangle$ is a $2^{n(S(\rho)+\delta)}$-qubit state and $\mathcal{A}$ is a CP map. If $(I - E)$ clicks, take the output state, cooresponding to input the state $|\Phi_{NS}(x_1, x_2, \ldots x_n)\rangle$ (appeared in the non-spectral ensemble for $\rho^{\otimes n}$), as $\rho_{junk}(x_1, x_2, \ldots, x_n)$.**

Recalling Schumacher compression (continued)

- **Thus we see that, after this measurement, a general state** $|\Phi_{NS}(x_1, x_2, \ldots, x_n)\rangle\langle\Phi_{NS}(x_1, x_2, \ldots, x_n)| \mapsto \sum_{|\Phi_{typical}\rangle\in\Lambda} |\langle\Phi_{typical}|\Phi_{NS}(x_1, x_2, \ldots, x_n)\rangle|^2 |\Psi_{comp}\rangle\langle\Psi_{comp}| + \rho_{junk}(x_1, x_2, \ldots, x_n)\langle\Phi_{NS}(x_1, x_2, \ldots x_n)|(I - E)|\Phi_{NS}(x_1, x_2, \ldots, x_n)\rangle \equiv \rho_{NS}(x_1, x_2, \ldots, x_n)$**, where** $\{|\Phi_{typical}\rangle\}$ **is a complete orthonormal basis of the** $2^{n(S(\rho)+\delta)}$**-dimensional Hilbert space and** $\{|\Psi_{comp}\rangle\}$ **is the corresponding encoded set.**

Recalling Schumacher compression (continued)

- **The average fidelity F of this measurement $=$ encoding scheme:**
$F = \sum_{(x_1,x_2,\ldots,x_n)\in X} P_{NS}(x_1,x_2,\ldots,x_n)\times$
$\langle\Phi_{NS}(x_1,x_2,\ldots,x_n)|\rho_{NS}(x_1,x_2,\ldots,x_n)|\Phi_{NS}(x_1,x_2,\ldots,x_n)\rangle$
$> 1-2\epsilon$**.**
- **Once E clicks, because of knowledge of the encoding scheme $\mathcal{A}(|\Phi_{typical}\rangle\langle\Phi_{typical}|) = |\Psi_{comp}\rangle$, one can now perform a (unitary) decoding scheme $U(|\Psi_{comp}\rangle\otimes|0\rangle) = |\Phi_{typical}\rangle$, by appending extra dimension to $2^{n(S(\rho)+\delta)}$-dimensional Hilbert space.**

Recalling Schumacher compression (continued)

- **This compression scheme is <u>optimal</u>: If we want to encode states of Λ by states of a $2^{n(S(\rho)-\delta)}$-dimensional Hilbert space Λ', one can generate (through the scheme: measurement $\rightarrow$ encoding $\rightarrow$ <u>unitary</u> decoding) only a $2^{n(S(\rho)-\delta)}$-dimensional subspace (Λ'', say) of $H_S^{\otimes n}$. If E'' is the projector on the subspace Λ'', then $Tr(\rho^{\otimes n} E'') \leq$ sum of the first $2^{n(S(\rho)-\delta)}$ no. of largest eigen values of $\rho^{\otimes n}$, which, in turn, less that ϵ (follows from properties of typical subspace). So the average fidelity $F \leq Tr(\rho^{\otimes n} E'') < \epsilon$.**

Recalling Schumacher compression (continued)

- In the case when $|\phi_x\rangle$'s are pairwise orthogonal, the incompressible information content (in terms of qubits) in $\rho = \sum_{x\in X} p_x|\phi_x\rangle\langle\phi_x|$ is $S(\rho) = H(X)$. By the compression scheme, the decoder will be able to distinguish the encoded states perfectly, in the large n limit. For the case of non-orthogonal $|\phi_x\rangle$'s, the incompressible information content is $S(\rho)$, which is strictly less than $H(X)$. But we have to pay the price: the decoder will not be able to distinguish the encoded state.

Data compression for mixed state encoding

- **In the case of mixed state encoding with the ensemble $\mathcal{E} = \{p_x, \rho_x | x \in X\}$ with $\rho = \sum_{x \in X} p_x \rho_x$, it can be shown that the incompressible information content (in terms of qubits) in ρ will be less than or equal to the Holevo bound $\chi(\mathcal{E}) \equiv S(\rho) - \sum_{x \in X} p_x S(\rho_x)$. But whether one can achieve the limit $\chi(\mathcal{E})$ in an asymptotic compression scheme, is still not fully settled.**

Data compression for mixed state encoding (continued)

- **That one needs less that $S(\rho)$ no. of qubits as incompressible information content for mixed state encoding, in general, can be seen in the trivial case: Consider $p_{x_0} = 1$ and $p_x = 0$ for all $x \in (X - \{x_0\})$, where it is assumed that $S(\rho_{x_0}) > 0$. So it is certain that the ensemble $\mathcal{E}$ has been prepared in the state ρ_{x_0}. So there is nothing to be compressed – everything is known, even though $S(\rho_{x_0})$ is positive!**

Mutual information vs. Holevo bound

- **In classical information theory, the mutual information $I(Y;X) = H(Y) - H(Y|X)$ tells us how much, on the average, the Shannon entropy of the random variable $Y = \{y, q_y | y \in Y\}$ is reduced once we learn the values of $X = \{x, p_x | x \in X\}$. Similarly, in quantum information theory, the Holevo bound $\chi(\mathcal{E})$ tells us how much, on the average, von Neumann entropy $S(\rho)$ (with $\rho = \sum_{x \in X} p_x \rho_x$) is reduced once we know which preparation procedure was chosen to prepare ρ.**

Mutual information vs. Holevo bound (continued)

- $I(Y;X) \geq 0$; $\chi(\mathcal{E}) = S(\rho) - \sum_{x\in X} p_x S(\rho_x) = S(\sum_{x\in X} p_x \rho_x) - \sum_{x\in X} p_x S(\rho_x) \geq 0$ **(due to concavity of von Neumann entropy).**

Mutual information vs. Holevo bound (continued)

- **Lindblad-Uhlman monotonicity: If** $\mathcal{T} : \mathcal{D}(H_S) \to \mathcal{D}(H_S)$ **is a (trace-preserving) CP map, then** $\chi(\mathcal{E}') \leq \chi(\mathcal{E})$ **where** $\mathcal{E}' = \{p_x, \mathcal{T}(\rho_x) | x \in X\}$ **while** $\mathcal{E} = \{p_x, \rho_x\}$**. Note that this inequality does not always hold good for von Neumann entropy.**
- **Due to the analogy of** $\chi(\mathcal{E})$ **with** $I(Y; X)$**, the Holevo bound** $\chi(\mathcal{E})$ **may be interpreted as the amount of <u>classical</u> information about the signal** x **that one can extract from the ensemble** $\mathcal{E} = \{p_x, \rho_x | x \in X\}$ **by performing measurements.**

Accessible information

- **Suppose Alice prepares a density matrix ρ_x of S with probability p_x according to the ensemble $\mathcal{E} = \{p_x, \rho_x | x \in X\}$. She then sends the state ρ_x to Bob via a noiseless quantum channel. Bob does know the ensemble $\mathcal{E}$ but he doesn't know which state is sent by Alice. Bob wants to know the signal x (i.e., Bob wants to extract the (classical) information about the random variable $X = \{x, p_x | x \in X\}$) by performing a POVM $\{E_y | y \in Y\}$ on the state with measurement outcomes $y \in Y$ and probabilities $p(y|x) = Tr(E_y \rho_x)$.**

Accessible information (continued)

- **Using Bayes' rule, we have** $p(x|y) = (p(y|x)p_x)/(\sum_{x \in X} p(y|x)p_x)$**. So the mutual information** $I(X;Y) \equiv H(X) - H(X|Y) = -\sum_{x \in X} p_x \log_2 p_x + \sum_{y \in Y} (\sum_{x \in X} p(y|x)p_x) \times H(X|y)$**, which is the amount of reduction of (classical) information about** X**, on an average, after Bob performs the measurement. Bob wants to <u>maximize</u> this reduction by choosing appropriate POVM.**

Accessible information (continued)

- **Accessible information:**

$$I_{acc} \equiv \max\{I(X;Y)| \text{ over all POVMs } \{E_y|y \in Y\}\}$$

- **In general, it is difficult to calculate I_{acc}.**
- **Holevo (1973) has shown that**

$I_{acc} \leq S(\rho) - \sum_{x\in X} p_x S(\rho_x) \equiv \chi(\mathcal{E})$.

Accessible information (continued)

- Proof of this result uses <u>strong subadditivity</u> of von Neumann entropy, for which we need three systems: An preparation system P of dimension $|X|$ having a complete orthonormal basis $\{|x\rangle : x \in X\}$, the original system S which is prepared in state ρ_x with probability p_x (corresponding to the ensemble $\mathcal{E} = \{p_x, \rho_x | x \in X\}$ for the average state $\rho_S = \sum_{x \in X} p_x \rho_x$), and the measuring apparatus M of dimension $\geq |Y|$ with an orthonomal basis (possibly incomplete) $\{|y\rangle : y \in Y\}$.

Quantum conditional entropy

- **Quantum conditional entropy: For any density matrix ρ_{AB} of a bipartite system $A+B$, with reduced density matrices $\rho_A = Tr_B(\rho_{AB})$ and $\rho_B = Tr_A(\rho_{AB})$, the quantum conditional entropy $S(A|B) \equiv S(\rho_{AB}) - S(\rho_B)$. Unlike classical case, $S(A|B)$ can be negative!**

Quantum mutual entropy

- **Quantum mutual entropy:** For any density matrix ρ_{AB} of a bipartite system $A+B$, with reduced density matrices $\rho_A = Tr_B(\rho_{AB})$ and $\rho_B = Tr_A(\rho_{AB})$, $S(A;B) \equiv S(\rho_A) - S(A|B) \equiv S(\rho_A) + S(\rho_B) - S(\rho_{AB})$. It is always non-negative due to subadditivity property of von Neumann entropy.

Results from strong subadditivity

- **(i) Ignoring subsystem:** $S(A;B,C) \geq S(A;B), S(A;C)$.

- **(ii) One-sided CP map: If** $\mathcal{T} : \mathcal{B}(H_B) \to \mathcal{B}(H_{B'})$ **is a trace-preserving CP map such that** $(I \otimes \mathcal{T})(\rho_{AB}) = \rho'_{AB'} \in \mathcal{D}(H_A \otimes H_{B'})$ **for each** $\rho_{AB} \in \mathcal{D}(H_A \otimes H_B)$**, then** $S(A;B) \geq S(\rho'_A) + S(\rho'_{B'}) - S(\rho'_{AB}) \equiv S(A';B')$.

Useful property of von Neumann entropy

- **(iii) If $\{|x\rangle : x \in X\}$ is an orthonormal basis of the $|X|$-dimensional preparation system P, then $S(\sum_{x\in X} p_x|x\rangle\langle x| \otimes \rho_x) = H(X) + \sum_{x\in X} p_x S(\rho_x)$.**
- **This is true because of the fact that here $S(\sum_{x\in X} p_x|x\rangle\langle x| \otimes \rho_x) \equiv -\sum_{x\in X} Tr((p_x|x\rangle\langle x| \otimes \rho_x)\log_2(p_x|x\rangle\langle x| \otimes \rho_x))$, as $|x\rangle\langle x| \otimes \rho_x$'s are pairwise orthogonal.**

Proof of $I_{acc} \le \chi(\mathcal{E})$

- **Alice prepares the initial ensemble, described by $\rho_{PSM} = \sum_{x \in X} p_x |x\rangle\langle x| \otimes \rho_x \otimes |0\rangle\langle 0|$, where $|0\rangle$ is a fixed state of M. She then sends (undisturbedly) the systems S and M to Bob.**
- **So $S(\rho_{PSM}) = S(\rho_{PS})$. Now $S(P;S,M) = S(\rho_P) + S(\rho_{SM}) - S(\rho_{PSM}) = S(\sum_{x \in X} p_x |x\rangle\langle x|) + S(\sum_{x \in X} p_x \rho_x \otimes |0\rangle\langle 0|) - S(\rho_{PS}) = S(\rho_P) + S(\rho_S) - S(\rho_{PS}) = S(P;S)$.**

Proof of $I_{acc} \leq \chi(\mathcal{E})$ (continued)

- **Bob applies now a CP map** $\mathcal{T} : \mathcal{D}(H_S \otimes H_M) \rightarrow \mathcal{D}(H_S \otimes H_M)$ **corresponding to implimentation of the POVM** $E_y | y \in Y\}$**, given by the Kraus representation** $\mathcal{T}(\sigma_S \otimes |0\rangle\langle 0|) = \sum_{y \in Y} (E_y^{1/2} \otimes U_y)(\sigma_S \otimes |0\rangle\langle 0|)(E_y^{1/2} \otimes U_y^{\dagger})$**, where** σ_S **is any state of** S **and for each** $y \in Y$**,** $U_y : H_M \rightarrow H_M$ **is an unitary operator for which** $U_y|0\rangle = |y\rangle$**.**

Proof of $I_{acc} \leq \chi(\mathcal{E})$ (continued)

- **Let** $\rho'_{PSM} = (I_P \otimes \mathcal{T})(\rho_{PSM}) = \sum_{x\in X,\ y\in Y} p_x |x\rangle\langle x| \otimes (E_y^{1/2} \rho_x E_y^{1/2}) \otimes |y\rangle\langle y|$**. Then by property (ii),** $S(P;S,M) \geq S(P';S',M')$**.**
- **So by property (i), we have** $S(P';S',M') \geq S(P';M')$**.**
- **So we have finally** $S(P';M') \leq S(P;S)$**.**

Proof of $I_{acc} \leq \chi(\mathcal{E})$ (continued)

- **Now** $S(P;S) = S(\rho_P) + S(\rho_S) - S(\rho_{PS}) = H(X) + S(\rho_S) - H(X) - \sum_{x \in X} p_x S(\rho_x)$ **(by property (iii)).** **Thus** $S(P;S) = S(\rho_S) - \sum_{x \in X} p_x S(\rho_x) = \chi(\mathcal{E})$**.**
- $P(P';M') = S(\rho'_P) + S(\rho'_M) - S(\rho'_{PM}) = S(\sum_{x \in X,\ y \in Y} p_x p(y|x) |x\rangle\langle x|) + S(\sum_{x \in X,\ y \in Y} p_x p(y|x) |y\rangle\langle y|) - S(\sum_{x \in X,\ y \in Y} p_x p(y|x) |xy\rangle\langle xy|) = H(X) + H(Y) - H(X,Y) = I(X;Y)$**.**
- **Thus we have** $I(X;Y) \leq \chi(\mathcal{E})$**, whatever be POVM** $\{E_y | y \in Y\}$**.**

Attainability of the Holevo bound

- **Example 1:** Let $|\psi_1\rangle$ and $|\psi_2\rangle$ be two non-orthogonal states (spanning a two-dimensional Hilbert space H), supplied with equal probability. Thus $\mathcal{E} = \{|\psi_1\rangle, p(|\psi_1\rangle) = 1/2 \text{ and } |\psi_2\rangle, p(|\psi_2\rangle) = 1/2\}$. Here the accessible information $I_{acc}(\mathcal{E})$ can be shown to be the value of $I(X;Y)$ corresponding to the optimal POVM $\{F_1 = (1/(1 + |\langle\psi_1|\psi_2\rangle|))|\psi_2^{\perp}\rangle\langle\psi_2^{\perp}|, F_2 = (1/(1 + |\langle\psi_1|\psi_2\rangle|))|\psi_1^{\perp}\rangle\langle\psi_1^{\perp}|, F_3 = I_H - F_1 - F_2\}$.

Attainability of the Holevo bound (Exmaple 1)

- **Example 1 (continued): Here** $I_{acc}(\mathcal{E}) = 1 - |\langle\psi_1|\psi_2\rangle|$ **while** $\chi(\mathcal{E}) = H((1 + |\langle\psi_1|\psi_2\rangle|)/2, (1 - |\langle\psi_1|\psi_2\rangle|)/2)$**. So** $I_{acc} \leq \chi(\mathcal{E})$**, equality holds iff either** $|\langle\psi_1|\psi_2\rangle| = 0$ **or** $|\langle\psi_1|\psi_2\rangle| = 1$**.**

Attainability of the Holevo bound (Example 2)

- **Example 2: Let $|\psi_1\rangle$ and $|\psi_2\rangle$ be same as in Example 1, supplied with equal probability. Now, instead of taking the initial ensemble as $\mathcal{E}$, let us take $\mathcal{E}^{(2)} = \{|\Psi_1\rangle \equiv |\psi_1\rangle \otimes |\psi_1\rangle, p(|\Psi_1\rangle) = 1/2 \text{ and } |\Psi_2\rangle \equiv |\psi_2\rangle \otimes |\psi_2\rangle, p(|\Psi_2\rangle) = 1/2\}$. Let H' be the Hilbert space spanned by $|\Psi_1\rangle$ and $|\Psi_2\rangle$. Here the accessible information $I_{acc}(\mathcal{E}^{(2)})$ can be shown to be the value of $I(X;Y)$ corresponding to the optimal POVM $\{G_1 = (1/(1 + |\langle\Psi_1|\Psi_2\rangle|))(I_{H'} - |\Psi_2\rangle\langle\Psi_2|), G_2 = (1/(1 + |\langle\Psi_1|\Psi_2\rangle|))(I_{H'} - |\Psi_1\rangle\langle\Psi_1|, G_3 = I_{H'} - G_1 - G_2\}$.**

Attainability of the Holevo bound (Exmaple 2) (continued)

- **Example 2 (continued): Here** $I_{acc}(\mathcal{E}^{(2)}) = 1 - |\langle\psi_1|\psi_2\rangle|^2$ **while** $\chi(\mathcal{E}^{(2)}) = H((1 + |\langle\psi_1|\psi_2\rangle|^2)/2, (1 - |\langle\psi_1|\psi_2\rangle|^2)/2)$**. So** $I_{acc}(\mathcal{E}^{(2)}) \leq \chi(\mathcal{E}^{(2)})$**, equality holds iff either** $|\langle\psi_1|\psi_2\rangle| = 0$ **or** $|\langle\psi_1|\psi_2\rangle| = 1$**. Note that** $\{\chi(\mathcal{E}^{(2)}) - I_{acc}(\mathcal{E}^{(2)})\} \leq \{\chi(\mathcal{E}) - I_{acc}(\mathcal{E})\}$ **and equality holds iff either** $|\langle\psi_1|\psi_2\rangle| = 0$ **or** $|\langle\psi_1|\psi_2\rangle| = 1$**.**

Entanglement dilution for pure states

- **Far apart parties Alice and Bob want to share n copies of a bipartite pure entangled state $|\psi\rangle_{AB}$, for large n, starting from minimum no. (k_{min}, say) of the shared siglet state $|\phi^+\rangle = (1/\sqrt{2})(|00\rangle + |11\rangle)$, by using LOCC only. k_{min}/n, in the limit $n \to \infty$, is called as the entanglement of formation $E_F(|\psi\rangle_{AB})$ of $|\psi\rangle_{AB}$.**

Entanglement concentration for pure states

- **Far apart parties Alice and Bob are sharing share n copies of a bipartite pure entangled state $|\psi\rangle_{AB}$, for large n, and they want to generate now, by using LOCC only, maximum no. (k_{max}, say) of siglet states $|\phi^+\rangle$. The value k_{max}/n, in the limit $n \rightarrow \infty$, is called the distillable entanglement $E_D(|\psi\rangle_{AB})$ of $|\psi\rangle_{AB}$.**

$$E_D(|\psi\rangle) \leq E_F(|\psi\rangle)$$

- **As entanglement does not increase under LOCC (which is known as "irreversibility of entanglement under the thermodynamic process of LOCC"), therefore:** $k_{max}/n \leq k_{min}/n$ **in the large** n **limit. So**

$$E_D(|\psi\rangle_{AB}) \leq E_F(|\psi\rangle_{AB}). \tag{1}$$

- **Using Schumacher data compression, one can show that, in the large** n **limit, Alice and Bob can share** n **no. of entangled state** $|\psi\rangle_{AB}$**, by using LOCC only, starting from** $nS(\rho_A^\psi)$ **no. of shared siglet states** $|\phi^+\rangle$**, where** $\rho_A^\psi = Tr_B(|\psi\rangle_{AB}\langle\psi|)$ **.**

$$E_D(|\psi\rangle) \leq E_F(|\psi\rangle)$$

- **There is an entanglement concentration scheme, using which Alice and Bob can finally share $nS(\rho_A^{\psi})$ no. of singlet state, in the large n limit, starting from n copies of the shared state $|\psi\rangle_{AB}$.**
- **So we have** $nS(\rho_A^{\psi})/n \leq E_D(|\psi\rangle_{AB}) \leq E_F(|\psi\rangle_{AB}) \leq nS(\rho_A^{\psi})/n$. **Thus:** $E_D(|\psi\rangle_{AB}) = E_F(|\psi\rangle_{AB}) = S(\rho_A^{\psi})$.

Entanglement dilution scheme

- **Let n be large and let Alice and Bob share $nS(\rho_A^\psi)$ no. of copies of the singlet state $|\phi^+\rangle_{AB}$.**
- **Alice now locally prepares n copies of the state $|\psi_{AD}\rangle$, having Schmidt decomposition $|\psi\rangle_{AD} = \sum_{i=1}^d \sqrt{\lambda_i}|e_i\rangle_A \otimes |f_i\rangle_D$. So $\rho_D^\psi = \sum_{i=1}^d \lambda_i |f_i\rangle\langle f_i|$.**
- **Alice now uses Schumacher data compression on the string $(\rho_D^\psi)^{\otimes n}$ to compress it to the typical subspace of $nS(\rho_D^\psi) = nS(\rho_A^\psi)$ qubits.**
- **Note that $S(\rho_A^\psi) = -\sum_{i=1}^d \lambda_i \log_2 \lambda_i \equiv H(\lambda_1, \lambda_2, \ldots, \lambda_d)$.**

Entanglement dilution scheme (continued)

- **Alice now sends the states of this typical subspace to Bob, by using standard teleportation protocol using the shared $nS(\rho_A^\psi)$ no. of shared singlet states.**
- **So, in this teleportation scheme, $2nS(\rho_A^\psi)$ bits of classical communication from Alice to Bob is necessary.**
- **After receving the states of the typical subspaces, Bob now decompress them to the string $(\rho_B^\psi)^{\otimes n}$.**
- **So now Alice and Bob share n copies of the state $|\psi\rangle_{AB}$.**
- **Thus, from the definition of E_F, we have:**

$$E_F(|\psi\rangle_{AB}) \leq H(\lambda_1, \lambda_2, \ldots, \lambda_d). \tag{2}$$

Entanglement concentration scheme

- **For simplicity, we consider that Alice and Bob are sharing n copies of the state $|\psi\rangle = a|00\rangle_{AB} + b|11\rangle_{AB}$, where $a, b > 0$ and $a^2 + b^2 = 1$ and n is large enough. We would like to extract as many singlet states $|\phi^+\rangle_{AB}$ as possible, using LOCC only.**
- **So the joint state of Alice and Bob will be now a linear superposition of product states of the form $|x\rangle_A \otimes |x\rangle_B$ where $x = \sum_{j=0}^{n-1} a_j 2^j \equiv a_0 a_1 \ldots a_{n-1}$. So $|x\rangle$ is a product state of k no. of single-qubit states, each in $|0\rangle$ and $(n-k)$ no. of single-qubit states, each in $|1\rangle$, and we write $|x\rangle$ as $|0^k 1^{n-k}\rangle$.**

Entanglement concentration scheme (continued)

- **The square of the modulus of the coefficient of the state $|x\rangle_A \otimes |x\rangle_B$ is $a^{2k}b^{2(n-k)}$, and the total no. of such states is ${}^nC_k$. Let $|\psi(0^k; 1^{n-k})\rangle_{AB}$: normalized equal superposition of <u>all</u> such states. So the probability associated with $|\psi(0^k; 1^{n-k})\rangle_{AB}$ is $P(a^2, k; n) \equiv {}^nC_k \times a^{2k}b^{2(n-k)}$.**
- **The dimension of the subspace ($H_A(0^k; 1^{n-k})$, say) of $(\mathbb{C}^2)^{\otimes n}$, spanned by the ${}^nC_k$ no. of pairwise orthogonal states $|0^k 1^{n-k}\rangle$, is ${}^nC_k$. Let $P_A(0^k; 1^{n-k})$ be the projector on $H_A(0^k; 1^{n-k})$.**
- **Alice now performs the projective measurement: $\{P_A(0^k; 1^{n-k}), I_{2^n \times 2^n} - P_A(0^k; 1^{n-k})\}$ on her n-qubit system when Alice and Bob share the state $|\psi\rangle^{\otimes n}$.**

Entanglement concentration scheme (continued)

- **Note that** $(P_A(0^k;1^{n-k}) \otimes I^{(B)}_{2^n \times 2^n})(|\psi\rangle^{\otimes n}_{AB}) = \sqrt{{}^nC_k}\, a^k b^{n-k} |\psi(0^k;1^{n-k})\rangle_{AB}$. **So, when** $P_A(0^k;1^{n-k})$ **clicks in the measurement of Alice (which occurs with probability** ${}^nC_k \times a^{2k} b^{2(n-k)}$**), the shared state between Alice and Bob will be** $|\psi(0^k;1^{n-k})\rangle_{AB}$.
- **What should be the value of** k**, in the large** n **limit, so that the above-mentioned probability of occurrance of** $P_A(0^k;1^{n-k})$ **can become very much close to unity (and also approaches unity as** $n \to \infty$**)? Stirling's approximation provides us this value:** $k = na^2$**! And so, for this value of** k**,** ${}^nC_k \approx 2^{nH(a^2,b^2)}$.

Entanglement concentration scheme (continued)

- **Thus the finally shared state $|\psi(0^k;1^{n-k})\rangle_{AB}$ will have, in this case, $2^{nH(a^2,b^2)}$ no. of terms $|x\rangle_A \otimes |x\rangle_B$, all with <u>same</u> coefficient. How much entanglement is there in the state $|\psi(0^k;1^{n-k})\rangle_{AB}$?**
- **For the product states $|x\rangle_A \otimes |x\rangle_B$, appearing in $|\psi(0^k;1^{n-k})\rangle_{AB}$, we can arrange all these $2^{nH(a^2,b^2)}$ no. of non-negative numbers $x = \sum_{j=0}^{n-1} a_j 2^j \equiv a_0 a_1 \ldots a_{n-1}$ (where eaxctly $k = na^2$ no. of a_i's are equal to 0 and rest $(n-k)$ no. of a_i's are equal to 1) in the increasing order and call them as 1, 2, 3, $\ldots$, $2^{nH(a^2,b^2)}$. Thus $|\psi(0^k;1^{n-k})\rangle_{AB} = 2^{-nH(a^2,b^2)/2}(|1\rangle_A \otimes |1\rangle_B + |2\rangle_A \otimes |2\rangle_B + |3\rangle_A \otimes |3\rangle_B + \ldots + |2^{nH(a^2,b^2)}\rangle_A \otimes |2^{nH(a^2,b^2)}\rangle_B)$.**

Entanglement concentration scheme (continued)

- **As $|1\rangle_A$, $|2\rangle_A$, $\ldots$, $|2^{nH(a^2,b^2)}\rangle_A$ are pairwise orthonormal states forming a basis for an $nH(a^2,b^2)$-qubit Hilbert space $\mathcal{C}^2_{A_1} \otimes \mathcal{C}^2_{A_2} \otimes \ldots \otimes \mathcal{C}^2_{A_{nH(a^2,b^2)}}$, therefore, without loss of any generality, we can write**
$|1\rangle_A = |0\rangle_{A_1} \otimes |0\rangle_{A_2} \otimes \ldots \otimes |0\rangle_{A_{nH(a^2,b^2)-1}} \otimes |0\rangle_{A_{nH(a^2,b^2)}}$,
$|2\rangle_A = |0\rangle_{A_1} \otimes |0\rangle_{A_2} \otimes \ldots \otimes |0\rangle_{A_{nH(a^2,b^2)-1}} \otimes |1\rangle_{A_{nH(a^2,b^2)}}$, $\ldots$,
$|2^{nH(a^2,b^2)}\rangle_A = |1\rangle_{A_1} \otimes |1\rangle_{A_2} \otimes \ldots \otimes |1\rangle_{A_{nH(a^2,b^2)-1}} \otimes |1\rangle_{A_{nH(a^2,b^2)}}$.

- **Thus we can now write: $|\psi(0^k; 1^{n-k})\rangle_{AB} = |\phi^+\rangle_{A_1B_1} \otimes |\phi^+\rangle_{A_2B_2} \otimes \ldots \otimes |\phi^+\rangle_{A_{nH(a^2,b^2)}B_{nH(a^2,b^2)}}$, which is nothing but $nH(a^2,b^2)$ no. of two-qubit singlet state $|\phi^+\rangle$, shared between Alice and Bob.**

Entanglement concentration scheme (continued)

- **Thus we see that, starting from large n copies of the shared two-qubit non-maximally entangled state $|\psi\rangle_A B = a|0\rangle_A \otimes |0\rangle_B + b|1\rangle_A \otimes |1\rangle_B$, Alice and Bob can distill out, using LOCC only, $nH(a^2, b^2)$ no. of shared two-qubit singlet state $|\phi^+\rangle_{AB}$, with probability of success approaching to 1 as $n \to \infty$.**

- **So we have:**

$$E_D(|\psi\rangle_{AB}) \geq H(a^2, b^2). \tag{3}$$

- **Note that in this concentration scheme, no classical communication is needed.**

Entanglement concentration scheme (continued)

- **From equations (1), (2) and (3), it follows that:** $E_D(|\psi\rangle_{AB}) = E_F(|\psi\rangle_{AB}) = H(a^2, b^2)$**.**
- **For entanglement concentration of large** n **copies of a state** $|\psi\rangle_{AB}$ **having Schmidt decomposition** $|\psi\rangle_{AB} = \sum_{i=1}^{d} \sqrt{\lambda_i}|e_i\rangle_A \otimes |f_i\rangle_B$**, one should look (as earlier) for those product states** $|e_1^{n\lambda_1} e_2^{n\lambda_2} \ldots e_d^{n\lambda_d}\rangle_A \otimes |f_1^{n\lambda_1} f_2^{n\lambda_2} \ldots f_d^{n\lambda_d}\rangle_B$**, in the expansion of** $|\psi\rangle_{AB}^{\otimes n}$**, where, in** $|e_1^{n\lambda_1} e_2^{n\lambda_2} \ldots e_d^{n\lambda_d}\rangle_A$ **there are** $n\lambda_1$ **no. of states** $|e_1\rangle$**,** $n\lambda_2$ **no. of states** $|e_2\rangle$**, . . .,** $n\lambda_d$ **no. of states** $|e_d\rangle$**, and in** $|f_1^{n\lambda_1} f_2^{n\lambda_2} \ldots f_d^{n\lambda_d}\rangle_B$ **there are** $n\lambda_1$ **no. of states** $|f_1\rangle$**,** $n\lambda_2$ **no. of states** $|f_2\rangle$**, . . .,** $n\lambda_d$ **no. of states** $|f_d\rangle$**.**

Entanglement concentration scheme (continued)

- **The probability of occurrance of each such pairwise orthogonal product state is $\lambda_1^{n\lambda_1}\lambda_2^{n\lambda_2}\ldots\lambda_d^{n\lambda_d}$ and the total no. of such product states is $n!/((n\lambda_1)! \times (n\lambda_2)! \times \ldots \times (n\lambda_d)!) \approx 2^{nH(\lambda_1,\lambda_2,\ldots,\lambda_d)}$. So all of these product states when added, with equal coefficients, will form the state $|\psi(e_1^{n\lambda_1}e_2^{n\lambda_2}\ldots e_d^{n\lambda_d}; f_1^{n\lambda_1}f_2^{n\lambda_2}\ldots f_d^{n\lambda_d})\rangle_{AB}$, the later state will occur in the expansion of $|\psi\rangle_{AB}^{\otimes n}$ with probability $\{n!/((n\lambda_1)! \times (n\lambda_2)! \times \ldots \times (n\lambda_d)!)\} \times \lambda_1^{n\lambda_1}\lambda_2^{n\lambda_2}\ldots\lambda_d^{n\lambda_d} \approx 1$.**

Entanglement concentration scheme (continued)

- **Let $H(e_1^{n\lambda_1} e_2^{n\lambda_2} \ldots e_d^{n\lambda_d})$ be the $n!/((n\lambda_1)! \times (n\lambda_2)! \times \ldots \times (n\lambda_d)!)$-dimensional subspace of Alice's Hilbert space $(\mathbb{C}^d)^{\otimes n}$, spanned by the $n!/((n\lambda_1)! \times (n\lambda_2)! \times \ldots \times (n\lambda_d)!)$ no. of pairwise orthogonal states $|e_1^{n\lambda_1} e_2^{n\lambda_2} \ldots e_d^{n\lambda_d}\rangle_A$.**
- **Alice now performs the projective measurement $\{P(e_1^{n\lambda_1} e_2^{n\lambda_2} \ldots e_d^{n\lambda_d}), I_{d^n \times d^n} - P(e_1^{n\lambda_1} e_2^{n\lambda_2} \ldots e_d^{n\lambda_d})\}$ on her system, where $P(e_1^{n\lambda_1} e_2^{n\lambda_2} \ldots e_d^{n\lambda_d})$ is the projector on $H(e_1^{n\lambda_1} e_2^{n\lambda_2} \ldots e_d^{n\lambda_d})$. If $P(e_1^{n\lambda_1} e_2^{n\lambda_2} \ldots e_d^{n\lambda_d})$ clicks (which will happen with probability $\{n!/((n\lambda_1)! \times (n\lambda_2)! \times \ldots \times (n\lambda_d)!)\} \times \lambda_1^{n\lambda_1} \lambda_2^{n\lambda_2} \ldots \lambda_d^{n\lambda_d}$), the shared final state will be $|\psi(e_1^{n\lambda_1} e_2^{n\lambda_2} \ldots e_d^{n\lambda_d}; f_1^{n\lambda_1} f_2^{n\lambda_2} \ldots f_d^{n\lambda_d})\rangle_{AB}$.**

Entanglement concentration scheme (continued)

- **Representing all the $2^{nH(\lambda_1,\lambda_2,\ldots,\lambda_d)}$ (approximately) no. of terms $|e_1^{n\lambda_1}e_2^{n\lambda_2}\ldots e_d^{n\lambda_d}\rangle_A$ as $|0\rangle_{A_1}\otimes|0\rangle_{A_2}\otimes\ldots\otimes|0\rangle_{A_{nH(\lambda_1,\lambda_2,\ldots,\lambda_d)-1}}\otimes|0\rangle_{A_{nH(\lambda_1,\lambda_2,\ldots,\lambda_d)}}$, $|0\rangle_{A_1}\otimes|0\rangle_{A_2}\otimes\ldots\otimes|0\rangle_{A_{nH(\lambda_1,\lambda_2,\ldots,\lambda_d)-1}}\otimes|1\rangle_{A_{nH(\lambda_1,\lambda_2,\ldots,\lambda_d)}}$, $\ldots$, $|1\rangle_{A_1}\otimes|1\rangle_{A_2}\otimes\ldots\otimes|1\rangle_{A_{nH(\lambda_1,\lambda_2,\ldots,\lambda_d)-1}}\otimes|1\rangle_{A_{nH(\lambda_1,\lambda_2,\ldots,\lambda_d)}}$, as above, we can express $|\psi(e_1^{n\lambda_1}e_2^{n\lambda_2}\ldots e_d^{n\lambda_d};f_1^{n\lambda_1}f_2^{n\lambda_2}\ldots f_d^{n\lambda_d})\rangle_{AB}$ as $|\phi^+\rangle_{A_1B_1}\otimes|\phi^+\rangle_{A_2B_2}\otimes\ldots\otimes|\phi^+\rangle_{A_{nH(\lambda_1,\lambda_2,\ldots,\lambda_d)}B_{nH(\lambda_1,\lambda_2,\ldots,\lambda_d)nH(\lambda_1,\lambda_2,\ldots,\lambda_d)}}$.**
- **Thus we see that, with probability of success approacing unity (as $n\to\infty$), Alice and Bob will share $nH(\lambda_1,\lambda_2,\ldots,\lambda_d)$ no. of copies of the singlet state $|\phi^+\rangle_{AB}$ starting from n copies of $|\psi\rangle_{AB}$. (No classical communication is needed.)**

Topics not covered

- **Different capacities of quantum channels.**
- **Classical and quantum error correcting codes**
- **Different no-go theorems of quantum information**

References

(1) Chapters 1, 2, 8, 9, 11 and 12 of the book: *Quantum Computation and Quantum Information* **by Michael A. Nielsen and Issac L. Chuang (Cambridge Univ. Press, 2002).**

(2) Chapters 2, 3, 4 and 5 of John Preskill's Caltech lectures on Quantum Information and Computation, available at his website.