Lecture 8.

1) Accessible information & Holevo's bound.

2) Schumacher's source coding theorem.

In this chapter we prove two theorems of quantum information theory namely the Holevo bound and Schumacher's quantum source coding theorem. The Holevo bound is a general bound on the "information that can be extracted from a quantum state ρ". As we will see later it has important applications for channel coding.

Schumacher's source coding theorem is the analog of Shannon's source coding theorem. The method of proof is also via typicality.

1) Accessible information & Holevo's bound.

We know that a set of non-orthogonal quantum states is not perfectly distinguishable. The Holevo bound quantifies this statement. The setting is as follows. Suppose a system with Hilbert space $\mathcal{H}$ is prepared in a mixed state $\{p_x, \rho_x\}$ where ρ_x are density matrices. (Here the state is a "mixture of mixed states").

The total density matrix is

$$\rho = \sum_x p_x \rho_x .$$

Now suppose that Alice has prepared $\{ p_x, \rho_x \}$ and gives ρ to Bob who wants to extract some information about the preparation out of measurements performed on ρ.

<u>Let us formalize this problem.</u>

* The preparation of Alice is described by a classical random variable X taking value x with $\mathrm{Prob}(X = x) = p_x$.

<u>example:</u> Alice flips a coin. If F is obtained with $p_F = 1/2$ she prepares a photon in state $|0\rangle\langle 0|$ and if Tail is obtained with $p_T = 1/2$ is obtained she prepares a photon in state $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \frac{1}{\sqrt{2}}(\langle 0| + \langle 1|)$. Here

$$X = \begin{cases} F & ; \; p_F = 1/2 \quad \text{or} \quad x = 0 = F \\ T & ; \; p_T = 1/2 \quad \text{or} \quad x = 1 = T. \end{cases}$$

$$(p_x, \rho_x) = \begin{cases} \frac{1}{2} & ; \; |0\rangle\langle 0| \\ \frac{1}{2} & ; \; \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\frac{1}{\sqrt{2}}(\langle 0| + \langle 1|). \end{cases}$$

* Bob is given $\rho = \sum_x p_x \rho_x$ but he does not know the random variable X.

<u>in the example:</u> $\rho = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\frac{1}{\sqrt{2}}(\langle 0| + \langle 1|)$

$$\rho = \frac{3}{4}|0\rangle\langle 0| + \frac{1}{4}|1\rangle\langle 1| + \frac{1}{4}|0\rangle\langle 1| + \frac{1}{4}|1\rangle\langle 0|$$

$$= \begin{pmatrix} 3/4 & 1/4 \\ 1/4 & 1/4 \end{pmatrix}.$$

* Bob makes measurements with an apparatus corresponding to a "measurement basis" $\{P_y\}$ where $P_y^2 = P_y$ are projectors and $\sum_y P_y = \mathbb{1}$. The outcomes of a measurement is a random variable Y s.t

$$\text{Prob}\,(Y = y \mid X = x) = \text{Tr}\, \rho_x P_y .$$

by the measurement postulate [Indeed given that the state is ρ_x the prob that y is the outcome is $\text{Tr}\, \rho_x P_y$].

The joint distribution $p(x,y)$ is

$$p(x,y) = p_x\, p(y|x) = p_x \cdot \text{Tr}\, \rho_x P_y .$$

and the marginal $p(y)$ is

$$p(y) = \sum_x p_x \,\text{Tr}\, \rho_x P_y = \text{Tr}\, \rho P_y .$$

Note that the last equation could have been written directly by applying the measurement postulate to ρ.

in the example: let Bob use $\{P_y\} = \{ |0\rangle\langle 0| \,;\, |1\rangle\langle 1| \}$.

$$p(y|x) = \begin{cases} p(0|0) = 1 \\ p(1|0) = 0 \end{cases} ; \quad \begin{matrix} p(0|1) = \frac{1}{2} \\ p(1|1) = \frac{1}{2} . \end{matrix}$$

$$p(y,x) = \begin{cases} p(0,0) = \frac{1}{2} \\ p(1,0) = 0 \end{cases} \begin{matrix} ; & p(0,1) = \frac{1}{4} \\ ; & p(1,1) = \frac{1}{4} . \end{matrix}$$

$$p(y) = \begin{cases} p(0) = 3/4 \\ p(1) = 1/4 . \end{cases}$$

* The mutual information $I(X;Y)$ is the information that Bob can extract from X by looking at his measurement results Y. We define the accessible information as the maximum possible mutual information obtained by the best possible measurement

$$Acc(\rho) \triangleq \sup_{\{P_y\}} I(X;Y).$$

where $I(X;Y) = H(X) + H(Y) - H(X,Y)$
$$= H(X) - H(X|Y) = H(Y) - H(Y|X).$$

in the example: with the particular measurement $P_y = \{|0\rangle\langle 0| ; |1\rangle\langle 1|\}$

we have $H(X) = -\frac{1}{2}\ln\frac{1}{2} - \frac{1}{2}\ln\frac{1}{2} = \ln 2.$

$$H(Y) = -\frac{3}{4}\ln\frac{3}{4} - \frac{1}{4}\ln\frac{1}{4} = \ln 4 - \frac{3}{4}\ln 3$$

$$H(X;Y) = -\frac{1}{2}\ln\frac{1}{2} - \frac{1}{4}\ln\frac{1}{4} - \frac{1}{4}\ln\frac{1}{4} = \frac{1}{2}\ln 2 + \frac{1}{2}\ln 4 = \frac{3}{2}\ln 2.$$

so $I(X;Y) = \ln 2 + \ln 4 - \frac{3}{4}\ln 3 - \frac{3}{2}\ln 2 = \frac{3}{2}\ln 2 - \frac{3}{4}\ln 3.$

$$I(X;Y) = 0,215 = (0,31)\times\ln 2 = \underline{0,31 \text{ bits}}.$$

If the states where classical we would always have (with perfect measurements) $Y = X$ so that $I(X;Y) = H(X) = \ln 2 = 1$ bit.

So Bob does not have access to all the info about the preparation.

The Holevo bound tells that whatever the measurement that Bob does his accessible information is bounded by

$S(\rho) - \sum_x p_x S(\rho_x)$. Thus because of the bound on ρ's in lect 7 it is less than $H(X)$. [Classical we would have $\sup_{\{P_y\}} I(X;Y) = H(X)$].

Theorem [Holevo's bound].

let X a classical r.v $\{p_x, X = x\}$ and $\{p_x, \rho_x\}$ the preparation of a mixed quantum state. let Y the r.v describing measurement outcomes in some basis $\{P_y\}$. Then

$$I(X;Y) \leq \chi(\rho).$$

where the "Holevo-χ" is defined by

$$\chi(\rho) = S(\rho) - \sum_x p_x S(\rho_x).$$

In the example:

the states ρ_x are pure so $S(\rho_x) = 0$. So $\chi(\rho) = S(\rho)$ and $S(\rho) = -\operatorname{Tr} \rho \ln \rho$. The eigenvalue of $\rho = \begin{pmatrix} 3/4 & 1/4 \\ 1/4 & 1/4 \end{pmatrix}$ are $\lambda_\pm = \frac{1}{2} \pm \frac{\sqrt{2}}{4}$. So the Von-Neuman entropy is

$$S(\rho) = -\left(\frac{1}{2} + \frac{\sqrt{2}}{4}\right) \ln\left(\frac{1}{2} + \frac{\sqrt{2}}{4}\right) - \left(\frac{1}{2} - \frac{\sqrt{2}}{4}\right) \ln\left(\frac{1}{2} - \frac{\sqrt{2}}{4}\right) = 0{,}1351 + 0{,}2813 = 0{,}41$$

$$= 0{,}59 \times \ln 2 = 0{,}59 \text{ bits}.$$

We remark that with the particular measurement basis

$\{ |0\rangle\langle 0| \ ; \ |1\rangle\langle 1| \}$ we do not attain 0,59 bits $= \chi(\rho)$.

Exercice: Compute $Acc(\rho) = \sup\limits_{\{P_y\}} I(X;Y)$ by

taking a general measurement basis

$$\{P_y\} = \{ (\alpha|0\rangle + \beta|1\rangle)(\bar{\alpha}\langle 0| + \bar{\beta}\langle 1|) \ ;$$
$$(\gamma|0\rangle + \delta|1\rangle)(\bar{\gamma}\langle 0| + \bar{\delta}\langle 1|) \} .$$

where $\alpha\bar{\gamma} + \beta\bar{\delta} = 0 \quad ; \quad |\alpha|^2 + |\beta|^2 = 1 \quad ; \quad |\gamma|^2 + |\delta|^2 = 1.$

Is it possible to achieve Holevo's bound ?.

Achievability of Holevo's bound.

* In general Holevo's bound is not attained. It is not known what would be a sharp bound to $Acc(\rho)$.

* It is possible to achieve the bound in an assymptotic sense for ensembles of the type $\rho^{(n)}_{x_1 \dots x_n} = \rho_{x_1} \otimes \rho_{x_2} \dots \otimes \rho_{x_n}$ with probabilities

$$p^{(n)}_{x_1 \dots x_n} = p_{x_1} p_{x_2} \dots p_{x_n}$$

Here $\rho^{(n)}_{typ} = \sum\limits_{\substack{x_1 \dots x_n \\ \text{typical sequences}}} p_{x_1} p_{x_2} \dots p_{x_n} \ \rho_{x_1} \otimes \rho_{x_2} \dots \otimes \rho_{x_n}$.

This will be of importance in the Noisy Channel coding Theorem. (next chapter).

Proof of Holevo bound.

The proof uses strong subadditivity or conditioning.

Bob is given $\rho = \sum_x p_x \rho_x$. This is a mixed state in the Hilbert space $\mathcal{H}_Q$. Here Q is just an index serving to designate the system Q.

We introduce a larger Hilbert space $\mathcal{H}_X \otimes \mathcal{H}_Q \otimes \mathcal{H}_Y$ and a state

$$\rho_{XQY} = \sum_x p_x \, |x\rangle\langle x| \otimes \rho_x \otimes |0\rangle\langle 0|$$

The interpretation of ρ_{XQY} is as follows: $|x\rangle\langle x|$ are the mutually orthogonal states describing Alice's preparation (or r.v X) and $|0\rangle\langle 0|$ is a "blank state" where Bob will record his measurement outcomes. Note that

$\dim \mathcal{H}_X$ = # of values of x (say D).

$\dim \mathcal{H}_Q$ = dim of Hilbert space of state given to Bob. (e.g. 2 if this a qbit).

$\dim \mathcal{H}_Y = \dim \mathcal{H}_Q$ since $\mathcal{H}_Y$ records measurement outcomes.

The measurement basis of Bob is $\{ \Pi_y \}$ and his measurement results are described by mutually orthogonal states $|y\rangle\langle y|$.
(Essentially $\Pi_y = |y\rangle\langle y|$ and we can make copies of these states)

We introduce the unitary operation:

$$U_{XQY} = \mathbb{1} \otimes U_{QY} \quad ,$$

$$U_{QY} : \quad |\varphi\rangle_Q \otimes |a\rangle_Y \longmapsto U_{QY} |\varphi\rangle_Q \otimes |a\rangle_Y$$

$$= \sum_y P_y |\varphi\rangle_Q \otimes |y\rangle_Y .$$

It is easy to check it is unitary. Indeed,

$$\langle\psi|\otimes\langle b| U^{\dagger}_{QY} \; U_{QY} \; |\varphi\rangle \otimes |a\rangle = \sum_{y,y'} \left(\langle\psi| P^{\dagger}_{y'} \otimes \langle y'|\right)\left(|y\rangle \otimes P_y |\varphi\rangle\right)$$

$$= \sum_{y,y'} \underbrace{\langle\psi| P^{\dagger}_{y'} P_y |\varphi\rangle}_{\langle\psi| P_y |\varphi\rangle} \; \underbrace{\langle y'|y\rangle}_{\delta_{yy'}}$$

$$= \sum_y \quad \langle\psi| P_y |\varphi\rangle$$

$$= \langle\psi| \sum_y P_y |\varphi\rangle = \langle\psi|\varphi\rangle .$$

Thus $\mathbb{1} \otimes U_{QY}$ preserves the scalar product, thus it is unitary.

Now

$$U_{XQY} \cdot \rho_{XQY} \cdot U^{\dagger}_{XQY} = \sum_{x,y,y'} p_x \, |x\rangle\langle x| \otimes P_y \rho_x P_{y'} \otimes |y\rangle\langle y'| .$$

$$= \rho'_{XQY} .$$

Now ρ'_{XQY} and ρ_{XQY} have the same eigenvalues since they are unitarily related. Thus their Von-Neumann entropies are the same

$$S(\rho_{XQY}) = S(\rho'_{XQY}).$$

Since U_{XQY} acts as the identity over $\mathcal{H}_X$ we have that the partial density matrices

$$\rho_{QY} = \mathrm{Tr}_X\, \rho_{XQY} = \sum_x p_x\, \rho_x \otimes |0\rangle\langle 0| = \rho \otimes |0\rangle\langle 0|.$$

and

$$\rho'_{QY} = \mathrm{Tr}_X\, \rho'_{XQY} = \sum_x p_x\, P_y\, \rho_x\, P_y \otimes |y\rangle\langle y'|.$$

are also unitarily related (in fact $U_{QY}\, \rho_{QY}\, U^{+}_{QY} = \rho'_{QY}$). Thus

$$S(\rho_{QY}) = S(\rho'_{QY}).$$

Now we apply (conditioning) or strong subadditivity to ρ'_{XQY} :

$$S'(X|QY) \leq S'(X|Y) \qquad \text{i.e}$$

$$S(\rho'_{XQY}) - S(\rho'_{QY}) \leq S(\rho'_{XY}) - S(\rho'_{Y}).$$

Thus we get (using the unitary relations.) :

$$S(\rho_{XQY}) - S(\rho_{QY}) \leq S(\rho'_{XY}) - S(\rho'_Y).$$

Now we compute all these entropies:

* $S(\rho_{XQY}) = S\left(\sum_x p_x |x\rangle\langle x| \otimes \rho_x\right)$

$$= H(X) + \sum_x p_x S(\rho_x).$$

[Check of this last relation: use $\rho_x = \sum_{a_x} \lambda_{a_x} |a_x\rangle\langle a_x|$ spectral decomposition. Then

$$\sum_x p_x |x\rangle\langle x| \otimes \rho_x = \sum_{x, a_x} p_x \lambda_{a_x} \left(|x\rangle\langle x| \otimes |a_x\rangle\langle a_x|\right)$$

Since this a superposition of mutually orthogonal states we have

$$S = -\sum_{x, a_x} p_x \lambda_{a_x} \ln p_x \lambda_{a_x} = H(X) - \sum_x p_x \sum_{a_x} \lambda_{a_x} \ln \lambda_{a_x}$$

$$= H(X) - \sum_x p_x S(\rho_x). \quad].$$

* $S(\rho_{QY}) = S(\rho)$. Indeed $\rho_{QY} = \rho \otimes |0\rangle\langle 0|$.

* $S(\rho'_{XY}) = ?$

First $\rho'_{XY} = \mathrm{Tr}_Q\, \rho'_{XQY} = \sum_{x,y,y'} p_x |x\rangle\langle x| \otimes |y\rangle\langle y'|$

$$\cdot \mathrm{Tr}_Q\, P_y \rho_x P_{y'}$$

by cyclicity of the trace:

$$\mathrm{Tr}_Q\, P_y\, \rho_x\, P_{y'} = \mathrm{Tr}_Q\, P_{y'} P_y\, \rho_x = \delta_{yy'}\, \mathrm{Tr}_Q\, P_y\, \rho_x$$
$$= \delta_{yy'}\, p(y|x).$$

Thus $\rho'_{XY} = \sum_{x,y} \underbrace{p_x \cdot p(y|x)}_{p(x,y)} \, |x\rangle\langle x| \otimes |y\rangle\langle y|.$

This density matrix is just another representation for the classical r.v (X,Y). Indeed the states $|x\rangle\langle x| \otimes |y\rangle\langle y|$ are mutually orthogonal (so distinguishable). The Von-Neumann entropy is

$$S(\rho'_{XY}) = H(X,Y).$$

* $S(\rho'_Y) = ?$

First $\rho'_Y = \mathrm{Tr}_X\, \rho'_{XY} = \sum_{x,y} p(x,y)\, |y\rangle\langle y|$

$$= \sum_y p(y)\, |y\rangle\langle y|.$$

Thus $S(\rho'_Y) = H(Y)$.

Putting all these results back in strong subadditivity inequality:

$$H(X) + \sum_x p_x S(\rho_x) - S(\rho) \leq H(X,Y) - H(Y).$$

$$\Rightarrow \quad I(X;Y) \leq S(\rho) - \sum_x p_x S(\rho_x) = \chi(\rho).$$