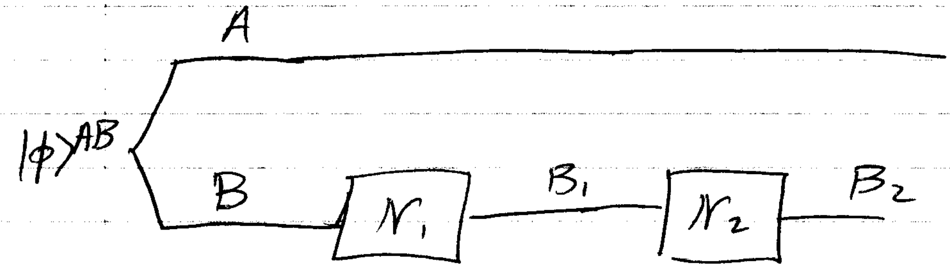
3/3/2011

Lecture 16

Quantum Data Processing Inequality

Alice & Bob share $|\phi\rangle^{AB}$



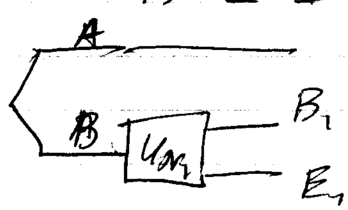
Quantify quantum correlations by coherent information $I(A\rangle B)_\phi$

Q. Data Processing: each step reduces quantum correlations

$$I(A\rangle B) \geq I(A\rangle B_1) \geq I(A\rangle B_2)$$

Proof:

1st part



$$I(A\rangle B) = H(B) - H(AB)$$
$$= H(B)$$
$$= H(A)$$

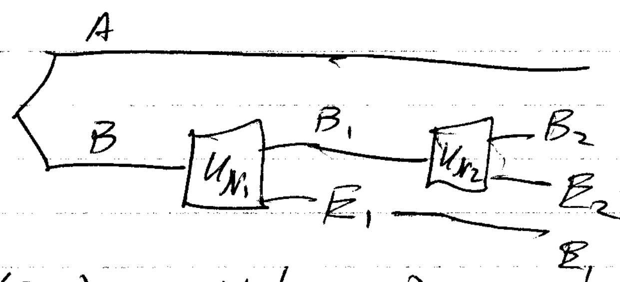
3/3/2011

$$I(A \rangle B_1) = H(B_1) - H(AB_1)$$
$$= H(AE_1) - H(E_1)$$
$$= H(A|E_1)$$

Finally, $H(A) \geq H(A|E_1)$

$\therefore \quad I(A \rangle B) \geq I(A \rangle B_1)$

Next part,



$$I(A \rangle B_2) = H(B_2) - H(AB_2)$$
$$= H(AE_2E_1) - H(E_2E_1)$$
$$= H(A|E_2E_1)$$

SSA

Holds that $H(A|E_1) \geq H(A|E_2E_1)$

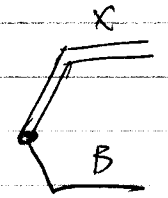
$\therefore \quad I(A \rangle B_1) \geq I(A \rangle B_2)$

can prove also that

$$I(A;B) \geq I(A;B_1) \geq I(A;B_2)$$

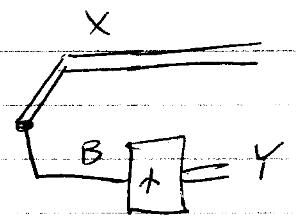
Holevo bound follows from QDP

Suppose



$$\sum_x p(x)|x\rangle\langle x|^X \otimes \rho_x^B$$

Bob can perform some measurement $\{\Lambda_y\}$ to learn about X



resulting state is

$$\sum_{x,y} p(x)\, |x\rangle\langle x|^X \otimes \sqrt{\Lambda_y}\rho_x^B\sqrt{\Lambda_y} \otimes |y\rangle\langle y|^Y$$ and he

traces out B, giving

$$\sum_{x,y} p(x)|x\rangle\langle x| \otimes \mathrm{Tr}\{\Lambda_y \rho_x\}\, |y\rangle\langle y|^Y$$

accessible info is $I(X;Y)$ optimized over POVMs

3/3/2011

By QDP,

$$I(X;B) \geq \max_{\{\Lambda_y\}} I(X;Y)$$

Continuity of Quantum Entropy

For any ρ^{AB} & σ^{AB} where

$$\|\rho^{AB} - \sigma^{AB}\|_1 \leq \epsilon \quad ,$$

$$|H(A|B)_\rho - H(A|B)_\sigma| \leq 4\epsilon \log d_A + 2 H_2(\epsilon)$$

Proof: suppose $\|\rho^{AB} - \sigma^{AB}\|_1 = \epsilon$ & $\epsilon < 1/2$

Define Density operators

$$\tilde{\rho}^{AB} \equiv \frac{1}{\epsilon} |\rho^{AB} - \sigma^{AB}|$$

$$\tilde{\sigma}^{AB} \equiv \frac{1-\epsilon}{\epsilon}(\rho^{AB} - \sigma^{AB}) + \tilde{\rho}^{AB}$$

Define $\gamma^{XAB} \equiv (1-\epsilon)\, |0\rangle\langle 0|^X \otimes \rho^{AB} +$
$\epsilon\, |1\rangle\langle 1|^X \otimes \tilde{\rho}^{AB}$

Observe that $\gamma^{AB} = (1-\epsilon)\, \rho^{AB} + \epsilon \tilde{\rho}^{AB}$

Also $\gamma^{AB} = (1-\epsilon)\, \sigma^{AB} + \epsilon\, \tilde{\sigma}^{AB}$

$[= (1-\epsilon)\, \sigma^{AB} + \epsilon(1-\epsilon)(\rho^{AB} - \sigma^{AB}) + \epsilon \tilde{\rho}^{AB}]$

3/3/2011

Now prove that

$$\left| H(A|B)_\rho - H(A|B)_\gamma \right| \le 2\epsilon \log d_A + H_2(\epsilon)$$

First, do

$$H(A|B)_\rho - H(A|B)_\gamma \le 2\epsilon \log d_A + H_2(\epsilon)$$

We know that

$$H(A|B)_\gamma \ge H(A|BX)_\gamma$$

$$= (1-\epsilon)\, H(A|B)_\rho + \epsilon\, H(A|B)_{\tilde{\rho}}$$

$\therefore$ $H(A|B)_\rho - H(A|B)_\gamma \le$

$$H(A|B)_\rho - \left[(1-\epsilon)\, H(A|B)_\rho + \epsilon\, H(A|B)_{\tilde{\rho}}\right]$$

$$= \epsilon \left[H(A|B)_\rho - H(A|B)_{\tilde{\rho}} \right]$$

$$\le \epsilon\, 2 \log d_A$$

$$\le \epsilon\, 2 \log d_A + H_2(\epsilon)$$

3/3/2011

Now prove other way

$$H(A|B)_\gamma - H(A|B)_\rho \le 2\epsilon \log d_A + H_2(\epsilon)$$

We know

$$H(B)_\gamma \ge H(B|X)_\gamma$$
$$= (1-\epsilon) H(B)_\rho + \epsilon H(B)_{\tilde{\rho}}$$

Also know that

$$H(X|AB)_\gamma \ge 0$$ whenever X is classical

so,

$$H(AB)_\gamma \le H(ABX)_\gamma$$
$$= H(AB|X)_\gamma + H(X)_\gamma$$
$$= (1-\epsilon) H(AB)_\rho + \epsilon H(AB)_{\tilde{\rho}} + H_2(\epsilon)$$

Combining both, we get

$$H(A|B)_\gamma \le (1-\epsilon) H(A|B)_\rho + \epsilon H(A|B)_{\tilde{\rho}} + H_2(\epsilon)$$

$$\therefore\ H(A|B)_\gamma - H(A|B)_\rho \le \epsilon \{H(A|B)_{\tilde{\rho}} - H(A|B)_\rho\} + H_2(\epsilon)$$
$$\le \epsilon 2 \log d_A + H_2(\epsilon)$$

3/3/2011

By same development, can get

$$|H(A|B)_\sigma - H(A|B)_\gamma| \leq 2\epsilon \log d_A + H_2(\epsilon)$$

triangle gives

$$|H(A|B)_\rho - H(A|B)_\sigma| \leq 4\epsilon \log d_A + H_2(\epsilon)$$

Classical Typicality

Suppose information source randomly emitting sequences $x^n \equiv x_1 x_2 \cdots x_n$ according to some IID distribution $p_{X^n}(x^n) = p_X(x_1)\, p_X(x_2) \cdots p_X(x_n)$

would expect that source is becoming convergent when n is large, so that

$$p_{X^n}(x^n) \approx p_X(a_1)^{n p_X(a_1)} \cdots p_X(a_{|X|})^{n p_X(a_{|X|})}$$

3/3/2011

sample entropy of sequence x^n is

$$\overline{H}(x^n) \equiv -\frac{\log p_{X^n}(x^n)}{n} \approx -\sum_{i=1}^{|X|} p_X(a_i) \log p_X(a_i) = H(X)$$

Define typical set as the set of all sequences w/ sample entropy close to true entropy:

$$T_\delta^{X^n} \equiv \{x^n : |\overline{H}(x^n) - H(X)| \leq \delta\}$$

Typical Set has three properties:

1) unit probability

$$\forall \epsilon > 0 \quad \Pr\{X^n \in T_\delta^{X^n}\} \geq 1-\epsilon$$

can write as

$$\Pr\{X^n \in T_\delta^{X^n}\} = \sum_{x^n \in T_\delta^{X^n}} p_{X^n}(x^n) \geq 1-\epsilon$$

LLN

2) Exponentially small Cardinality

$$|\mathcal{T}_\delta^{X^n}| \le 2^{n\{H(X)+\delta\}}$$

& $\forall \epsilon > 0 \quad |\mathcal{T}_\delta^{X^n}| \ge (1-\epsilon)\, 2^{n\{H(X)-\delta\}}$

& sufficiently large n

Proof:
$$\begin{aligned}
1 &= \sum_{x^n \in \mathcal{X}^n} P_{X^n}(x^n) \\
&\ge \sum_{x^n \in \mathcal{T}_\delta^{X^n}} P_{X^n}(x^n) \\
&\ge \sum_{x^n \in \mathcal{T}_\delta^{X^n}} 2^{-n\{H(X)+\delta\}} \\
&= |\mathcal{T}_\delta^{X^n}|\, 2^{-n\{H(X)+\delta\}}
\end{aligned}$$

Other way

$$\begin{aligned}
1-\epsilon &\le \Pr\{X^n \in \mathcal{T}_\delta^{X^n}\} \\
&= \sum_{x^n \in \mathcal{T}_\delta^{X^n}} P_{X^n}(x^n) \\
&\le \sum_{x^n \in \mathcal{T}_\delta^{X^n}} 2^{-n\{H(X)-\delta\}} \\
&= |\mathcal{T}_\delta^{X^n}|\, 2^{-n\{H(X)-\delta\}}
\end{aligned}$$

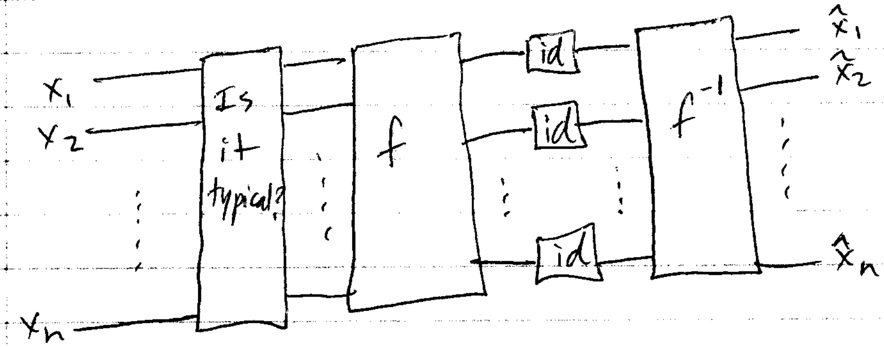
3/3/2011

3) Equipartition:

Probabilities of sequences in typicalset are approximately uniform

$$2^{-n[H(X)+\delta]} \le p_{X^n}(x^n) \le 2^{-n[H(X)-\delta]}$$

Shannon Compression



First ask: is sequence x^n typical?

w/ prob. $> 1-\epsilon$ yes.

if so,

$$f : T_\delta^{X^n} \to \{0,1\}^{nR}$$

where $R = H(X) + \delta$

transmit bits, Bob performs inverse + decodes

Schumacher compression will work very similarly

Other Kind of Typicality

Strong Typicality

Example: Suppose IID dist. is
$p(0) = 1/4$ & $p(1) = 3/4$ &
source generates

$x^n = 011\ 0111\ 010$

compare empirical distribution to true distribution

$N(0|x^n) = 4$ $\quad N(1|x^n) = 6$

empirical dist. is

$$\frac{N(0|x^n)}{n} = \frac{4}{10} = \frac{2}{5} \qquad \frac{N(1|x^n)}{n} = \frac{6}{10} = \frac{3}{5}$$

deviation from true distribution is

$$\max \left\{ \left|\frac{1}{4} - \frac{2}{5}\right|, \left|\frac{3}{4} - \frac{3}{5}\right| \right\} = \frac{3}{20}$$

Define Strongly Typical Set in this way

$$T_\delta^{X^n} \equiv \left\{ x^n : \forall x \in \mathcal{X} \left| \frac{N(x|x^n)}{n} - p_X(x) \right| \leq \delta \right\}$$

implies weak typicality (AKA entropic typicality)