- Shannon's noisy channel coding theorem (classical):

Given a classical channel: $x = \boxed{N} =$ output $= y$ wp $p(y|x)$

use n times iid : $\forall x_1, \ldots, x_n \in \Sigma_{in}^{\otimes n}$

$x_1 = \boxed{N} =$ output $= y_1$ wp $p(y_1|x_1)$

$x_2 = \boxed{N} = \;\; - - - \;\; y_2 \quad p(y_2|x_2)$

$\vdots \qquad\qquad\qquad \vdots$

$x_n = \boxed{N} = \qquad y_n \quad p(y_n|x_n)$

↑ can choose from a code book, arbitrary correlation between the x_i's

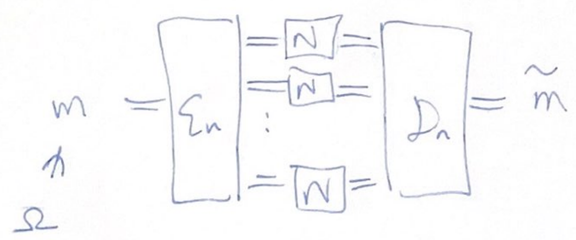
so output $= y_1 y_2 \cdots y_n$

wp $p(y_1|x_1) \cdot p(y_2|x_2) \cdots p(y_n|x_n)$

"memoryless" : eg $p(y_n | x_1 \cdots x_n) = p(y_n|x_n)$

- For any distribution $p(x)$ on Σ_{in}, let $p(xy) = p(y|x) \cdot p(x)$.

$\forall n$, there is an encoder and decoder : (a mathematical step used to define a family of ECC's)



s.t $Pr(m \neq \tilde{m}) \to 0$ as $n \to \infty$

$|\Omega| \sim 2^{n I(X:Y)}$, $I(X:Y)$ evaluated on $p(xy)$.

ie $\approx I(X:Y)$ bits can be sent <u>per use</u> of N.

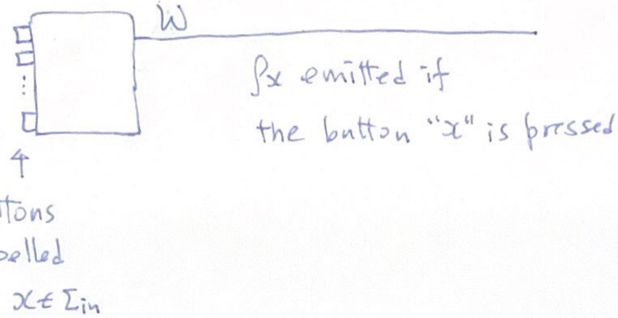
- Capacity of $N = \max_{p(x)} I(X:Y)$ (proof see Corr & Thomas or QIC 890 / CO781 / CS867 F2020)

- Holevo 73: What if the channel is quantum?
  (and Alice wants to send classical data to Bob)

Answer is very complex....

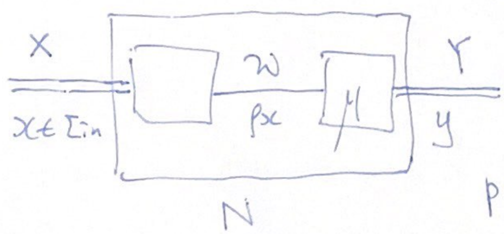
Let's ask some simpler questions.

Q1. Consider a quantum device, called a "C-Q channel" (input: C, output: Q) (or Q-box in QIC 890)



Note this is a big restriction!

(*) If X is measured according to some measurement M, with POVM $\{M_y\} \subseteq \mathrm{Pos}(W)$, $\sum_y M_y = \mathbb{I}_W$, we obtain a classical channel



$$\mathrm{prob}(y|x) = \mathrm{tr}(M_y \rho_x)$$

Add another restriction

(*) If we fix a distribution $p(x)$ for the input but allow M to be optimized,

that is, for the ensemble $\mathcal{E} = \{ p(x), \rho_x \}$, or

$$\Lambda = \sum_{x \in \Sigma_{in}} p(x) \; |x\rangle\langle x|_X \otimes \rho_{x\,W}$$

define the accessible info $I_{acc}(\mathcal{E})$ as:

$$\max_{M} I(X:Y)_{I \otimes M(\Lambda)}$$

Remarks:

① the optimal M can be attained!

In fact, the # measurement outcomes $\leq (\dim(X))^2$, all M_y rank 1.

Pf is by E Davis (78) Convexity, Caratheodory's Thm ++

② Unlike state discrimination, optimization is quite difficult, known only for very simple ensembles with symmetry

objective function is NOT linear in the variable $\{M_y\}$

③ Note that I_{acc} is defined for a fixed distn $p(x)$.

If we also max over $p(x)$:

$$\max_{\{p(x)\}} \max_{M} I(X:Y)_{I \otimes M(\Lambda)}$$

we have the capacity of the "Q-box + individual meas" combo.

Q2. For any ensemble $\Lambda = \sum_{x \in I_{in}} p(x)\, |x\rangle\langle x|_X \otimes \rho_{x\,W}$

(so the dist^n $p(x)$ is fixed and given)

The Holevo info of the ensemble is

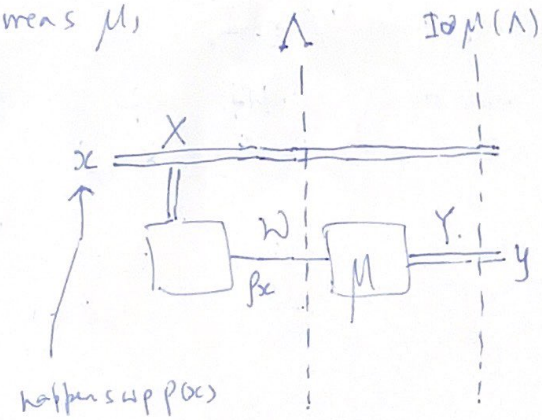
$$S(X:W)_\Lambda = S(X) + S(W) - S(XW)$$

$$= H(p) + S(\underbrace{\textstyle\sum_x p(x)\rho_x}_{\text{average state}}) - [\underbrace{\textstyle\sum_x p(x) S(\rho_x) + H(p)}_{S(XW) \text{ from Thm NCOO}}]$$

$$= S(\textstyle\sum_x p(x)\rho_x) - \sum_x p(x) S(\rho_x)$$

Thm 12.1 (Holevo's theorem): for any ensemble, the accessible info is upper-bounded by the Holevo info!

Pf: for any meas μ,

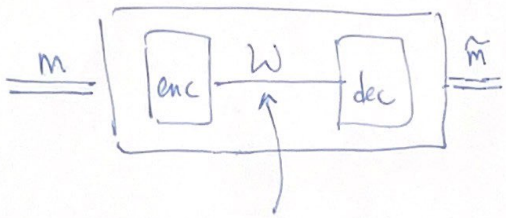


• By mono of QMI:

$$\forall \mu.\quad S(X:W)_\Lambda \geq I(X:Y)_{I\otimes\mu(\Lambda)}$$

$$\therefore\ S(X:W)_\Lambda \geq \max_\mu I(X:Y)_{I\otimes\mu(\Lambda)} = I_{acc}(\{p(x), \rho_x\})$$

Q3. What about unlimited comm strategy via noiseless Q channels?



some Q sys has to be sent, call it W

Thm 12.3 (Fanos inequality):

Let A, B be r.v's over a common sample space Ω, with joint distⁿ $p(ab)$.

Let $q = \text{prob}(A \neq B) = \sum_a \sum_{b \neq a} p(ab)$.

Then $H(A|B) \le h(q) + q \log(|\Omega| - 1)$,

where $h(q) = -q\log q - (1-q)\log(1-q)$ is the binary entropy function.

rv has outcome 0, 1 wp q, $1-q$.

Pf: define an error rv $E = \begin{cases} 1 & \text{if } A \neq B \\ 0 & \text{otherwise.} \end{cases}$

Note $H(EA|B) = \underbrace{H(AB) - H(B)} + \underbrace{H(EAB) - H(AB)}$

$= H(A|B) + \underbrace{H(E|AB)}_{=\,0} = \underbrace{H(E|B)}_{\le H(E) \,=\, h(q)} + \underbrace{H(A|EB)}_{=\,(1-q)H(A|B\,E=0) + q\,H(A|B\,E=1)}$

SWAP A, E

Conditioning reduces entropy (true classically)

$q\,H(A|B\,E=1) \le q \log(|\Omega| - 1)$

Applying Fano's ineq to Alice's message & Bob's estimate,
we want $\Pr(m \neq \tilde{m})$ to vanish.

$\therefore$ Let A be r.v with outcome m,
B - - - - $\tilde{m}$.

$$H(A|B) \le h(q) + q \log(|\Omega| - 1) \to 0 \text{ as } q \to 0.$$

$$\therefore I(A:B) = H(A) - H(A|B) \ge H(A) - h(q) - q\log(|\Omega| - 1)$$

$\uparrow$ log(# message) with classical data compression and with $q \to 0$

Quantum

$$I(A:B) \overset{\text{mono/Holevo}}{\le} S(A:W) \le S(W) \le \log(\dim W)$$

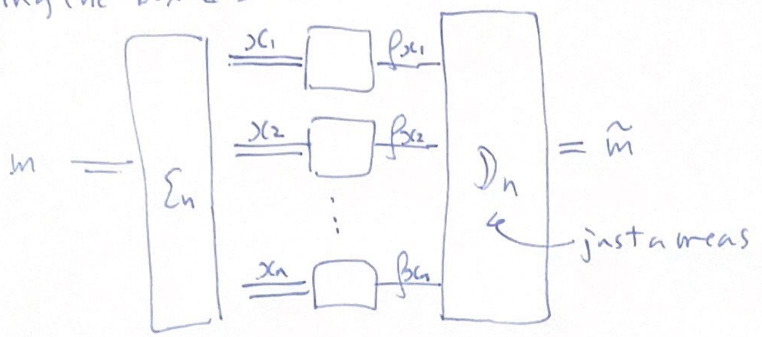
$\therefore$ Can at best send n bits with n qubits of noiseless comm.

For classical - Quantum states:

$$S(A:W) = S\left(\sum_m p_m \rho_m\right) - \sum_m p_m S(\rho_m) \le S(W).$$

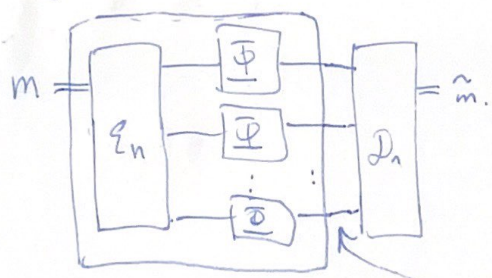
Q4 finally, given n uses of noisy Q channel Φ, how many bits can be sent?

(a) using the box described in Q1:



If joint meas on all n outputs are allowed, then for any dist^n $p(x)$, $\exists\, \mathcal{E}_n, \mathcal{D}_n$ for transmitting the Holevo info for ensemble $\{p(x), \rho_x\}$ per use of the box.

(b) Using channel $\Phi \in C(\mathcal{X}, \mathcal{Y})$



this is a box like the one describe in Q1 !!

(See Watrous bk or QIC 890 F2020)

(The HSW Thm)

Holevo-Schumacher-Westmoreland

a giant $\rho_m = \Phi^{\otimes n}(\xi_m)$, $\xi_n \in D(\mathcal{X}^{\otimes n})$

so $\frac{1}{n}$ Holevo info of $\left(\{p_m, \Phi^{\otimes n}(\xi_m)\} \right)$ is achievable

and also optimal....

sup n sup p_m, ξ_m

but need not be the same without sup n