CO781 / QIC 890:

Theory of Quantum Communication

Topics 4, part 5

Encoding classical information in quantum states and retrieving it

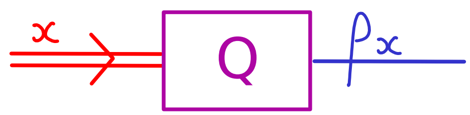
Scenario 4: classical capacity of quantum channels

LAST TIME

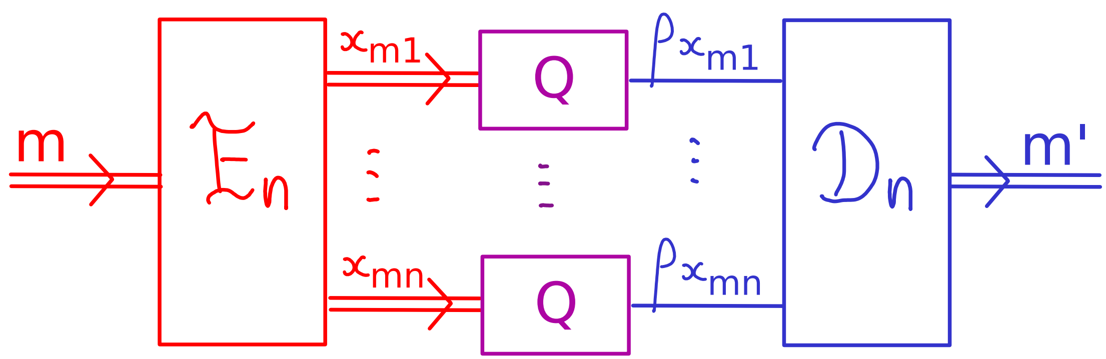
<u>Definition</u>:

A Q-box is specified by $\{\rho_x\}_{x \in \Omega}$

If Alice inputs x, then, Bob gets ρ_x :



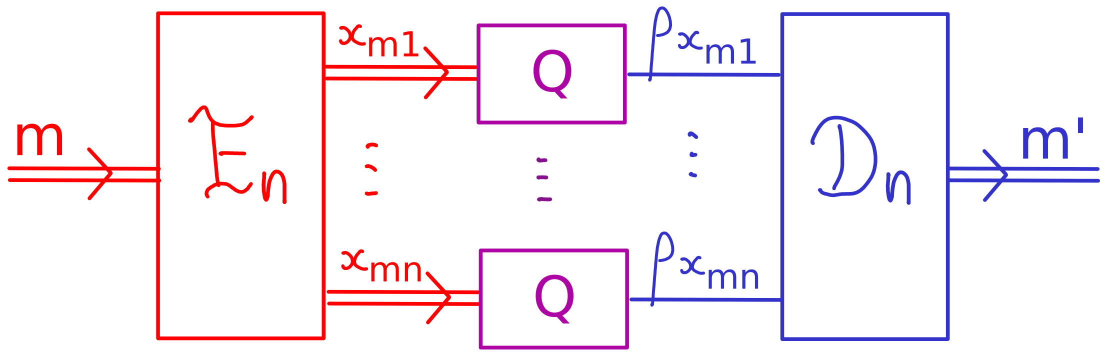
<u>Most general communication protocol using Q-boxes n times</u>:



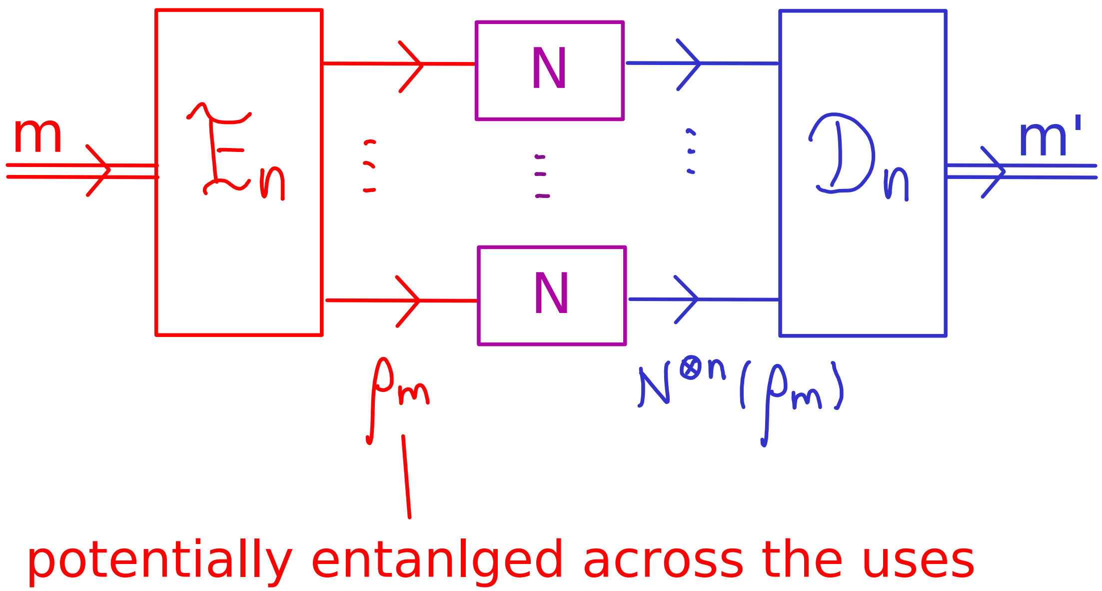
<u>Theorem</u>: Capacity of Q-box, $C(Q) = \max_{p(x)} \chi(\{p_x, \rho_x\})$

TODAY: classical capacity of quantum channels (the HSW theorem)

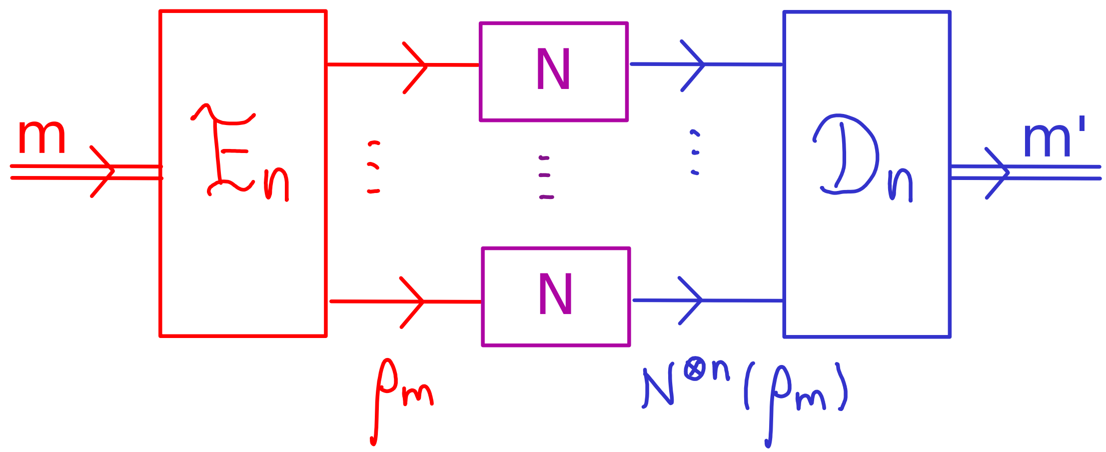
Most general communication protocol using Q-boxes n times:



Most general comm protocol using a quantum channel n times:



Most general comm protocol using a quantum channel n times:



1. Alice's message is m

2. She looks up code book to find the q input ρ_m for n uses of N — picked & fixed for each m

3. She enters the input

4. Bob gets the output q systems in the state $N^{\otimes n}(\rho_m)$

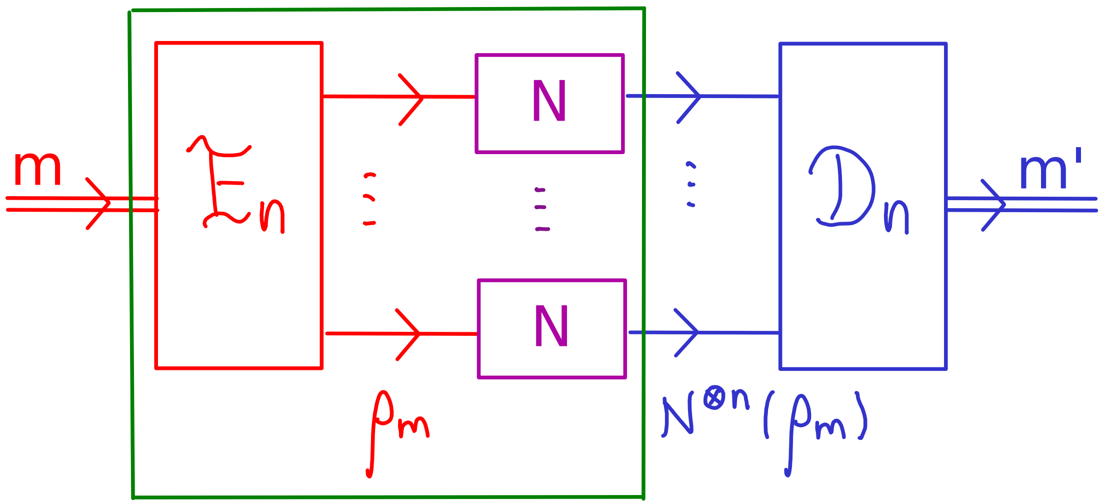
5. He applies a measurement $\mathcal{D}_n$ that outputs m' — jointly for optimality

Concepts as defined before:

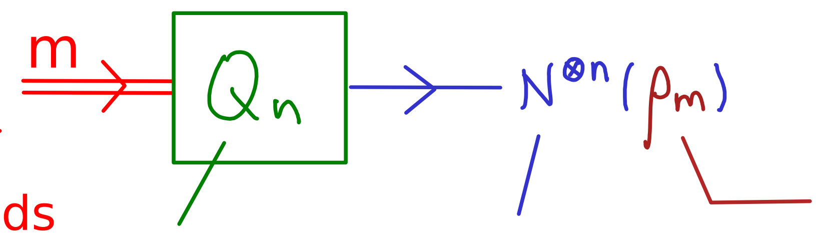
for each channel: (M,n) codes
for each code: error for each message, average error, worse case error
for each channel: achievable rates, capacity

Relating the capacity of quantum channels to that of Q-boxes:

Most general comm protocol using a quantum channel n times:



Identify a Q-box (a very big one) in the above n-use protocol:



range depends on R and n

depends on n

no choice

reduces distinguishability of ρ_m

determined by the code which Alice & Bob agree to and can optimize later

Useful result from last lecture $C(Q_n) = \max_{p_m} \chi(\{p_m, N^{\otimes n}(\rho_m)\})$

Will show that:

$\frac{1}{n}$ capacity of Qn, optimized over n and ρ_m = capacity of N

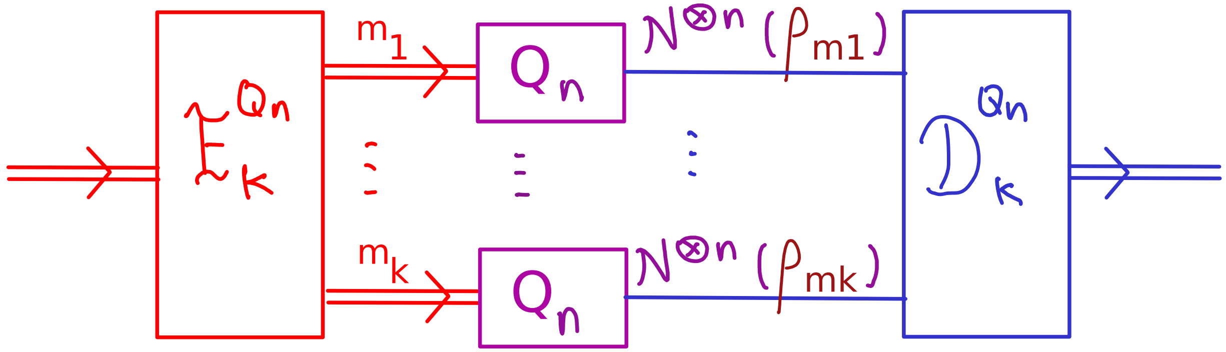
Will do so in 2 (simple) steps:

(1) LHS is achievable rate for N (direct coding thm for N based on Qn)

(2) Converse - that no rate above the LHS is achievable

Use Qn k times for large k:

(1) DIRECT CODING THM
use Qn as a big Q-box



$\approx k\, C(Q_n)$ bits comm with vanishing error, using N nk times

So, $\frac{1}{n} C(Q_n)$ is an achievable rate for N for any $\{\rho_m\}$ and any n

$$\therefore\; C(N) \geq \sup_n \max_{\{\rho_m\}} \frac{1}{n} C(Q_n)$$

↖ inputs to n uses of N

$$= \sup_n \max_{\{\rho_m\}} \frac{1}{n} \max_{p_m} \chi\left(\{p_m, N^{\otimes n}(\rho_m)\}\right)$$

$$= \sup_n \left[\frac{1}{n} \max_{\{p_m, \rho_m\}} \chi\left(\{p_m, N^{\otimes n}(\rho_m)\}\right)\right]$$

We next show the above IS an also an upper bound to any achievable rate, thus it is the capacity ...

What is the code?

- If you know $C(N)$, to achieve rate $C(N) - \delta$

$\exists r$, r-use ensemble $\{\rho_x, p_x\}$ with $\frac{1}{r}\chi(\{p_x, N^{\otimes r}(\rho_x)\}) \geq C(N) - \frac{\delta}{2}$

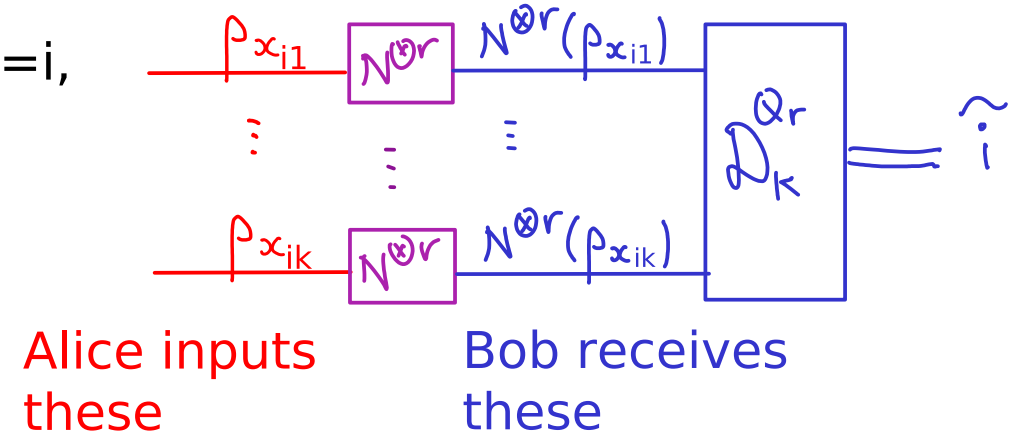
These defines a Qr-box $x \Rightarrow \boxed{Q_r} \rightarrow N^{\otimes r}(\rho_x)$

- Now code for Qr-box: take random code of length k (large k):

$$\begin{matrix} x_{11} & \cdots & x_{1j} & \cdots & x_{1k} \\ x_{i1} & \cdots & x_{ij} & \cdots & x_{ik} \\ x_{M1} & \cdots & x_{Mj} & \cdots & x_{Mk} \end{matrix} \qquad \text{for } M = 2^{k\left[\chi(\{p_x, N^{\otimes r}(\rho_x)\}) - \frac{\delta}{2}\right]}$$

where each x_{ij} drawn iid ~ $p(x)$, reject i-th row if not strongly typical

- To transmit m=i,



$$\therefore\ C(N) \geq \sup_{n} \max_{\{\rho_m\}} \frac{1}{n} C(Q_n)$$

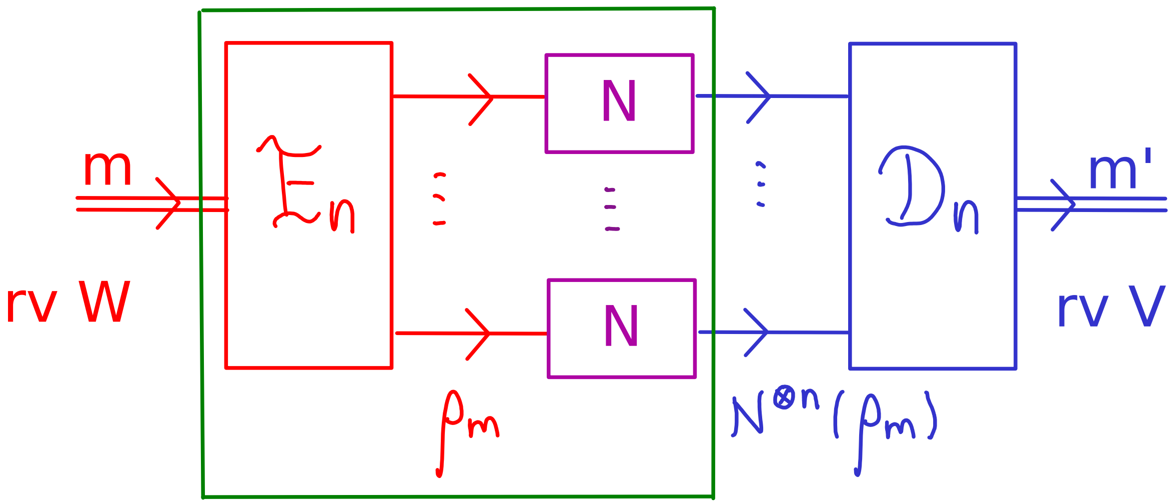
$\{\rho_m\}$ ← inputs to n uses of N

$$= \sup_{n} \left[\frac{1}{n} \max_{\{p_m, \rho_m\}} \chi\left(\{ p_m, N^{\otimes n}(\rho_m) \} \right) \right]$$

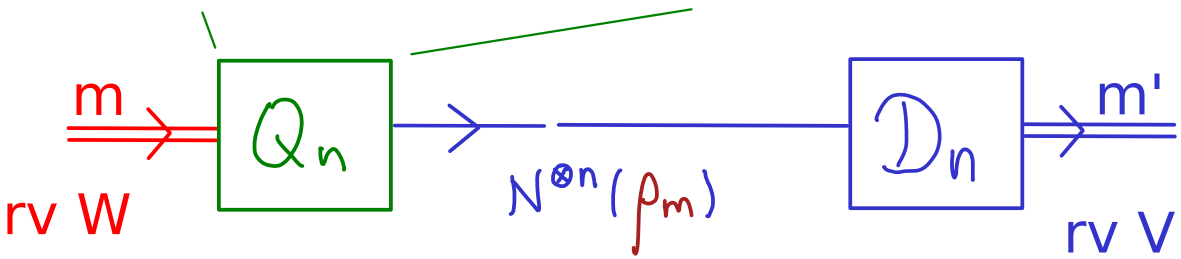
We next show the RHS IS an also an upper bound for any achievable rate, thus it is the capacity ...

(2) CONVERSE

Most general comm protocol using N n times:



or:



no matter how large n is, how well p_m, ρ_m are chosen,
if R achievable , $\epsilon_n \to 0$,

$$nR \lesssim I(W:V) \le I_{acc}(\{p_m, N^{\otimes n}(\rho_m)\}) \le \max_{\{p_m,\rho_m\}} I_{acc}(\{p_m, N^{\otimes n}(\rho_m)\})$$

Fanos ineq

$$\le \max_{\{p_m,\rho_m\}} \chi(\{p_m, N^{\otimes n}(\rho_m)\})$$

$$\therefore R \le \sup_n \frac{1}{n} \max_{\{p_m,\rho_m\}} \chi(\{p_m, N^{\otimes n}(\rho_m)\}) .$$

Putting (1) and (2) together, we obtain:

Theorem (Holevo-Schumacher-Westmoreland) (HSW Thm):

$$C(N) = \sup_r \frac{1}{r} \max_{\{p_x, \rho_x\}} \chi(\{p_x, N^{\otimes n}(\rho_x)\}) =: \sup_r \chi^{(r)}(N)$$

where $$\chi(N) := \max_{\{p_x, \rho_x\}} \chi(\{p_x, N(\rho_x)\})$$ 1-shot Holevo info of N

p_x: arbitrary label x

ρ_x: input to 1 channel labeled by x

r-shot Holevo info of N

$$\chi^{(r)}(N) := \frac{1}{r} \chi(N^{\otimes r}) = \frac{1}{r} \max_{\{p_x, \rho_x\}} \chi(\{p_x, N^{\otimes r}(\rho_x)\})$$

p_x: arbitrary label x

ρ_x: input to r channels labeled by x

The capacity expression is called "regularized", optimized over r, then an optimization involving r uses of N. (Classical capacity of classical channels and Q-boxes are "single-letter" -- optimization involving 1 use of N.)

* Holevo: IEEE TIT 44 p269 (1998)
Schumacher and Westmoreland: PRA 56 p131 (1999)

Putting (1) and (2) together, we obtain:

Theorem (Holevo-Schumacher-Westmoreland) (HSW Thm):

$$C(N) = \sup_r \frac{1}{r} \max_{\{p_x, \rho_x\}} \chi\left(\{p_x, N^{\otimes n}(\rho_x)\}\right) =: \sup_r \chi^{(r)}(N)$$

where $\chi(N) := \max_{\{p_x, \rho_x\}} \chi(\{p_x, N(\rho_x)\})$ 1-shot Holevo info of N

p_x: arbitrary label x

ρ_x: input to 1 channel labeled by x

r-shot Holevo info of N

$$\chi^{(r)}(N) := \frac{1}{r}\chi(N^{\otimes r}) = \frac{1}{r} \max_{\{p_x, \rho_x\}} \chi(\{p_x, N^{\otimes r}(\rho_x)\})$$

p_x: arbitrary label x

ρ_x: input to r channels labeled by x

if output states are product over the r uses, (e.g., if all input states are product or if the channel is "entanglement breaking") then, proof of converse for capacity of Q-boxes applies, and $\chi^{(r)}(N) = \chi(N)$

Optimizing the Holevo information of a channel:

Def: in the expression $\chi(N) := \max_{\{p_x, \rho_x\}} \chi(\{p_x, N(\rho_x)\})$

$\{p_x, \rho_x\}$ is called the "optimal ensemble" for N
if the max is attained on $\{p_x, \rho_x\}$

1. Finiteness of the optimal ensemble

Uhlmann 9701014, Schumacher and Westmoreland 9912122

(a) ρ_x's can be chosen pure, AND

(b) d^2 states are sufficient, where d = min (d_{in}, d_{out})

input, output dims of N

NB. thus the "max".

Proof: A3 Q3.

Optimizing the Holevo information of a channel:

Def: in the expression $\chi(N) := \max_{\{p_x, \rho_x\}} \chi(\{p_x, N(\rho_x)\})$

$\{p_x, \rho_x\}$ is called the "optimal ensemble" for N if the max is attained on $\{p_x, \rho_x\}$

Def: χ is strongly additive on N if $\forall N'$, $\chi(N \otimes N') = \chi(N) + \chi(N')$

Def: χ is weakly additive on N if $\forall r$, $\chi^{(r)}(N) = \chi(N)$

Lemma: $\forall N, N'$, $\chi(N \otimes N') \geq \chi(N) + \chi(N')$

Proof sketch :

Let $\{p_x, \rho_x\}$, $\{q_y, \sigma_y\}$ be optimal ensembles for N, N' respectively.

$\chi(N \otimes N') \geq \chi(\{p_x q_y, N \otimes N'(\rho_x \otimes \sigma_y)\})$ try product ensemble $\{p_x q_y, \rho_x \otimes \sigma_y\}$ for $N \otimes N'$

rewrite as QMI of $\Lambda \otimes \Lambda'$
equate to sum of QMI of Λ and QMI of Λ'

$= \chi(\{p_x, N(\rho_x)\}) + \chi(\{q_y, N'(\sigma_y)\})$

$= \chi(N) + \chi(N')$

Optimizing the Holevo information of a channel:

Def: in the expression $\chi(N) := \max_{\{p_x, \rho_x\}} \chi(\{p_x, N(\rho_x)\})$

$\{p_x, \rho_x\}$ is called the "optimal ensemble" for N if the max is attained on $\{p_x, \rho_x\}$

Def: χ is <u>strongly</u> additive on N if $\forall N'$, $\chi(N \otimes N') = \chi(N) + \chi(N')$

Def: χ is <u>weakly</u> additive on N if $\forall r$, $\chi^{(r)}(N) = \chi(N)$

Lemma: $\forall N, N'$, $\chi(N \otimes N') \geq \chi(N) + \chi(N')$

2. Nonadditivity of Holevo information

equiv:
nonadditivity of ent of formation
nonadditivity of min output entropy

Shor 0305035 + Hastings 0809.3972

$\exists N, N'$ s.t. $\chi(N \otimes N') > \chi(N) + \chi(N')$

$\exists N$ s.t. $\chi^{(2)}(N) > \chi(N)$

no known explicit example

See also:
Brandao, (M) Horodecki 0907.3210
Fukuda, King, Moser 0905.3697
Aubrun, Szarek, (E) Werner 0910.1189

Useful consequences of lemma $\forall N, N', \; \chi(N \otimes N') \geq \chi(N) + \chi(N')$

(a) lower bound for capacity:

$$C(N) = \sup_r \chi^{(r)}(N) \geq \chi^{(r)}(N) \geq \chi(N)$$

(b) characterization of zero capacity channels:

$$C(N) = 0 \iff \chi(N) = 0 \iff \forall \rho \;\; N(\rho) = \sigma \text{ constant}$$

Proof: (i) $C(N) = 0 \Rightarrow \chi(N) = 0$

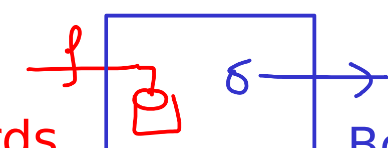
(ii) If $\chi(N) = 0 = \max_{\{p_x, \rho_x\}} S(X:Q)_\Lambda$, $\Lambda = \sum_x p_x |x\rangle\langle x| \otimes N(\rho_x)$

then $\forall_{\{p_x, \rho_x\}}$ $\Lambda = \sum_x p_x |x\rangle\langle x| \otimes N(\rho_x)$ is a product state

then $\forall \rho \;\; N(\rho) = \sigma$ constant

(iii) If $\forall \rho \;\; N(\rho) = \sigma$ constant

then N can be simulated without communication!



Alice discards her input

Bob produces σ himself

If $C(N) > 0$, violates C_2

$\therefore C(N) = 0$

Optimizing the Holevo information of a channel:

3. Examples

$\mathcal{N}_p(\rho) = (1-p)\rho + p\frac{I}{d}$ d-dim depolarizing channel

$$\chi(\mathcal{N}_p) = \max_{\{p_x, |\psi_x\rangle\}} S\Big(\sum_x p_x \mathcal{N}(|\psi_x\rangle\langle\psi_x|)\Big) - \sum_x p_x S\big(\mathcal{N}(|\psi_x\rangle\langle\psi_x|)\big)$$

For the 2nd term, $\mathcal{N}(|\psi_x\rangle\langle\psi_x|) = (1-p)|\psi_x\rangle\langle\psi_x| + p\frac{I}{d} = \begin{bmatrix} 1-p & & 0 \\ & 0 & \\ 0 & & \ddots \; 0 \end{bmatrix} + \frac{p}{d}\begin{bmatrix} 1 & & 0 \\ & 1 & \\ 0 & & \ddots \; 1 \end{bmatrix}$

Spectrum: $1-p+\frac{p}{d}, \underbrace{\frac{p}{d}, \ldots, \frac{p}{d}}_{d-1 \text{ times}}$ indep of $|\psi_x\rangle$ in any basis including $|\psi_x\rangle$

$$S\big(\mathcal{N}(|\psi_x\rangle\langle\psi_x|)\big) = -\left(1-p+\frac{p}{d}\right)\log\left(1-p+\frac{p}{d}\right) - (d-1)\frac{p}{d}\log\frac{p}{d} =: \tau$$

2nd term $= \sum_x p_x S\big(\mathcal{N}(|\psi_x\rangle\langle\psi_x|)\big) = \sum_x p_x \tau = \tau$

$$\therefore \chi(\mathcal{N}_p) = \underbrace{\max_{\{p_x, |\psi_x\rangle\}} S\Big(\sum_x p_x \mathcal{N}(|\psi_x\rangle\langle\psi_x|)\Big)} - \tau$$

attained when $\sum_x p_x \mathcal{N}(|\psi_x\rangle\langle\psi_x|) = \frac{I}{d}$ when $p_x = \frac{1}{d}$, $|\psi_x\rangle = |x\rangle$

(check: exercise)

$$= (\log d) - \tau$$

Ex: find $\chi(N)$ for $N(\rho) = 0.8\rho + 0.15 X\rho X + 0.05 Z\rho Z$

(ρ: 2×2; X, Z: Pauli's)

mixed Pauli channel

Ex: find $\chi(E_p)$ for $E_p(\rho) = (1-p)\rho + p|e\rangle\langle e|$

erasure channel

($|e\rangle$: erasure symbol orthogonal to any input)

Optimizing the Holevo information of a channel:

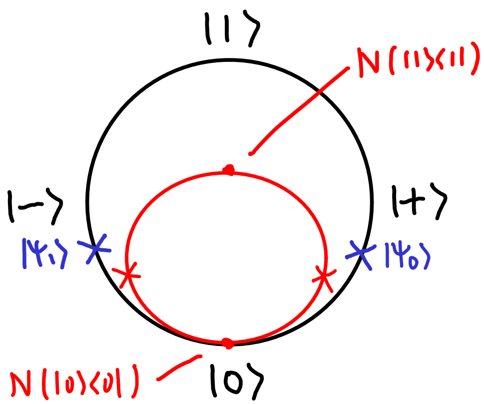
4. More distinguishable inputs need not have higher Holevo info !

$$\mathcal{N}_\gamma(\rho) = A_0 \rho A_0^\dagger + A_1 \rho A_1^\dagger$$ amplitude damping channel (AD)

$$A_0 = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{bmatrix}, \quad A_1 = \begin{bmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{bmatrix}$$

$$a|0\rangle + b|1\rangle \xrightarrow{A_0} a|0\rangle + b\sqrt{1-\gamma}\,|1\rangle$$
$$a|0\rangle + b|1\rangle \xrightarrow{A_1} b\sqrt{\gamma}\,|0\rangle$$

If $|0\rangle, |1\rangle$ ground and excited states then, AD describes de-excitation.



for any pair of orthogonal inputs

χ max at $p_0 = p_1 = \frac{1}{2}$, $\chi = 0.4567$.

for 2 such nonorthogonal inputs,

$\langle \psi_0 | \psi_1 \rangle \approx \cos 80^\circ$, $\chi \approx 0.4717$

Fuchs PRL 79 1162 (1997) first example, 9912122 AD channel.

Optimizing the Holevo information of a channel:

5. Hardness to estimate $\chi(N)$ Beigi & Shor 0707.2090

Let $c \in \mathbb{R}^+$. To decide whether $\chi(N) > c$ or $\chi(N) < c - \varepsilon$ for $\varepsilon = \frac{1}{\text{poly}(d)}$ is NP complete.

(d: input dim)

6. Continuity of C(N):   Leung & Smith 0810.4931

$$|C(N) - C(M)| \leq 8 \, \|N - M\|_\diamond \log d_{out} + 4 h(\|N-M\|_\diamond)$$

Optimizing the Holevo information of a channel:

7. Special channels with known additive Holevo info

(a) If N is entanglement breaking, then χ is strongly additive on N

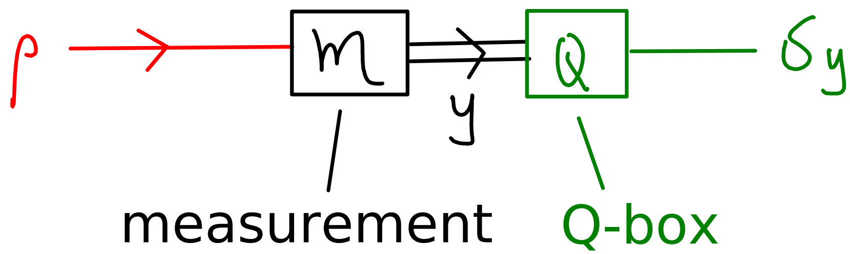
Cor: $C(N) = \chi(N)$

Def: N is entanglement breaking if,

$\forall \rho_{RA}$, $I \otimes N(\rho_{RA})$ separable (i.e., being a mixture of product states)

e.g., classical channels and Q-boxes are entanglement breaking

Aside: characterization of entanglement breaking channels



Ex: show the characterization.
Hint: apply def of ent break to Choi matrix.

Optimizing the Holevo information of a channel:

7. Special channels with known additive Holevo info (King 0103156)

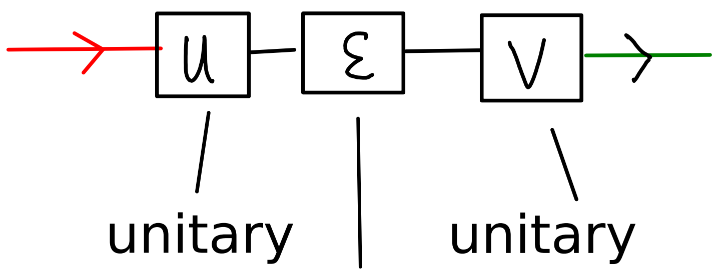
(b) χ is strongly additive on d-dim depolarizing channel $\mathcal{N}_p$

Cor: $C(N) = \chi(N) = (\log d) - \tau$

(b) χ is strongly additive on qubit unital channels $(N(I) = I)$

Cor: $C(N) = \chi(N)$

Aside: characterization of qubit unital channels



$$\mathcal{E}(\rho) = (1-p_x-p_y-p_z)\,\rho + p_x X\rho X + p_y Y\rho Y + p_z Z\rho Z$$

(c) Amplitude dampling channel is NOT known to have additive χ