CO781 / QIC 890:

Theory of Quantum Communication

Topics 4, part 4

Encoding classical information in quantum states and retrieving it

Scenario 3: classical capacity of Q-boxes (cq-channels)

Definition:

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

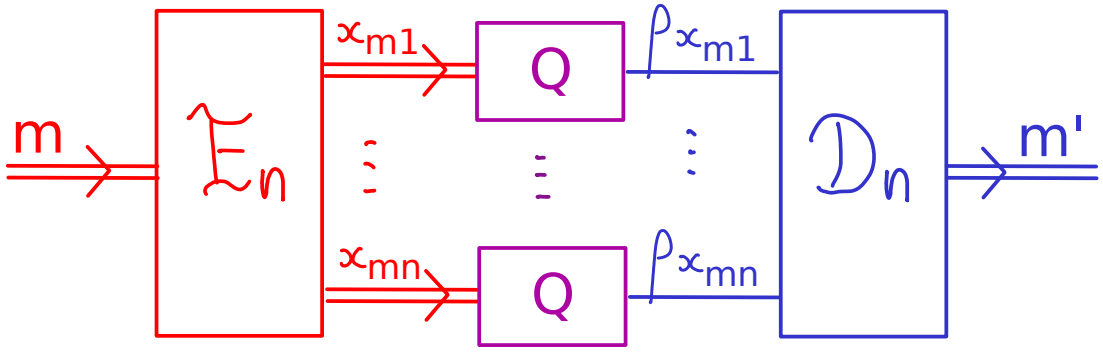
ρ_x: states on a common space, say, a d-dim system

Ω: input alphabet

If Alice inputs x, then, Bob gets ρ_x :

How to optimize communication rate using Q-boxes?

Most general communication protocol using Q-boxes n times:



1. Alice's message is m
2. She looks up code book to find n classical inputs to the Q-boxes $x_{m1} \dots x_{mn}$
3. She enters the inputs
4. Bob gets the output q systems in the state $\rho_{x_{m1}} \otimes \cdots \otimes \rho_{x_{mn}}$
5. He applies a measurement $\mathcal{D}_n$ that outputs m'

jointly for optimality

Codebook (known to Alice, Bob):

m=1: $x_{11} \dots x_{1j} \cdots x_{1n}$

m= i $x_{i1} \dots x_{ij} \cdots x_{in}$

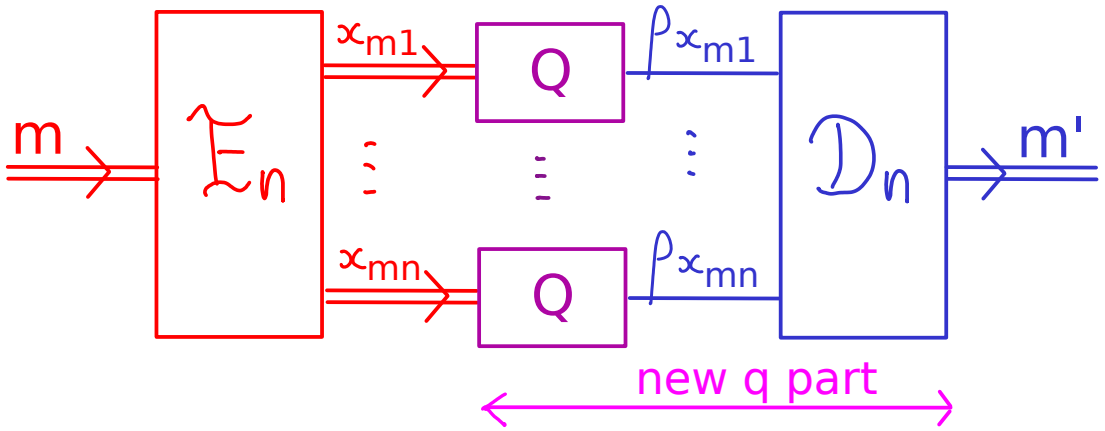
m=M: $x_{M1} \dots x_{Mj} \cdots x_{Mn}$

picked and fixed so, the n states are correlated for each m

given, no choice

$\{\rho_x\}_{x \in \Omega}$

Most general communication protocol using Q-boxes n times:



Classical notions still apply:

(M,n) code $\mathcal{C}_n$ where n: #uses of Q-box, M: #messages (fcn of n)

For each $\mathcal{C}_n$

* Specification:

Codebook (known to both):

m=1: $x_{11} \cdots x_{1j} \cdots x_{1n}$

m= i $x_{i1} \cdots x_{ij} \cdots x_{in}$

m=M: $x_{M1} \cdots x_{Mj} \cdots x_{Mn}$

Bob's meas $\mathcal{D}_n$ (known to both)

* Performance

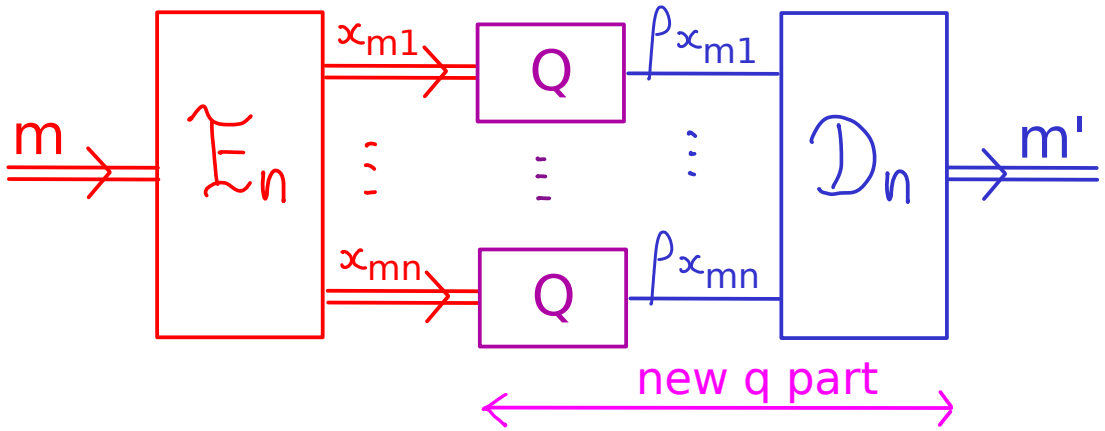
$P_e(m) = \text{prob}(m \neq m')$

(nonortho of q states, Bob's meas cannot perfectly distinguish m)

$\mathbb{E}\, P_e(\mathcal{C}_n) = \frac{1}{M} \sum_m P_e(m)$

$P_e(\mathcal{C}_n) = \max_m P_e(m)$

Most general communication protocol using Q-boxes n times:



Classical notions still apply:

(M,n) code $\mathcal{C}_n$ where n: #uses of Q-box, M: #messages (fcn of n)

For each $\mathcal{C}_n$ Codebook Bob's meas $P_e(\mathcal{C}_n)$

A rate R>0 is achievable if,

$\exists$ (M,n) codes, $\frac{1}{n}\log(M) \geq R$, $P_e(\mathcal{C}_n) \to 0$

NB 1: single-letter

2. smaller Iacc if Bob can't meas jointly

Theorem: Capacity of Q-box, $C(Q) = \max_{p(x)} S(X:Q)$ $\quad \Lambda = \sum_x p_x |x\rangle\langle x| \otimes \rho_x$

cf classical case $C(N) = \max_{p(x)} I(X:Y)$ $\quad \Lambda = \sum_x p_x |x\rangle\langle x| \otimes \sum_y p(y|x) |y\rangle\langle y|$

Theorem: Capacity of Q-box, $C(Q) = \max_{p(x)} S(X:Q)$ $\Lambda = \sum_x p_x |x\rangle\langle x| \otimes \rho_x$

$= \max_{p(x)} \chi(\{p_x, \rho_x\})$

Proof ideas:

(1) For the direct coding theorem:

Conceptual:

(a) random codes for the codebook

(b) pretty good measurement for decoding

Technical:

(c) the packing lemma

(d) the gentle measurement lemma

(2) For the converse:

Fanos inequality

cq-channel

product structure of Bob's received states

(1) For the direct coding theorem:

Fix any Q-box specified by $\{\rho_x\}_{x\in\Omega}$, any distribution p(x)

(a) consider random codes for the codebook

m=1: $c_1 = x_{11} \cdots x_{1j} \cdots x_{1n}$

m= i : $c_i = x_{i1} \cdots x_{ij} \cdots x_{in}$

m=M: $c_M = x_{M1} \cdots x_{Mj} \cdots x_{Mn}$

where each x_{ij} is drawn iid ~ p(x),

and reject c_i if it is not strongly typical

To transmit m=i, Alice inputs $x_{i1} \cdots x_{ij} \cdots x_{in}$ into $Q^{\otimes n}$

Bob receives $Y_i = \rho_{x_{i1}} \otimes \cdots \rho_{x_{ij}} \otimes \cdots \rho_{x_{in}}$

e.g., Let $|\psi_0\rangle = |0\rangle$

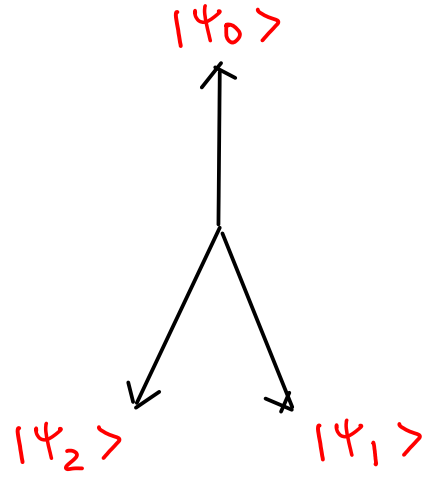
$|\psi_1\rangle = \cos\frac{\pi}{3}|0\rangle + \sin\frac{\pi}{3}|1\rangle$

$|\psi_2\rangle = \cos\frac{\pi}{3}|0\rangle - \sin\frac{\pi}{3}|1\rangle$

s.t. $0.4\,|\psi_0\rangle\langle\psi_0| + 0.3\,|\psi_1\rangle\langle\psi_1| + 0.3\,|\psi_2\rangle\langle\psi_2| = \frac{I}{2}$

Let $\rho_x = 0.9\,|\psi_x\rangle\langle\psi_x| + 0.1\,\frac{I}{2}$

For example, the Q-box tries to emit $|\psi_x\rangle$ when the input is x, but with 10% "garbage" $\frac{I}{2}$



closer pair

Let p(0) = 0.4, p(1) = p(2) = 0.3.

For n = 20, codebook of M messages looks like:

Bob's state

c1 = 10012 00210 12120 12200 $\quad \gamma_1 = \rho_1\otimes\rho_0\otimes\rho_0\otimes\rho_1\otimes\rho_2\otimes\cdots\otimes\rho_0\otimes\rho_0$

c2 = 02101 20120 02211 02112 $\quad \gamma_2 = \rho_0\otimes\rho_2\otimes\rho_1\otimes\rho_0\otimes\rho_1\otimes\cdots\otimes\rho_1\otimes\rho_2$

...

c_M = 22011 10021 20010 10220 $\quad \gamma_M = \rho_2\otimes\rho_2\otimes\rho_0\otimes\rho_1\otimes\rho_1\otimes\cdots\otimes\rho_2\otimes\rho_0$

roughly 8 "0"'s, 6 "1"'s, 6 "2"'s

how distinguishable are they?

how large can M be?

what measurement to use?

e.g., c2 = 02101 20120 02211 02112

$\gamma_2 = \rho_0 \otimes \rho_2 \otimes \rho_1 \otimes \rho_0 \otimes \rho_1 \otimes \rho_2 \otimes \rho_0 \otimes \rho_1 \otimes \rho_0 \otimes \rho_2 \otimes \rho_2 \otimes \rho_1 \otimes \rho_1 \otimes \rho_0 \otimes \rho_2 \otimes \rho_1 \otimes \rho_2 \otimes \rho_1$

(1) allowing a small approx,

can project the 6 ρ_0's to typical space of roughly $2^{6S(\rho_0)}$ dims

7 ρ_1's to typical space of roughly $2^{7S(\rho_1)}$ dims

7 ρ_2's to typical space of roughly $2^{7S(\rho_2)}$ dims

(from p3-6 topic-2.2.pdf, precursor to Schumacher compression)

thus can project γ_2 to a subspace of roughly $2^{6S(\rho_0)}\, 2^{7S(\rho_1)}\, 2^{7S(\rho_2)}$ dims

Since all ci's are strongly typical, the same holds for each γ_i

(2) meanwhile, γ_2 consists of 20 iid draws of $\{p(x), \rho_x\}$

Let $\rho = \sum_x p(x)\, \rho_x$ (which by coincidence, is I/2)

Projecting γ_2 onto typical space of $\rho^{\otimes 20}$ of dim $2^{20\, S(\rho)}$

doesn't change it very much (Schumacher compression).

Back to general Q-box, general p(x) ...

$$\gamma_i = \rho_{x_{i1}} \otimes \cdots \rho_{x_{ij}} \otimes \cdots \rho_{x_{in}}$$

For large n, for each γ_i for each x

there are approx np(x) ρ_x's in known positions

which can be projected whp onto the $2^{n\,p(x)\,S(\rho_x)}$ dim typical space

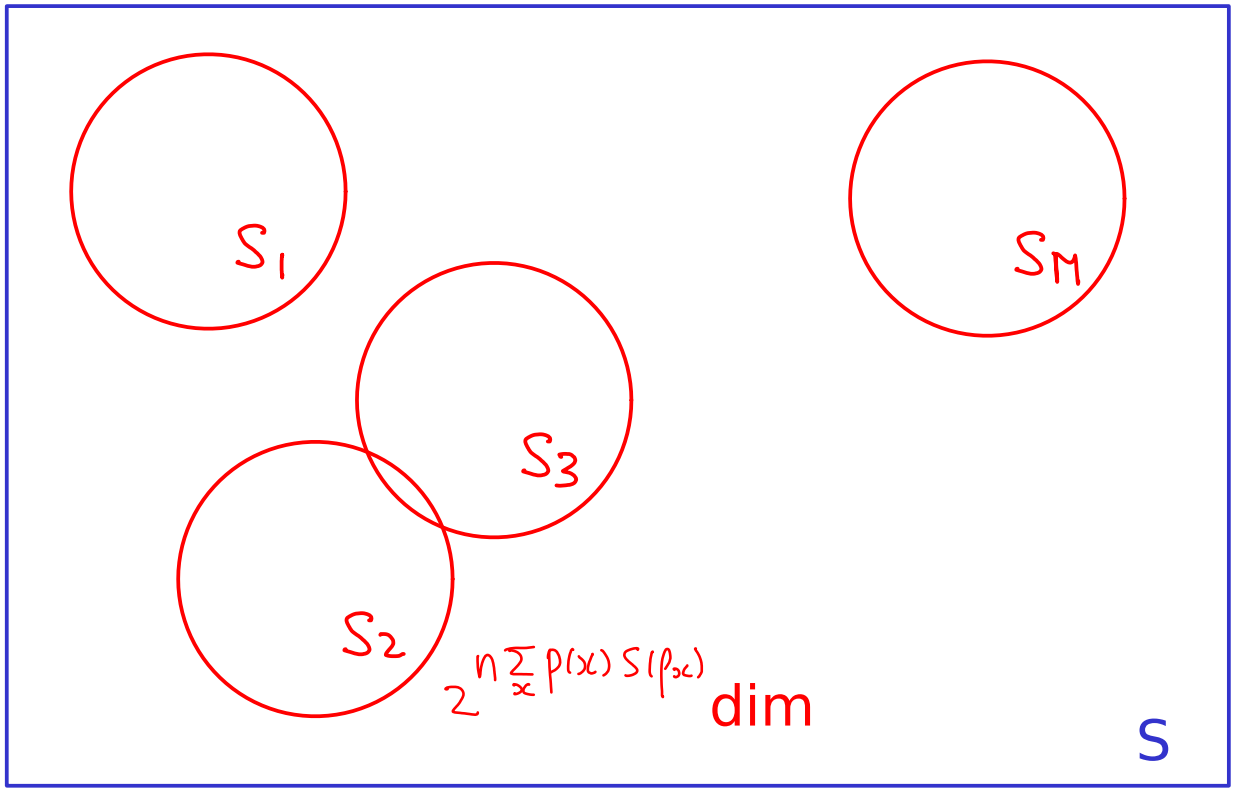
So, γ_i can be projected whp onto a $2^{n\sum_x p(x)\,S(\rho_x)}$ dim space.

So, each γ_i lives in a $2^{n\sum_x p(x)\,S(\rho_x)}$ dim space S_i.

For large n, all γ_i also approx live in the typical space of $\rho^{\otimes n}$

where $\rho = S\left(\sum_x p(x)\,\rho_x\right)$. This typical space S has $2^{n\,S\left(\sum_x p(x)\,\rho_x\right)}$ dim.

(This is by Schumacher compression of the n states, and the same typical space contains all γ_i.)



If the Y_i's (or S_i's) don't overlap too much, there is a measurement for Bob to decode with small prob of error.

We expect to be able to pack $\approx \dfrac{2^{n S(\sum_x p(x) \rho_x)}}{2^{n\sum_x p(x) S(\rho_x)}}$ Y_i's into S.

$$= 2^{n\chi(\{p(x), \rho_x\})} \approx M$$

More technical detail:

The distortions caused by various projections are upper bounded by:

(d) The gentle measurement lemma

(Winter IEEE TIT 45(7) p2481-2485, 1999)

Let $\rho \geq 0$, $tr(\rho) \leq 1$, $0 \leq E \leq I$

one element in the POVM for a meas

If $tr(\rho E) \geq 1-\eta$ ← if E corr to a likely outcome for ρ

then $\| E^{\frac{1}{2}} \rho E^{\frac{1}{2}} - \rho \|_1 \leq \sqrt{3\eta}$ ← then meas preserves ρ

In our problem, E is the projector onto some typical subspace.

More technical detail:

Bob uses the "pretty-good-measurement" (PGM) for the decoding.

(c) The pretty good measurement (Belavkin 75)

Let $\sigma_1, \sigma_2, \ldots, \sigma_k \in Pos(\mathbb{C}^d)$ (d-dim, subnormalized states or states weighted by probs)

The PGM for $\{\sigma_i\}_{i=1}^k$ has POVM elements

- $M_i = \Gamma^{-\frac{1}{2}} \sigma_i \Gamma^{-\frac{1}{2}}$ for i = 1,2, ..., k
- $M_{k+1} = I - \sum_i M_i$

where $\Gamma = \sum_{i=1}^k \sigma_i$ — positive semidefinite

$= \sum_j \lambda_j |e_j\rangle\langle e_j|$ — spectral decomp

($\lambda_j > 0$, $|e_i\rangle$ orthonormal)

$\Gamma^{-\frac{1}{2}} = \sum_j \lambda_j^{-\frac{1}{2}} |e_j\rangle\langle e_j|$ (applying inv-square-root to the supp)

(saw last lecture, for double trine)

NB $\forall i \;\; M_i = \Gamma^{-\frac{1}{2}} \sigma_i \Gamma^{-\frac{1}{2}} \geq 0$,

$$\sum_{i=1}^{K} M_i = \sum_{i=1}^{K} \Gamma^{-\frac{1}{2}} \sigma_i \Gamma^{-\frac{1}{2}} = \Gamma^{-\frac{1}{2}} \sum_{i=1}^{K} \sigma_i \Gamma^{-\frac{1}{2}}$$

$$= \Gamma^{-\frac{1}{2}} \Gamma \Gamma^{-\frac{1}{2}} = \sum_j |e_j\rangle\langle e_j| \leq I$$

- $M_i = \Gamma^{-\frac{1}{2}} \sigma_i \Gamma^{-\frac{1}{2}}$
- $M_{K+1} = I - \sum_i M_i$

$$\Gamma = \sum_{i=1}^{K} \sigma_i = \sum_j \lambda_j |e_j\rangle\langle e_j|$$

$$\Gamma^{-\frac{1}{2}} = \sum_j \lambda_j^{-\frac{1}{2}} |e_j\rangle\langle e_j|$$

$\therefore$ Each $M_i \leq I$. $\therefore \forall i \;\; 0 \leq M_i \leq I$.

Also, $M_{K+1} = I - \sum_{i=1}^{K} M_i = I - \underbrace{\sum_j |e_j\rangle\langle e_j|}_{\text{projector}} \geq 0$

$\therefore 0 \leq M_{K+1} \leq I$ $\leq I$

$\therefore$ $M_1, M_2, \ldots, M_{K+1}$ form a POVM.

More technical detail:

Bob uses the "pretty-good-measurement" (PGM) for the decoding.

(b) The pretty good measurement (Belavkin 75)

Let $\sigma_1, \sigma_2, \ldots, \sigma_K \in Pos(\mathbb{C}^d)$

The PGM for $\{\sigma_i\}_{i=1}^K$ has POVM elements

- $M_i = \Gamma^{-\frac{1}{2}} \sigma_i \Gamma^{-\frac{1}{2}}$ for i = 1,2, ..., k
- $M_{K+1} = I - \sum_i M_i$

where $\Gamma = \sum_{i=1}^K \sigma_i = \sum_j \lambda_j |e_j\rangle\langle e_j|$, $\Gamma^{-\frac{1}{2}} = \sum_j \lambda_j^{-\frac{1}{2}} |e_j\rangle\langle e_j|$

In our problem, $\sigma_i = \gamma_i$, $K = M = 2^{n\chi(\{p(x), \rho_x\})}$

More technical detail:

To show that (M,n) code exists for $M = 2^{n\chi(\{p(x), \rho_x\})}$ with error ->0

(c) The packing lemma gives a precise upper bound on the error probability averaged over messages and over the choice of code, for the PGM, as a function of the dim of the common space, and the dim of each γ_i .

$$\mathbb{E}\, P_e(\mathcal{C}_n) \leq \underbrace{2(\eta_2 + \sqrt{3\eta_1})}_{\text{prob of not in typical spaces}} + \frac{2^{n\sum_x p(x) S(\rho_x)}}{2^{n S(\sum_x p(x)\rho_x)}} \times M$$

can be $2^{n\chi(\{p(x), \rho_x\} - \delta)}$ for $\delta \to 0$

The rest are same as classical capacity through classical channel:

$\mathbb{E}\, P_e(\mathcal{C}_n)$ small $\Rightarrow$ $\exists\, \mathcal{C}_n$ s.t $\mathbb{E}\, P_e(\mathcal{C}_n)$ small

$\Rightarrow$ $\exists\, \mathcal{C}_n'$ s.t. $P_e(\mathcal{C}_n)$ small expunge bad words

More technical detail:

See notes for Lecture 12-13, Oct 25, 2016 for F2016 offerimg.

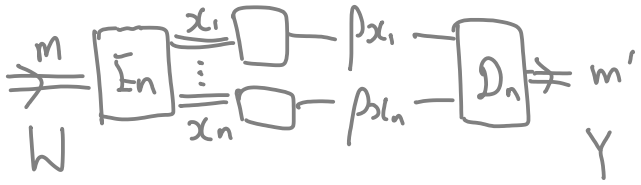
(2) For the converse:

Suppose R>0 is an achievable rate, so, there are $(2^{nR}, n)$ codes with vanishing error prob.

Suppose the code uses the codeword $x_1 x_2 \cdots x_n$ wp $p(x_1 x_2 \cdots x_n)$

By Fanos inequality, $\exists f(\epsilon_n) \to 0$ as $\epsilon_n \to 0$ s.t.

$$nR - f(\epsilon_n) \leq I(W:Y) \leq \chi_{acc}\left(\{p(x_1 x_2 \cdots x_n), \rho_{x_1} \otimes \rho_{x_2} \otimes \cdots \otimes \rho_{x_n}\}\right)$$



(2) For the converse:

Suppose R>0 is an achievable rate, so, there are $(2^{nR}, n)$ codes with vanishing error prob.

Suppose the code uses the codeword $x_1 x_2 \cdots x_n$ wp $p(x_1 x_2 \cdots x_n)$

By Fanos inequality,

$$nR \approx I(W:Y) \leq \mathcal{I}_{acc}\left(\{p(x_1 x_2 \cdots x_n), \rho_{x_1} \otimes \rho_{x_2} \otimes \cdots \otimes \rho_{x_n}\}\right)$$

$$\leq \chi\left(\{p(x_1 x_2 \cdots x_n), \rho_{x_1} \otimes \rho_{x_2} \otimes \cdots \otimes \rho_{x_n}\}\right)$$

$$= S(X_1 X_2 \cdots X_n : B_1 B_2 \cdots B_n)_\Lambda$$

$$\Lambda = \sum_{x_1 \cdots x_n} p(x_1 \cdots x_n) \underbrace{|x_1\rangle\langle x_1|}_{X_1} \otimes \cdots \otimes \underbrace{|x_n\rangle\langle x_n|}_{X_n} \otimes \underbrace{\rho_{x_1}}_{B_1} \otimes \cdots \otimes \underbrace{\rho_{x_n}}_{B_n}$$

$$= S(B_1 \cdots B_n) - S(B_1 \cdots B_n | X_1 \cdots X_n)$$

$$\wedge\!\vee \text{ SA} \qquad\qquad \| $$

$$\sum_i S(B_i) \qquad \sum_{x_1 \cdots x_n} p(x_1 \cdots x_n)\, S(B_1 \cdots B_n | x_1 \cdots x_n)$$

$$= \sum_i S(B_i) - \sum_{x_1 \cdots x_n} p(x_1 \cdots x_n)\, S(\rho_{x_1} \otimes \cdots \otimes \rho_{x_n})$$

$$= \sum_i S(B_i) - \sum_{x_1 \cdots x_n} p(x_1 \cdots x_n)\ S(\rho_{x_1} \otimes \cdots \otimes \rho_{x_n})$$

$$= \sum_i S(B_i) - \sum_{x_1 \cdots x_n} p(x_1 \cdots x_n) \sum_i S(\rho_{x_i})$$

product structure of
Bob's received states

$$= \sum_i S(B_i) - \sum_i p(x_i)\ S(\rho_{x_i})$$

$$\Lambda = \sum_{x_1 \cdots x_n} p(x_1 \cdots x_n) \underbrace{|x_1\rangle\langle x_1|}_{X_1} \otimes \cdots \otimes \underbrace{|x_n\rangle\langle x_n|}_{X_n} \otimes \underbrace{\rho_{x_1}}_{B_1} \otimes \cdots \otimes \underbrace{\rho_{x_n}}_{B_n}$$

$$S(B_i) = S\left(\sum_i p(x_i)\, \rho_{x_i}\right)$$

$$= \sum_i S\left(\sum_i p(x_i)\, \rho_{x_i}\right) - \sum_i p(x_i)\ S(\rho_{x_i})$$

$$= \sum_i \chi\left(\{p(x_i), \rho_{x_i}\}\right)$$

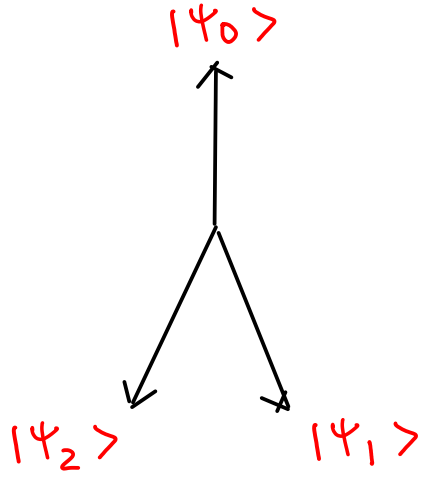
$$\leq n \max_{p(x)} \chi\left(\{p(x), \rho_x\}\right).$$

e.g., Let $|\psi_0\rangle$ $|\psi_1\rangle$ $|\psi_2\rangle$ be as before

s.t. $0.4\,|\psi_0\rangle\langle\psi_0| + 0.3\,|\psi_1\rangle\langle\psi_1| + 0.3\,|\psi_2\rangle\langle\psi_2| = \frac{I}{2}$

Let $\rho_x = 0.9\,|\psi_x\rangle\langle\psi_x| + 0.1\,\frac{I}{2}$

For example, the Q-box tries to emit $|\psi_x\rangle$ when the input is x, but with 10% "garbage" $\frac{I}{2}$



$$C(Q) = \max_{p_0,p_1,p_2} S\left(\sum_{x=1}^{3} p_x \rho_x\right) - \sum_{x=1}^{3} p_x S(\rho_x)$$

For each x, $\rho_x = 0.9\,|\psi_x\rangle\langle\psi_x| + 0.05\left(|\psi_x\rangle\langle\psi_x| + |\psi_x^{\perp}\rangle\langle\psi_x^{\perp}|\right)$

$= 0.95\,|\psi_x\rangle\langle\psi_x| + 0.05\,|\psi_x^{\perp}\rangle\langle\psi_x^{\perp}|$

$S(\rho_x) = h(0.05)$ indep of x

$$C(Q) = \max_{p_0,p_1,p_2} \underbrace{S\left(\sum_{x=1}^{3} p_x \rho_x\right)}_{\max \text{ at } \sum_{x=1}^{3} p_x \rho_x = \frac{I}{2}} - h(0.05) = 1 - h(0.05) \approx 0.7.$$

$\therefore\ p_0 = 0.4,\ p_1 = p_2 = 0.3$

saw the code earlier