CO781 / QIC890 Lec 11 Oct 18, 2016.

Encoding classical info in quantum states & retrieving it.

Abstract scenario. Let $p(x)$, ρ_x be fixed.

① Alice draws x with prob $p(x)$ ← rv X

② Alice prepares ρ_x

③ Alice sends ρ_x to Bob

④ Bob performs measurement $\mathcal{M}$ with POVM $\{M_y\}$. ← rv Y

This defines 2 classical rvs XY

where $p(xy) = p(y|x) \cdot p(x)$

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

Bob wants to learn about X via Y.

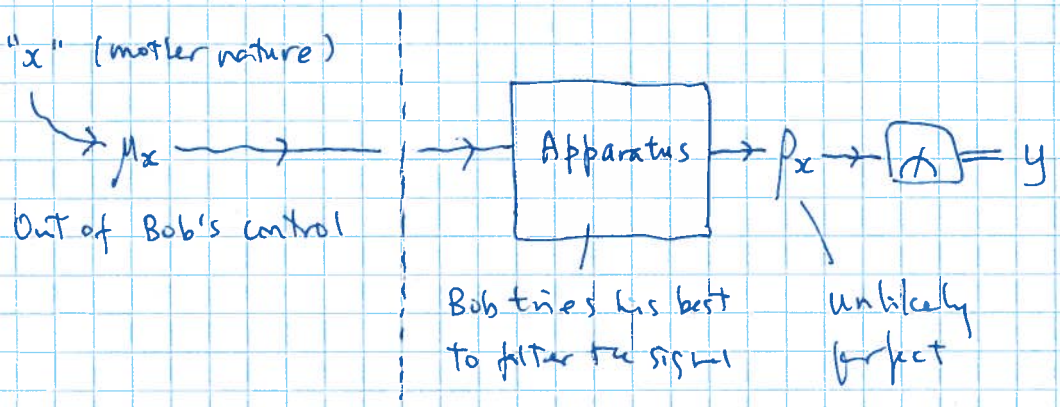
Why such scenario?

Why store x in non-orthogonal, possibly mixed quantum states?

Why $x \sim p(x)$?

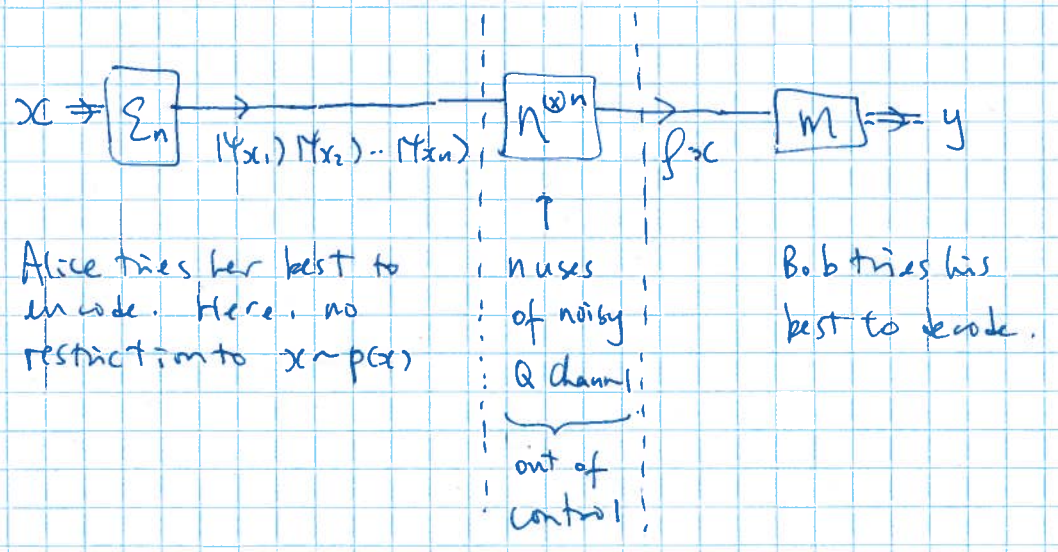
Partial answers:

- x arises from some physical process.
  Bob the experimentalist builds an apparatus to learn about the world.
  eg. gravitational waves in LIGO, frequency in atomic clocks @ NIST



- x accessible only via an oracle in query complexity
  eg it may label the marked item in Grover's algorithm

- In channel capacity problems:



Quantifying success:

- State discrimination problem with min prob of error.
  Here $\mathcal{Q}_X = \mathcal{Q}_Y$, goal is to min $\text{prob}(X \neq Y)$.
  Without further restriction on Bob's meas M,
  optimal solution can be obtained from an SDP (Watrous)

- Maximizing "accessible information".

Def [accessible info]:

Let $\Lambda = \sum_x p(x) \underset{X}{|x\rangle\langle x|} \otimes \underset{Q}{\rho_x}$

Let m be a measurement on Q with output space Y.

The accessible info for "the ensemble" $\mathcal{E} = \{p(x), \rho_x\}$ is

$$I_{acc}(\mathcal{E}) := \max_{m} \underset{I \otimes m(\Lambda)}{I(X:Y)}$$

NB I_{acc} is a commonly used measure.
But $I(X:Y)$ only makes sense with large # of iid draws.
We will discuss possible operational meaning later.
Stepping stone to classical capacity of quantum channels.

Upper bound for accessible info:

Recall given $\mathcal{E} = \{p(x), \rho_x\}$,

$$\Lambda = \sum_x p(x) \underbrace{|x\rangle\langle x|}_{X} \otimes \underbrace{\rho_x}_{Q}$$

$$S(X{:}Q) = \chi(\mathcal{E}) = S\left(\sum_x p(x)\rho_x\right) - \sum_x p(x) S(\rho_x)$$

$\chi(\mathcal{E})$: Holevo info of $\mathcal{E}$

Thm (Holevo bound): $I_{acc}(\mathcal{E}) \leq \chi(\mathcal{E})$

Pf: $I_{acc}(\mathcal{E}) = \max_{\mathcal{M}} I(X{:}Y)_{(\mathrm{Id}\otimes\mathcal{M})(\Lambda)} \leq S(X{:}Q)_{\Lambda}$

- $I(X{:}Y)_{(\mathrm{Id}\otimes\mathcal{M})(\Lambda)}$: $I = S$ for classical systems
- $\leq S(X{:}Q)_{\Lambda}$: monotonicity of QMI under TCP maps (incl. meas).

eg. $\chi(\mathcal{E}_1) = S\left(\frac{1}{3}\sum_{x=0}^{2} |\Psi_x\rangle\langle\Psi_x|\right) - 0 = S\left(\frac{\mathbb{I}}{2}\right) = 1$

$I_{acc}(\mathcal{E}_1) \doteq 0.5850$

Remark 1:

Consider an ensemble $\mathcal{E}$ of t 1-qubit pure states each drawn with prob $\frac{1}{t}$, and t can be arbitrarily large.

State preparation can be described as:

$$\sum_{x=1}^{t} \frac{1}{t} |x\rangle\langle x| \otimes \rho_x \;\rightarrow\; \sum_{x=1}^{t} \frac{1}{t} |x\rangle\langle x| \otimes \rho_x$$

X Q

However, without the label x, at most $I_{acc}(\mathcal{E})$ bits of info about x can be retrieved from sys Q.

$$I_{acc}(\mathcal{E}) \leq \chi(\mathcal{E}) \leq \log 2 = 1.$$

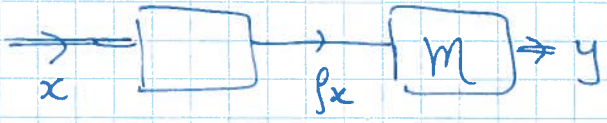
But we have $\log t$ bits of data in X!

So preparing a quantum state & forgetting the label is highly irreversible!

* Also makes teleportation more remarkable since neither Alice nor Bob can learn much about the classical description of the qubit to be sent (and it's transmitted with 2 cbits).

Remark 2

Holevo's bound says we cannot use 1 qbit to transmit more than 1 bit of data.



$p(x) = \frac{1}{|\Omega|}$

If $X = Y$ whp, $I(X:Y) \approx \log|\Omega|$

$$\leq \chi(\{p(x), \rho_x\})$$

$$\leq \log(\dim \mathcal{Q})$$

How to optimize M for I_{acc}?

Largely unresolved.

Some partial results:

① EB Davies IEEE TIT 24, P596, 1978

Let $\mathcal{E} = \{p(x), \rho_x\}$, $\rho_x \in B(\mathbb{C}^d)$

Then, optimal M with POVM $\{M_y\}_{y=1}^n$ can be chosen s.t.

ⓐ $\text{rank}(M_y) = 1$

ⓑ $d \leq n \leq d^2$

} simultaneously

Ideas behind: ⓐ If $M_i = \sum_k M_{i,k}$ where $M_{i,k}$ has rank 1, replace M_i by $M_{i,k}$ in the POVM to make a new measurement M'.

$M = M'$ followed by coarse graining outcomes corresponding to all (i,k) into outcome i.

So by the data processing inequality, M' gives at least as much mutual info about X as M.

ⓑ Careful use of convexity & Caratheodory's thm.

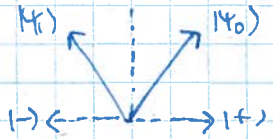
See T. Decker QS09122 Lemma 14 + Appendix for a proof.

Caratheodory's Theorem: Let $S \subseteq \mathbb{R}^n$, $\text{conv}(S)$ is convex hull. Then any $x \in \text{conv}(S)$ is a convex combination of at most $n+1$ elmts of S.

② Special ensembles

ⓐ Let $p(0) = p(1) = \frac{1}{2}$

$$|\psi_0\rangle = \cos\theta\,|0\rangle + \sin\theta\,|1\rangle$$
$$|\psi_1\rangle = \cos\theta\,|0\rangle - \sin\theta\,|1\rangle$$



Then optimal measurement (for I_{acc}) is along $\{|+\rangle, |-\rangle\}$

Pf: Chris Fuchs thesis P81, using a result by Levitin.

ⓑ If $\mathcal{E}$ is group covariant

ie $\exists$ finite group G s.t.

(i) can index ρ_x as ρ_g, $g \in G$

(ii) $p(g) = \frac{1}{|G|}$ $\forall g$

(iii) $\exists$ irrep s.t. $U_g \rho_h U_g^\dagger = \rho_{g\cdot h}$

[An irrep $\{U_g\} \subseteq \mathbb{U}(\dim(\rho_x))$ satisfies
$U_g \cdot U_h = U_{g\cdot h}$, $\forall g\ [U_g, K] = 0 \Rightarrow K \propto I$.]

then $\exists |\psi\rangle \in \mathbb{C}^d$ s.t. $\left\{\frac{d}{|G|} U_g^\dagger |\psi\rangle\langle\psi| U_g\right\}_{g\in G}$
is an optimal POVM.

Pf: Davies 78.

ⓒ Same as ⓑ, remove (iii), add condition

(iv) $U_g \cdot U_h = e^{i\phi(g,h)} U_{g\cdot h}$ ← $\{U_g\}$ "real"

$\{U_g\}$ acts irreducibly on $\mathbb{R}^d$

then, $\exists |\psi\rangle \in \mathbb{R}^d$ s.t. $\left\{\frac{d}{|G|} U_g^\dagger |\psi\rangle\langle\psi| U_g\right\}_{g\in G}$
is an optimal POVM.

Pf: 9812062

ⓓ Same as ⓑ, remove (iii), add orbits to POVM. Pf: Decker 0509122

Def: Let $\mathcal{E}_1 = \{p(x_1), \rho_{x_1}\}$, $\mathcal{E}_2 = \{q(x_2), \sigma_{x_2}\}$

The prob ensemble $\mathcal{E}_1 \otimes \mathcal{E}_2 := \{p(x_1)q(x_2), \rho_{x_1} \otimes \sigma_{x_2}\}$.

If we represent $\mathcal{E}_1$ as $\Lambda_1 = \sum_{x_1} p(x_1) \underset{X_1}{|x_1\rangle\langle x_1|} \otimes \underset{Q_1}{\rho_{x_1}}$

$\Lambda_2 = \sum_{x_2} q(x_2) \underset{X_2}{|x_2\rangle\langle x_2|} \otimes \underset{Q_2}{\sigma_{x_2}}$

then $\mathcal{E}_1 \otimes \mathcal{E}_2$ represented as

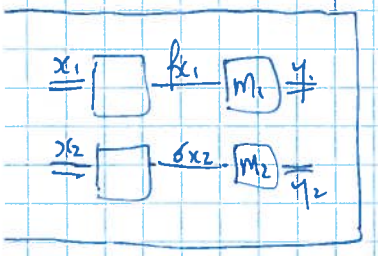
$$\Lambda = \sum_{x_1 x_2} p(x_1)q(x_2) \underbrace{|x_1\rangle\langle x_1| \otimes |x_2\rangle\langle x_2|}_{X_1X_2} \otimes \underbrace{\rho_{x_1} \otimes \sigma_{x_2}}_{Q_1Q_2}$$

$$= \Lambda_1 \otimes \Lambda_2$$

To evaluate $I_{acc}(\mathcal{E}_1 \otimes \mathcal{E}_2)$, we need to optimize over measurements on Q_1Q_2 yielding outcomes with most info on X_1X_2.

Thm: $I_{acc}(\mathcal{E}_1 \otimes \mathcal{E}_2) = I_{acc}(\mathcal{E}_1) + I_{acc}(\mathcal{E}_2)$

Pf: [$\geq$] Let M_1, M_2 be optimal for $\mathcal{E}_1$ & $\mathcal{E}_2$ respectively, with output spaces Y_1 & Y_2. Consider measurement on Q_1Q_2 which is $M_1 \otimes M_2$



Then $I(X_1X_2 : Y_1Y_2)$ (evaluated on $I \otimes M_1 \otimes M_2(\Lambda)$)

$$= H(X_1X_2) + H(Y_1Y_2) - H(X_1X_2Y_1Y_2)$$

X_1Y_1 indep from X_2Y_2

$$= H(X_1X_2) + H(Y_1Y_2) - H(X_1Y_1) - H(X_2Y_2)$$

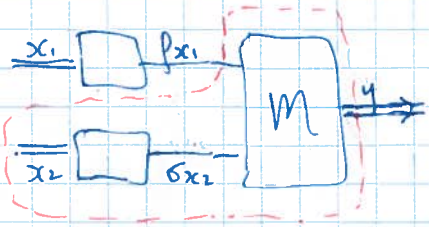
$$\leq H(X_1) + H(X_2) + H(Y_1) + H(Y_2) - H(X_1Y_1) - H(X_2Y_2)$$

(SA)

$$= I(X_1 : Y_1) + I(X_2 : Y_2) = I_{acc}(\mathcal{E}_1) + I_{acc}(\mathcal{E}_2).$$

$\therefore I_{acc}(\mathcal{E}_1 \otimes \mathcal{E}_2) \geq I_{acc}(\mathcal{E}_1) + I_{acc}(\mathcal{E}_2)$.

Pf: [≤] Let M be an arbitrary measurement on $Q_1 Q_2$ with output space Y.



$$I(X_1X_2:Y) = I(X_1:Y) + I(X_2:Y|X_1)$$

$$H(X_1X_2)+H(Y) \quad H(X_1)+H(Y) \quad H(X_2|X_1) - H(X_1)$$
$$-H(X_1X_2Y) \quad -H(X_1Y) \quad -H(X_2YX_1) + H(YX_1)$$

(*) Note that when X_1, X_2 independent, $\boxed{\quad}$ is a meas on Q_1 and $I(X_1:Y) \leq I_{acc}(\mathcal{E}_1)$.

(*) The 2nd term is $\sum_{x_1} p(x_1) \underbrace{I(X_2:Y \mid X_1=x_1)}_{\text{evaluated on prob } (x_2 y \mid X_1=x_1)}$.

$$\leq \max_{x_1} \cdot I(X_2:Y \mid X_1=x_1)$$

$$\leq I_{acc}(\mathcal{E}_2)$$

since preparing f_{x_1} for the maximizing x_1 in Q_1 and applying M on $Q_1 Q_2$ IS a particular measurement on Q_2.

$\therefore I(X_1X_2:Y) \leq I_{acc}(\mathcal{E}_1) + I_{acc}(\mathcal{E}_2)$

$\therefore I_{acc}(\mathcal{E}_1 \otimes \mathcal{E}_2) \leq I_{acc}(\mathcal{E}_1) + I_{acc}(\mathcal{E}_2)$

NB: argument was due to Wootters, and applies even for LOCC measurements if Σ_2 consists only of separable states.

See 0103098 Sec VII A.

NB: for <u>product ensembles</u>, it suffices to optimize & apply the measurement on each quantum system independently!

Eg. $I_{acc}(\xi_1 \otimes \xi_1) \doteq 2 \times 0.5850 \doteq 1.1700$

uniform distribution over $|\psi_i\rangle \otimes |\psi_j\rangle$
9 possible states.

eg 1. The ensemble $\mathcal{E}_1$ has:

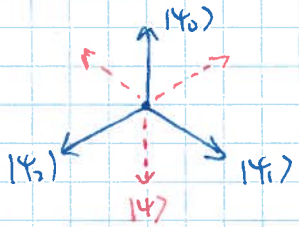
$p_0 = p_1 = p_2$, $\rho_{\mathcal{X}} = |\psi_x\rangle\langle\psi_x|$, and

$|\psi_0\rangle = |0\rangle$

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

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

They're called the "trine states".



Consider $G = \mathbb{Z}_3$, $U_g = e^{i\sigma_y \frac{2\pi}{3}\cdot g}$, $g = 0,1,2$.

Then conditions for 9812062 hold.

Optimal POVM $= \left\{ \frac{2}{3} U_g^\dagger |\psi\rangle\langle\psi| U_g \right\}$

where $|\psi\rangle$ can be chosen to be $|\psi_0^\perp\rangle$. (Pf: left as exercise.)

ie optimal POVM $= \underbrace{\left\{ \frac{2}{3} |\psi_y^\perp\rangle\langle\psi_y^\perp| \right\}}_{M_y}{}_{y=0,1,2}$

With this measurement:

$H(X|Y=0) = \log 2 = 1$ (rules out $|\psi_0\rangle$, $|\psi_{1,2}\rangle$ equiprobable)

Similarly for $Y = 1,2$.

$\therefore H(X|Y) = \sum_y H(X|Y=y) = 1$.

Also $H(X) = \log 3$. $\therefore I(X:Y) = H(X) - H(X|Y) = (\log 3) - 1 \doteq 0.5850$

$I_{acc}(\mathcal{E}_1) = 0.5850$.

eg2. The ensemble $\mathcal{E}_2$ is similar to $\mathcal{E}_1$ in eg 1

but now $\rho_x = |\psi_x\rangle\langle\psi_x|^{\otimes 2}$ (on $Q_1 Q_2$)

They're call the "double time" states.

Note: only 3 equiprobable states.

Note: $\mathcal{E}_2 \neq \mathcal{E}_1 \otimes \mathcal{E}_1$.

(I) Consider the following measurement M_2:

Step 1: apply optimal meas in eg 1 to Q_1.

If outcome is "a" (corresponding to $M_a = \frac{2}{3}|\psi_a^\perp\rangle\langle\psi_a^\perp|$)

then $|\psi_a\rangle^{\otimes 2}$ is mixed out.

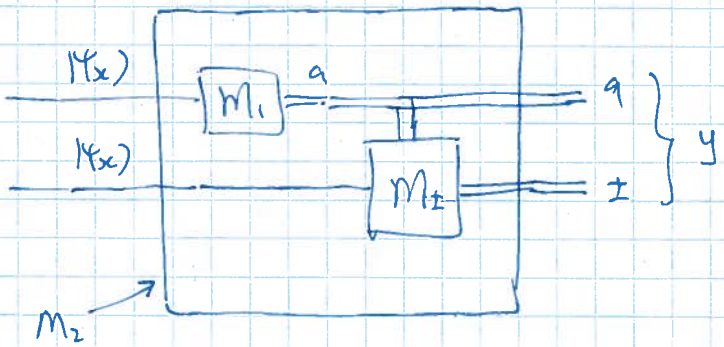
So state in Q_2 is $|\psi_b\rangle$ wp $\frac{1}{2}$

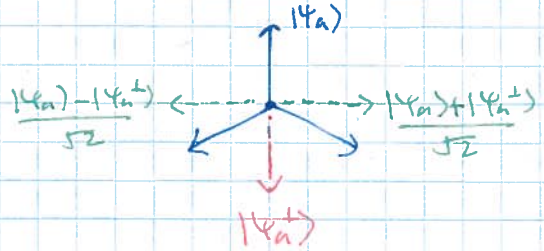
$|\psi_c\rangle$ wp $\frac{1}{2}$

where $b, c \in \{0,1,2\} \setminus \{a\}$, $b \neq c$.

Step 2: Measure Q_2 along the $\frac{1}{\sqrt{2}}(|\psi_a\rangle \pm |\psi_a^\perp\rangle)$ basis. Call this $M_\pm$.

optimal given the residual ensemble on Q_2.





In other words, M_2 optimizes the measurement on each system, allowing adaptation to all previous measurement outcomes.

Claim: $\mathbb{I}(X{:}Y) \approx 1.23038$ bits under M_2

Pf: AB.

NB: $\mathbb{I}(X{:}Y) > 2 \times \mathbb{I}_{acc}(\Sigma_1)$!

Lesson: restricting to 3 states in Σ_2 out of 9 states in $\Sigma_1^{\otimes 2}$ increases the accessible info!

NB: M_2 only uses 1-way classical comm from Q_1 to Q_2.

(*) Optimal $\mathbb{I}(X{:}Y)$ unknown for LOCC measurement.

(III) Consider the following measurement M_3:

Let $\Gamma = |\psi_0\rangle\langle\psi_0|^{\otimes 2} + |\psi_1\rangle\langle\psi_1|^{\otimes 2} + |\psi_2\rangle\langle\psi_2|^{\otimes 2}$

Let $M_y = \Gamma^{-\frac{1}{2}} |\psi_y\rangle\langle\psi_y|^{\otimes 2} \Gamma^{-\frac{1}{2}}$ for $y = 0,1,2$.

$M_3 = \mathbb{I} - M_0 - M_1 - M_2$.

Claim: $\mathbb{I}(X{:}Y) \approx 1.3691$ bits under M_3.

Pf: AB.

} call the Pretty Good measurement (PGM)

NB: M_3 is believed to be optimal for Σ_2 but no analytic proof. (0509122 Decker came close to understanding the optimal measurement but final step in argument was numerical.)

NB: The 3 states are related by $U_g \otimes U_g$ where U_g is defined in eq 1.

This is NOT irreducible so conditions for optimality do not apply.

NB: Wootters in 0506149 showed that there is a separable measurement (each POVM is a product operator) with the same $I(X:Y)$.

So if the PbM is proved optimal, optimal measurement for $I_{acc}(\mathcal{E}_2)$ can be chosen to be separable.

NB: M_3 achieves 1.3691 bits much higher than 1.23038 bits.

Lesson: Collective measurement (beyond LOCC) can be needed for achieving I_{acc}.

Conjectures: M_3 optimal & NOT achieved by LOCC

NB: Some recent work 1304.1555 showed the latter *under the success measure of min error of identifying the state.

Remark 3:

Recall the Araki-Lieb ineq and how much S & QMI can change when adding / discarding systems:

$$|S(AB) - S(A)| \leq S(B)$$

$$|S(A:BC) - S(A:C)| \leq 2S(B)$$

We now extend these bounds to the Holevo info:

Let $\mathcal{E}_1 = \{p(x), \rho_{x\,BC}\}$

$\mathcal{E}_2 = \{p(x), \sigma_{x\,C}\}$ where $\forall x$, $\sigma_{x\,C} = \mathrm{tr}_B\, \rho_{x\,BC}$

Then $|\chi(\mathcal{E}_1) - \chi(\mathcal{E}_2)|$

$= |S(A:BC) - S(A:C)|$ evaluated on $\sum_x p(x) |x\rangle\langle x|_A \otimes \rho_{x\,BC}$

$\leq 2S(B)$.

Remark 4:

Let $\mathcal{E}_1 = \{p(x), \rho_x\}$, $\mathcal{E}_2 = \{p(x), \eta(\rho_x)\}$ where η: TCP maps

Then $I_{acc}(\mathcal{E}_1) \geq I_{acc}(\mathcal{E}_2)$ by monotonicity of QMI under local TCP maps

$I_{acc}(\mathcal{E}_1) = S(X:Q)_\Lambda$, $I_{acc}(\mathcal{E}_2) = S(X:Q)_{I\otimes\eta(\Lambda)}$

where $\Lambda = \sum_x p(x) |x\rangle\langle x|_X \otimes \rho_{x\,Q}$

Remarks:

Thm [Cleve-van-Dam-Nielsen-Tapp 9708019]

Suppose Alice is allowed to send n_A qubits to Bob

Bob . . . n_B . . . Alice

in any order with any # of rounds.

Alice can communicate at most $n = n_A + n_B$ bits to Bob.

ie n_A qbits $\rightarrow$ + n_B qbit $\leftarrow$ $\geq \beta$ cbits $\rightarrow$ $\Rightarrow n_A + n_B \geq \beta$.

NB This generalizes Holevo's bdd to the interactive setting.

Pf: Let Alice's message be x, occurred with prob $p(x)$.

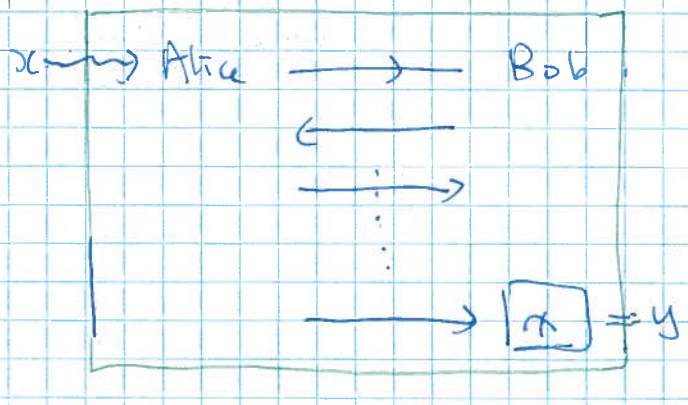
After the jth qubit of comm, let Bob's state be ρ_{xj}.

(How ρ_{xj} arises can be complicated...)

For each stage in the protocol, the rv x induces an ensemble

$\mathcal{E}_j = \{ p(x), \rho_{xj} \}$ and Bob has the q. part.

Let Bob's final estimate of the message be y.



$$I(X:Y) \leq I_{acc}(\mathcal{E}_n) \leq \chi(\mathcal{E}_n) \leq S\left(\sum_x p(x)\, \rho_{xn}\right) \quad \text{(by 1 qubit)}$$

(where $n = n_A + n_B$)

$$\leq S\left(\sum_x p(x)\, \rho_{x\,n-1}\right) + 1 \quad \text{(by 1 qubit)}$$

$$\vdots$$

$$\leq n_A + n_B = n.$$

Remark 6:

Locking of accessible info.

$\exists \{\xi_1 = \{p(x), \rho_{xBC}\}$ (and $\{\xi_2 = \{p(x), \mathrm{tr}_B \rho_{xC}\}$)

s.t $\left| I_{acc}(\xi_1) - I_{acc}(\xi_2) \right| \gg 2\, S(B)$.

In particular:

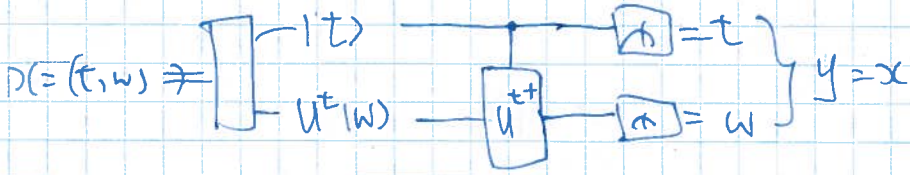
Take $p(x) = \frac{1}{2n}$, $x = wt$, $w \in \{0,1,\ldots,n-1\}$, $t \in \{0,1\}$

$$\rho_x = U^t |w\rangle\langle w| U^{t\dagger} \otimes |t\rangle\langle t|_B$$

$$U|w\rangle = \sum_l \omega^{wl} |l\rangle, \quad \omega^n = 1 \quad (U = FT).$$

In other words, we encode $0,1,\ldots,n-1$ in either the computation basis or the conjugate basis in B, and store the basis info in sys C.

$I_{acc}(\xi_1) = \log(2n)$, attained by measurement:



In 0303088, $I_{acc}(\xi_2) = \frac{1}{2}\log n$.

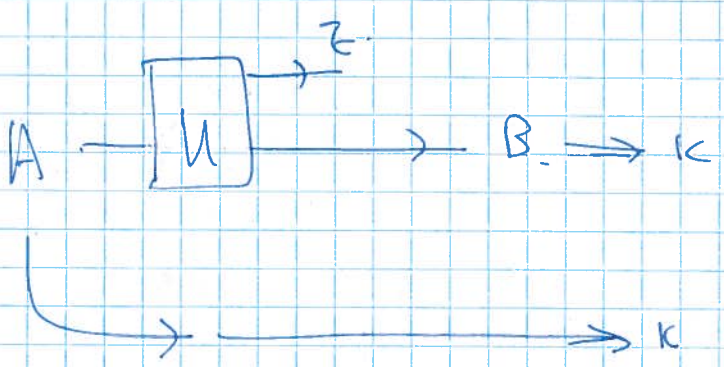
So dropping the 1-qubit sys B reduces I_{acc} from $1+\log n$ to $\frac{1}{2}\log n$! Diff $\gg 1$.

* Result has been improved on & extended.

* Accessible info assumes Bob makes a measurement.

But the measurement can be very different with a little side info, in the maps, it is the correct basis to decode w!

* This affects how we measure security of QKD.



Used to measure Iacc of Eve on K but that assumes Eve measures immediately after QKD protocol.

In real life, she can wait for more side info about K Before measuring.

In 0409078, we proposed more "composable" security measures for QKD.

Lower bound for I_{acc} [Jozsa, Robb, Wootters 94]

Def for a density matrix $\rho \in B(\mathbb{C}^d)$, with eigenvalues $\{\lambda_k\}_{k=1}^d$, the subentropy is

$$Q(\rho) := -\sum_{k=1}^{d} \underbrace{\left(\prod_{l \neq k} \frac{\lambda_k}{\lambda_k - \lambda_l}\right)}_{\text{additional part not in } S(\rho)} \lambda_k \log \lambda_k$$

Thm: for any ensemble $\mathcal{E} = \{p(x), \rho_x\}$

$$I_{acc}(\mathcal{E}) \geq Q\left(\sum_x p(x)\rho_x\right) - \sum_x p(x) Q(\rho_x)$$ ← like $\chi(\mathcal{E})$ with S replaced by Q

eg. If ρ_x pure for all x, and $\rho = \sum_x p(x)\rho_x = \frac{I}{d}$

then $I_{acc}(\mathcal{E}) \geq \log d - (\log e)\left(\frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{d}\right)$

$d = 2$, $I_{acc}(\mathcal{E}) \geq 0.2787$

$d \to \infty$, $I_{acc}(\mathcal{E}) \geq 0.60995$.

[Pf of Thm & eg: see PRA 49 p668 (1994).]

Also 1310.1312 for more on subentropy.