How much info can we learn about a quantum state by measuring it?

Suppose Alice draws $x$ and prepares $f_x$ and sends it to Bob.

Bob performs measurement $M$ with POVM $\{M_y\} \quad \mathcal{E} = \sum_y M_y \eta$

For classical RV's $X$, $Y$, $p(x,y) = p(y|x) p(x)$.

$p(y|x) = \text{tr} (M_y f_x)$, so $I(X;Y)$ depends only on $M$.

Def: for an ensemble $\mathcal{E} = \{p(x), p_x\}$

the accessible info $I_{acc}(\mathcal{E}) = \max_M I(X;Y)$.
Why such def? 

1. Why store classical info in non-orthogonal, possibly mixed states? 

2. Why $x \sim p(x)$? 

3. $I(X:Y)$ suggests iid draws, why? 

Some partial answers: 

- In physics, an experimenter (Bob) needs apparatus to measure some classical data about the world.
  
  e.g. gravity waves in LIGO, frequency in atomic clocks @ NIST.

"x" (Mother nature).

\[ \text{Apparatus} \quad \text{Bob's accuracy} \quad \text{Out of Bob's control} \quad \text{for}\ y \quad \text{Bob tries} \quad \text{unlimitedly}\ ]

\[ \text{by nature}, \]

- In channel capacities:

\[ x \xrightarrow{\text{E}_{m}} (y_1, y_2) \xrightarrow{\text{N}} (y_m) \xrightarrow{\text{P}_{x}} \]

\[ R = \text{h} \quad \text{Bob tries his best,}\]

Here signal need not obey $p(x)$ etc.
How to optimize $M$:

Major open problem, even for simple ensembles.

Some partial results:


Let $\mathcal{E} = \{p(x), p(x') | p(x) \in L_k(\mathbb{R}^d)\}.$

Optimal measure $\mu \in \mathcal{E}$ can be chosen s.t.

(a) $\text{rank}(M_\mu) = 1$

(b) $d \leq n \leq d^2$

Intuition:

(a) $M_\mu$ with rank $> 2$ is a positive sum of $M_k$'s, each with rank 1.

That we use $M_\mu = \sum_{i \in \mathbb{K}} M_{\mu_i}$ corresponds to "coarse graining" all outcomes $(\xi, k)$ into a single $\xi$, which cannot increase mutual info by monotonicity.

(b) Use convexity of and Caratheodory's theorem.
Special ensembles.

(a) \[ |\psi_1\rangle \quad |\psi_2\rangle \]

\[\langle \psi_1 | - - - - - - -\rangle \quad \langle \psi_2 | - - - - - - -\rangle\]

If \(|\psi_0\rangle = \cos \theta |0\rangle + \sin \theta |1\rangle\),

\(|\psi_1\rangle = \cos \theta |0\rangle - \sin \theta |1\rangle\).

Each drawn up \(\frac{1}{2}\).

Then optimal measurement is along \(1\langle \psi |.\)

Due to Chris finite thesis p81, using a result by Levitin.

(b) If \(G\) group covariant,

ie \(G\) finite group in it.

(i) can index \(G\) as \(f g \cdot g \in G\).

(ii) \(p(x) = \frac{1}{161}\) \(\forall x\)

(iii) \(\exists \text{ imp } \left[ \begin{array}{c} \{U_g \in U(G), \quad d = \exp f(x)\} \\ U_g \cdot U_h = U_{g \cdot h}, \quad "U_g K U_f = K \cdot g \Rightarrow K \cdot x \cdot K^{-1}" \end{array} \right] \text{ group op}\)

St. \(U_g \cdot U_f = U_{g \cdot f}\)

Then \(\exists 14) \in G\) such

\[\{U_g = \frac{d}{161} U^{\dagger} (14) U_h U^g \} \quad g \in G\]

is an optimal.

[DAÑES, 78].
(c) \(9 \text{, } 0\text{, } 6\text{, } 2\):

- **Remore condition (b) (iii):**

  Impose: \( \{ U_{g_1} \} \) is real
  \[
  [U_{g_1}, U_{g_2}] = C \phi (g_1, g_2) U_{g_1} U_{g_2}
  \]
  \( U_{g_1} \) acts irreducibly on \( \mathbb{R}^d \)

  Then: \( \exists \, \psi \in \mathbb{R}^d \) s.t.
  \[
  \left\{ \frac{d}{d \alpha} U_{g_1} \psi \right\}_\alpha < 4 \left( U_{g_1} \right)_\alpha^2 \quad g \in G.
  \]
  Is optimal POM.

(d) \(0 \text{, } 5\text{, } 0\text{, } 9\text{, } 1\text{, } 2\) (Decker)

  Reducible reps \( \{ U_{g_1} \} \), but complicated.
\[p_x = 1 \frac{1}{\sqrt{2}} \alpha x \gamma x 1\]
\[\alpha = 1, \quad \beta = 1\]
\[\gamma = \cos \frac{\pi}{3} \alpha + \sin \frac{\pi}{3} \beta\]
\[\xi = -\]

\[\text{equi probable.}\]

Consider \[\alpha = \xi\],
\[U_g = e^{i \varphi \alpha} g\]
\[\varphi = 0.112\]

Then conditions for \[98 \text{ (3052) held}\].

In fact \(14\) = \(14^0\), optm-() meas:
\[\hat{M}_2 = \frac{1}{3} 14^0, 14^1, 14^2\]
Here: \(H(x) = \log 3\).

\[H(x, y = 0) = 1 \quad \& \quad \text{equi probable}\]
\[14^0, 14^1\].

Similarly for \(y = 1, 2\).
\[H(x|y) = 1\].

\[\sum(x: y) = H(x) - H(x|y) = \log 3 - 1 \approx 0.5850\]

[\text{Hill from eq 1.}]

\[\sum_{ac} = \log 3 - 1\]
Upper bound to accessible $I_{\mathcal{X}}$:

Recall: Given $\mathcal{X} = \{ p(x), x \}$

\[
\Lambda_{x_{\mathcal{X}}} = \sum_{x} p(x) \langle x | x_{\mathcal{X}} \rangle \otimes p(x)
\]

\[
S(x_{\mathcal{X}}) = \chi(3)
\]

where $\chi(3) = S(\sum_{x} p(x) \psi_{x}^{c}) - \sum_{x} p(x) S(\psi_{x}^{c})$.

Holevo info for $\mathcal{X}$.

Then: $I_{\text{acc}}(3) \leq \chi(3)$.

Def: Let $\{ M_{y} \}$ be the optimal POVM.

Apply meas to $\mathcal{X}$, keep only classical outcome:

\[
\Lambda_{x_{\mathcal{X}}} \rightarrow \Lambda'_{x_{\mathcal{X}}} = \sum_{x} p(x) \langle x | \alpha_{x} \rangle \otimes \sum_{y} \text{tr}(M_{y} p(x) \psi_{y}^{c}) | y \rangle \langle y |
\]

\[
\chi(3) = S(x_{\mathcal{X}}) \Lambda' \geq S(x_{\mathcal{X}}) = I(x_{\mathcal{X}}) = I_{\text{acc}}(3)
\]

by assumption POVM opt-1.

Arg 1: $g = \frac{1}{2}$, $\chi(3_{1}) = S(p) - 0 = 1$

but $I_{\text{acc}}(3_{1}) = 0.5850$
Lower bound of $I_{\text{acc}}$: [Joza, Robb, Wooten 94]

Def. For a density matrix $\rho$ in $d$-fins with equals $f(x)$, subentropy

$$Q(\rho) := -\frac{d}{\log d} \sum_{x=1}^{d} \left[ -\log \rho_{x} \right] \rho_{x} \log \rho_{x}$$

Then: $\mathcal{E} = \{ p(x), \beta \in \mathbb{R} \}$

$$I_{\text{acc}}(\mathcal{E}) \geq Q \left( \sum_{x=1}^{d} p(x) \rho_{x} \right) - \sum_{x=1}^{d} p(x) Q(\rho_{x})$$

achieved by a $\mathcal{M}$ in random basis.

\[ \text{If } \forall x \text{ pure, } \rho_{x} = \frac{1}{d}, \text{ then } I_{\text{acc}} \geq \log d - \left( \log e \right) \left( \frac{1}{2} + \frac{1}{e} - \frac{1}{e^2} \right) \]

$d=2$, $I_{\text{acc}} \geq 0.2787$
$d=3$, $I_{\text{acc}} \geq 0.60995$
Getting complicated:

\[ \text{Eq. 2: } \mathcal{S}_2 = 14_x \langle 4_x | \mathcal{S}_1 \rangle_{x_1, x_2} \]

\[ \langle 4_x \rangle \text{ as before, } p(x) \text{ as before. } \]

\[ \exists \]

Consider the following measurement \( M_2: \)

1. Apply the measurement Eq. 1 on \( \mathcal{Q}_1 \). If outcome is "a", one of the three states, \( 1\frac{a}{a^2} \), is eliminated.

2. Now state in \( \mathcal{Q}_2 \) is \( 1\frac{b}{b^2} \) up \( \pm \)

\[ 1\frac{c}{c^2} \] up \( \pm \)

\[ \mathcal{b}, \mathcal{c} \in \{0, 1, 2\}; \mathcal{a}, \mathcal{b}, \mathcal{c}. \]

\[ 1\frac{a}{a^2} \left\langle \begin{array}{c} 1\frac{b}{b^2} \\ 1\frac{c}{c^2} \end{array} \right\rangle \left\langle \begin{array}{c} 1\frac{b}{b^2} \\ 1\frac{c}{c^2} \end{array} \right\rangle = \left\langle \begin{array}{c} 1\frac{a}{a^2} + 1\frac{b}{b^2} \\ 1\frac{a}{a^2} + 1\frac{c}{c^2} \end{array} \right\rangle \]

3. Apply measurement along \( \mathcal{a} \) \( \left( \frac{\mathcal{a}}{\mathcal{a}^2} \right) \) basis on \( \mathcal{Q}_2 \).

Claim: \( I(X:Y) \) due to \( M_2 \) is \( \approx 0.23038 \) bits

\[ > 0.5850 \times 2 = 2 \text{ bits } \]

Proof: Assignent 3

Note: meas on \( \mathcal{Q}_2 \) conditioned on meas on \( \mathcal{Q}_1 \).

Most general: joint meas on \( \mathcal{Q}_1, \mathcal{Q}_2 \).
Instead, consider the following measure \( M_3 \):

\[
\Lambda = \lambda_0 x_0 \Theta^2 + \lambda_1 x_1 \Theta^2 + \lambda_2 x_2 \Theta^2 .
\]

\[
M_i = \Lambda^{-1}_i (\Lambda x_i)^\top \Lambda^{-1}_i \text{ for } i = 0, 1, 2 .
\]

\[
M_3 = I - M_0 - M_1 - M_2.
\]

\[
\mathbb{I}(X;Y) \approx 1.369 \text{. } (\text{Assignment 3}) .
\]

Conjecture:

LDCC never achieves IIacc (\( \Sigma_2 \)).

\[\text{measure } \varphi_1, \varphi_2, \varphi_3, \ldots .\]

\[\text{allow subsequent measures to depend on previous outcomes}\]

Will be great if you can prove it, but this is NOT \( \Sigma_3 \).
Remark 1:

- Consider an ensemble $E$ of 1-qubit pure states.

In general, there can be many many different states.

But $I(x) \leq H(E) \leq \log 2 = 1$.

Preparing a state and forgetting the label is highly irreversible

[To make it more amazing telepate 3 qubits.]

[Without entanglement, Alice can at least measure input state and transmit 1 bit to Bob.]

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

\[
\begin{align*}
\hat{x} & \quad \xrightarrow{O} \quad y \\
(\hat{x}, p_x) & = X \quad \xrightarrow{M} \quad Y
\end{align*}
\]

if $Y = X$ w.h.p., $\mathbb{I}(X;Y) \approx \log |\Omega|$. 

\[
\leq H(\{p(x), \beta\})
\]

\[
\leq \log \dim \Omega
\]

- Likewise, data compression requires at least $n \cdot X(E)$ qubits for $n$ draws from $E$. 
Back to how much correlation can change when adding/disconnecting systems:

- We saw \( |S(AB) - S(A)| \leq S(B) \)
  
  \[ \text{Cor 1 (59)} \]

- \( |S(A:BC) - S(A:C)| \leq 2S(B) \)
  
  \[ \text{Cor 2 (59)} \]

- We can extend such results to Holevo info:

  Let \( E = \{ p(x), p_x \}_{x} \)
  
  lives in sys BC

  \( F = \{ p(x), p_x \}_{x} \)
  
  lives in sys C
  
  \( p_x := \text{Tr}_B \rho_x \)

  Then \( |\chi(E) - \chi(F)| \leq 2S(B) \)

  \[ \text{Cor 3 (69)} \]

**Proof:**

\( \rho = \sum_x p(x) \rho(x|x) \otimes \rho_{x:BC} \)

\( \rho = \sum_x \rho_{x:BC} \otimes \rho_{x:BC} \)

Then

\[ |\chi(E) - \chi(F)| \]

\[ = |S(A:BC) - S(A:C)| \leq 2S(B) \]

- So Holevo info can increase by at most \( 2S(B) \)
  
  if B is communicated, even in the presence of shared entanglement.
Thus suppose Alice is allowed to send $n_A$ qubits to Bob, Bob can send $n_B$ qubits back, in any order. Then Alice can communicate at most $n_A n_B$ bits to Bob.

\[ 2n_A \text{ qubit} + n_B \text{ qubit} \rightarrow 3 \text{ qubit} \rightarrow n_A n_B \geq 3. \]

**Proof.** Let Alice's message be $x$.

Suppose after the $3n$th qubit of comm.

Bob's state is $\rho_{xj}$ on sys $B_j$. Let $\Sigma_j = \{ p(x), f_{xj} \}$

Let Bob final output be $y$. 

\[
I(x:y) \leq I(3n) \leq S(\sum_x p(x) p(x|y) p(x|y)) \\
\leq S(\sum_x p(x) p(x|y) p(x|y)) + 1 \\
\vdots \\
\leq n_A + n_B
\]
Lacking of accessible data:

Let \( F = \mathcal{S} p(xt), \mathcal{S}x \mathcal{T} \)

where \( \mathcal{S}x = \Lambda U^t Ix \Lambda x \Lambda U^t + \)

\[ t \in \{0, 1\}, \quad x \in \{0, 1\}, \ldots, n = 3. \]

\[ p(xt) = \frac{1}{2^n}, \]

\[ \Lambda Ix = \sum_{i=0}^{n} w^i \Lambda i, \quad w^n = 1. \]

In 0303088, \( \text{Iacc} (F) = \frac{1}{2} \log n. \)

Now consider:

\[ \mathcal{E} = \{ p(xt), \mathcal{S}xt \in \Lambda txt \}. \]

A particular measurement is:

\[ (t) \quad \overrightarrow{\mathcal{S}t} = t \quad \overrightarrow{\mathcal{S}y} \]

\[ U^t (x) \quad \overrightarrow{U^t} = x \quad \overrightarrow{y} \]

State coming into Bob's lab

\[ \log(2n) = \log(2^n) = n \log n \]

\[ \mathcal{E}(\mathcal{S}) = 1 + \log n, \quad \text{Iacc} (\mathcal{S}) = 1 + \log n. \]

But \( F \) is obtained from \( \mathcal{E} \) by removing 1 qubit,

\[ \mathcal{E}(\mathcal{F}) = \mathcal{E}(\mathcal{S}) - 1 \log n. \]