Quantum Information Theory
- Homework 6 -

Guillermo Abad, David Ullrich

December 24, 2023

Part I

Consider the transmission of a classical random variable X through a classical-quantum channel with pure outputs such that the joint density matrix at the output of the channel is given by:

$$\rho_{XB} = \sum_x p_X(x)|x\rangle\langle x|_X \otimes |\theta_x\rangle\langle\theta_x|_B$$

A measurement POVM is applied to the B share to yield Y:

$$\rho_{XY} = \sum_{x,y} p_X(x)|x\rangle\langle x|_X \otimes \operatorname{tr}\{\Lambda_y|\theta_x\rangle\langle\theta_x|_B\}|y\rangle\langle y|_Y = \sum_{x,y} p_X(x)p_{Y|X}(y \mid x)\Big(|x\rangle\langle x|_X \otimes |y\rangle\langle y|_Y\Big)$$

For the binary, uniformly distributed case, $X \sim \mathrm{Bern}\left(\frac{1}{2}\right), |\mathcal{Y}| = |\mathcal{X}| = 2\dim(\mathcal{H}_B) = 2$ **and**

$$|\theta_0\rangle = \begin{pmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2} \end{pmatrix} \quad ; \quad |\theta_1\rangle = \begin{pmatrix} \cos\frac{\theta}{2} \\ -\sin\frac{\theta}{2} \end{pmatrix}$$

the optimal POVM (in terms of minimizing the error probability) is obtained in Lecture 5 slides and is given by $\{|+\rangle\langle+|, |-\rangle\langle-|\}$. **In this case, the probability of error is** $P_e = \frac{1}{2}(1 - \sin\theta)$**.**

Obtain and plot the accessible information $I(X;Y)$ **and the quantum mutual information** $I(X;B)_\rho$ **for** $\theta \in (0, \pi]$ **.**

Answer

The first part of the exercise asks us to determine $I(X;Y)$ as well as $I(X;B)_\rho$ for the given classical-quantum state before and after a POVM measurement.

We start with the slightly easier part of obtaining the mutual information before the POVM measurement, so $I(X;B)_\rho$:

$$I(X;B)_\rho = H(B)_\rho - H(B|X)_\rho = H(B)_\rho$$

where we have taken advantage of the fact that conditioned on the classical variable the state is pure. To determine $H(B)_\rho$, we need to acquire the ρ_B first by taking the partial trace, which gives:

$$\rho_B = \operatorname{Tr}_X\{\rho_{XB}\} = \frac{1}{2}\Big(|\theta_0\rangle\langle\theta_0| + |\theta_1\rangle\langle\theta_1|\Big) = \begin{bmatrix} \cos^2\left(\frac{\theta}{2}\right) & 0 \\ 0 & \sin^2\left(\frac{\theta}{2}\right) \end{bmatrix}$$

Finally, this leads to:

$$\boxed{I(X;B)_\rho = -\cos^2\left(\frac{\theta}{2}\right)\log\left\{\cos^2\left(\frac{\theta}{2}\right)\right\} - \sin^2\left(\frac{\theta}{2}\right)\log\left\{\sin^2\left(\frac{\theta}{2}\right)\right\}.}$$

Now, to compute $I(X;Y)$, we once again have to compute $H(Y)$ and $H(Y|X)$, but without the luxury of the conditional entropy being zero. On the other hand, we notice two important aspects. Firstly, that $p_Y(0) = p_Y(1)$, since:

$$p_Y(0) = \frac{1}{2}\Big(p_{Y|X}(0|0) + p_{Y|X}(0|1)\Big) = \frac{1}{2}\Big(1 - P_e + P_e\Big) = \frac{1}{2}\Big(p_{Y|X}(1|0) + p_{Y|X}(1|1)\Big) = \frac{1}{2} = p_Y(1)$$

which leads to:

$$H(Y) = -\sum_y p_Y(y)\log p_Y(y) = -\frac{1}{2}\log\frac{1}{2} - \frac{1}{2}\log\frac{1}{2} = 1.$$

And secondly, due to the symmetry of the channel, we can express the conditional entropy simply by using the given error probability, $P_e = \frac{1}{2}(1-\sin\theta)$, as:

$$H(Y|X) = H(Y|x=0)p_X(x=0) - H(Y|x=1)p_X(x=1) = -(1-P_e)\log(1-P_e) - P_e\log P_e$$

This eventually gives us the following expression for the mutual information

$$\boxed{I(X;Y) = 1 + \left(1 - \frac{1}{2}(1-\sin\theta)\right)\log\left\{1 - \frac{1}{2}(1-\sin\theta)\right\} + \frac{1}{2}(1-\sin\theta)\log\left\{\frac{1}{2}(1-\sin\theta)\right\}}$$

Notice the similarities to the classical BSC:

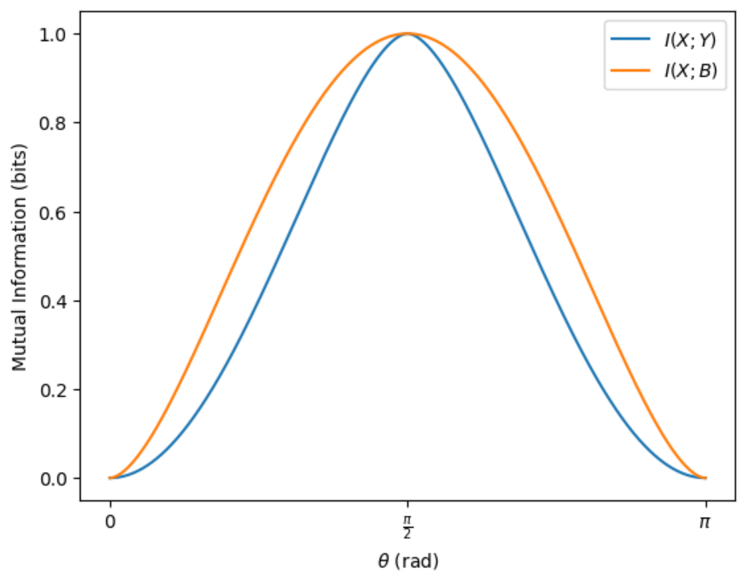

Figure 1: Mutual information $I(X;Y)$ and $I(X;B)_\rho$, as a function of the angle $\theta \in (0,\pi]$. Both are maximal for $\theta = \pi/2$ and minimal in the extremes.

As we can see $I(X;B)_\rho \geq I(X;Y)$ which is in accordance of the data processing inequality.

Both entities reach their maximum value at $\theta = \pi/2$, for which the states are $\{|+\rangle\langle+|, |-\rangle\langle-|\}$ proving as we knew, that those are the optimal POVMs. At this point, the error probability is zero and the channel is thereby noiseless, giving us the same value for $I(X;B)_\rho$ as well as $I(X;Y)$.

On the other hand, when the angle between the two states is 0 or π the two states cannot be distinguished $\Big(\{|0\rangle\langle 0|, |0\rangle\langle 0|\}$ or $\{|1\rangle\langle 1|, -|1\rangle\langle 1|\}\Big)$ leading to $I(X;Y) = I(X;B) = 0$.

The last thing to notice is that the maximum difference is given around an angle of $\pi/4$ when the states are transitioning from orthogonal to parallel.

Part II

We now consider $X \sim \text{Unif}([0,1,2,3]), |\mathcal{Y}| = |\mathcal{X}| = 4$. We use now three parallel quantum channels such that

$$\rho_{XB^3} = \sum_x p_X(x)|x\rangle\langle x|_X \otimes |\psi_x\rangle\langle\psi_x|_{B^3}$$

where

$$\begin{aligned}
|\psi_0\rangle_{B^3} &= |\theta_0\rangle_B \otimes |\theta_0\rangle_B \otimes |\theta_0\rangle_B := |\theta_{000}\rangle_B \\
|\psi_1\rangle_{B^3} &= |\theta_0\rangle_B \otimes |\theta_1\rangle_B \otimes |\theta_1\rangle_B := |\theta_{011}\rangle_B \\
|\psi_2\rangle_{B^3} &= |\theta_1\rangle_B \otimes |\theta_0\rangle_B \otimes |\theta_1\rangle_B := |\theta_{101}\rangle_B \\
|\psi_3\rangle_{B^3} &= |\theta_1\rangle_B \otimes |\theta_1\rangle_B \otimes |\theta_0\rangle_B := |\theta_{110}\rangle_B
\end{aligned}$$

Again, a measurement POVM is applied to the B^3 share to yield Y :

$$\rho_{XY} = \sum_{x,y} p_X(x)|x\rangle\langle x|_X \otimes \text{tr}\{\Lambda_y|\psi_x\rangle\langle\psi_x|_{B^3}\}|y\rangle\langle y|_Y = \sum_{x,y} p_X(x)p_{Y|X}(y \mid x)\ |x\rangle\langle x|_X \otimes |y\rangle\langle y|_Y$$

In this case, the optimal POVM is known as the square-root measurement, whose elements are of the form

$$\Lambda_y = \frac{1}{4}\left(\rho_{B^3}\right)^{-\frac{1}{2}}|\psi_y\rangle\langle\psi_y|\left(\rho_{B^3}\right)^{-\frac{1}{2}},\ \text{for } y \in [0,1,2,3]$$

and where $\rho_{B^3} = \text{tr}_X\{\rho_{XB^3}\}$.

Show that Λ_y is a proper POVM. Obtain (numerically) and plot the accessible information $I_3(X;Y)$ and the quantum mutual information $I_3(X;B^3)_\rho$ for $\theta \in (0,\pi]$.

Answer

We're now asked to show that the given square-root measurement is in fact a POVM.

$$\Lambda_y = \frac{1}{4}\left(\rho_{B^3}\right)^{-\frac{1}{2}}|\psi_y\rangle\langle\psi_y|\left(\rho_{B^3}\right)^{-\frac{1}{2}},\ \text{for } y \in [0,1,2,3]$$

For this we need to show that $\Lambda_y \geq 0\ \forall y$ and $\sum_y \Lambda_y = \mathbb{1}$. The positivity becomes obvious when realizing that every element consists of a projector namely $|\psi_y\rangle\langle\psi_y|$ and a positive maps acting on it from each side (The positive map results from the fact that ρ_{B^3} is a density matrix and thereby positive as well as its negative square root). This means that each Λ_y can be seen as a projector itself and the positivity becomes trivial.

To check that they sum up to the identity, we need to compute ρ_{B^3} first:

$$\rho_{B^3} = \text{tr}_X\{\rho_{XB^3}\} = \sum_x p_X(x)\,|\psi_x\rangle\langle\psi_x|_{B^3} = \frac{1}{4}\sum_x |\psi_x\rangle\langle\psi_x|_{B^3}$$

With this, we see that:

$$\sum_y \Lambda_y = \sum_y \frac{1}{4}\left(\rho_{B^3}\right)^{-\frac{1}{2}}|\psi_y\rangle\langle\psi_y|\left(\rho_{B^3}\right)^{-\frac{1}{2}} = \left(\rho_{B^3}\right)^{-\frac{1}{2}}\underbrace{\left(\frac{1}{4}\sum_y|\psi_y\rangle\langle\psi_y|\right)}_{\rho_{B^3}}\left(\rho_{B^3}\right)^{-\frac{1}{2}} = \mathbb{1}$$

For the quantum mutual information, the task is similar compared to Part I, because conditioned on X the state of B^3 is pure. By using the following expression for the mutual information, we get:

$$I_3\left(X;B^3\right) = H_3\left(B^3\right)_\rho - H_3\left(B^3|X\right)_\rho = H_3\left(B^3\right)_\rho = -\operatorname{tr}\left\{\rho_{B^3}\log\rho_{B^3}\right\}$$

So we just need to diagonalize ρ_{B^3}. Using the expression of ρ_{B^3} from the previous proof and writing it out explicitly, we get:

$$\rho_{B^3} = \frac{1}{4}\sum_x |\psi_x\rangle\langle\psi_x| = \left(\begin{array}{cccccccc} c^6 & 0 & 0 & 0 & 0 & 0 & 0 & c^3s^3 \\ 0 & c^4s^2 & 0 & 0 & 0 & 0 & c^3s^3 & 0 \\ 0 & 0 & c^4s^2 & 0 & 0 & c^3s^3 & 0 & 0 \\ 0 & 0 & 0 & c^2s^4 & c^3s^3 & 0 & 0 & 0 \\ 0 & 0 & 0 & c^3s^3 & c^4s^2 & 0 & 0 & 0 \\ 0 & 0 & c^3s^3 & 0 & 0 & c^2s^4 & 0 & 0 \\ 0 & c^3s^3 & 0 & 0 & 0 & 0 & c^2s^4 & 0 \\ c^3s^3 & 0 & 0 & 0 & 0 & 0 & 0 & s^6 \end{array}\right)$$

where we have defined $c := \cos\left(\frac{\theta}{2}\right)$ and $s := \sin\left(\frac{\theta}{2}\right)$. It's easy to see, that we just have 2×2 subspaces with 2 different matrices whose eigenvalues we can easily obtain.

The first one with multiplicity 1:

$$\left(\begin{array}{cc} c^6 & c^3s^3 \\ c^3s^3 & s^6 \end{array}\right) \Rightarrow \lambda = \left\{\begin{array}{l} 0 \\ c^6+s^6 = 1-\frac{3}{4}\sin^2\theta \end{array}\right.$$

The second one with multiplicity 3:

$$\left(\begin{array}{cc} c^4s^2 & c^3s^3 \\ c^3s^3 & c^2s^4 \end{array}\right) \Rightarrow \lambda = \left\{\begin{array}{l} 0 \\ c^4s^2+c^2s^4 = \frac{1}{4}\sin^2\theta \end{array}\right.$$

So the respective eigenvalues with multiplicities 4, 1 and 3 are:

$$\lambda = \left\{0\ ;\ 1-\frac{3}{4}\sin^2\theta\ ;\ \frac{1}{4}\sin^2\theta\right\}$$

Our final result is then:

$$\boxed{I_3\left(X;B^3\right)_\rho = -\left(1-\frac{3}{4}\sin^2\theta\right)\log\left(1-\frac{3}{4}\sin^2\theta\right) - \frac{3}{4}\sin^2\theta\log\left(\frac{1}{4}\sin^2\theta\right)}$$

Now, to compute accessible information we can once again apply the definition from before

$$I_3(X;Y) = H_3(Y) - H_3(Y|X)$$

which for the individual entropies, ρ_Y:

$$\rho_Y = \operatorname{tr}_X\left\{\rho_{XY}\right\} = \frac{1}{4}\sum_{x,y} p_X(x)p_{Y|X}(y|x)|y\rangle\langle y|$$

have the form:

$$H_3(Y|X) = \sum_x p_X(x)\cdot H(Y|X=x) = -\frac{1}{4}\sum_x p_X(x)\sum_y p_{Y|X}(y|x)\log p(y|x)$$

$$H_3(Y) = -\operatorname{tr}\left\{\rho_Y\log\left(\rho_Y\right)\right\} = -\frac{1}{4}\sum_{x,y} p_X(x)p_{Y|X}(y|x)\log\left(\sum_{x'} p_X(x')p_{Y|X}(y|x')\right)$$

If we find the marginal distribution of $p_Y(y)$, we can at least find $H_3(Y)$ algebraically, leaving only $H_3(Y|X)$ for posterior numerical computations:

$$
\begin{aligned}
p_Y(y) = \sum_x p_X(x) p_{Y|X}(y \mid x) &= \frac{1}{4}\sum_x \operatorname{tr}\left\{\Lambda_y \left|\psi_x\right\rangle\left\langle\psi_x\right|\right\} = \operatorname{tr}\left\{\Lambda_y\left(\frac{1}{4}\sum_x \left|\psi_x\right\rangle\left\langle\psi_x\right|\right)\right\} = \\
&= \operatorname{tr}\left\{\Lambda_y \rho_{B^3}\right\} == \frac{1}{4}\operatorname{tr}\left\{\left(\rho_{B^3}\right)^{-\frac{1}{2}}\left|\psi_y\right\rangle\left\langle\psi_y\right|\left(\rho_{B^3}\right)^{-\frac{1}{2}}\left(\rho_{B^3}\right)^{\frac{1}{2}}\left(\rho_{B^3}\right)^{\frac{1}{2}}\right\} = \\
&= \frac{1}{4}\operatorname{tr}\left\{\left(\rho_{B^3}\right)^{\frac{1}{2}}\left(\rho_{B^3}\right)^{-\frac{1}{2}}\left|\psi_y\right\rangle\left\langle\psi_y\right|\left(\rho_{B^3}\right)^{-\frac{1}{2}}\left(\rho_{B^3}\right)^{\frac{1}{2}}\right\} = \frac{1}{4}\operatorname{tr}\left\{\left|\psi_y\right\rangle\left\langle\psi_y\right|\right\} = \frac{1}{4}
\end{aligned}
$$

from where we see that $Y \sim \operatorname{Unif}([0,1,2,3]), |\mathcal{Y}| = 4$, and therefore its entropy is $H_3(Y) = \log(4) = 2$. Obtaining a finally accessible information:

$$
\boxed{I_3(X;Y) = 2 + \frac{1}{4}\sum_x p_X(x)\sum_y p_{Y|X}(y|x)\log p(y|x)}
$$

To compute this we need to determine all the conditional entropies as well as diagonalizing ρ_Y. We do both numerically (Python) and plot the result for the accessible information together with the quantum mutual information which we discussed before.

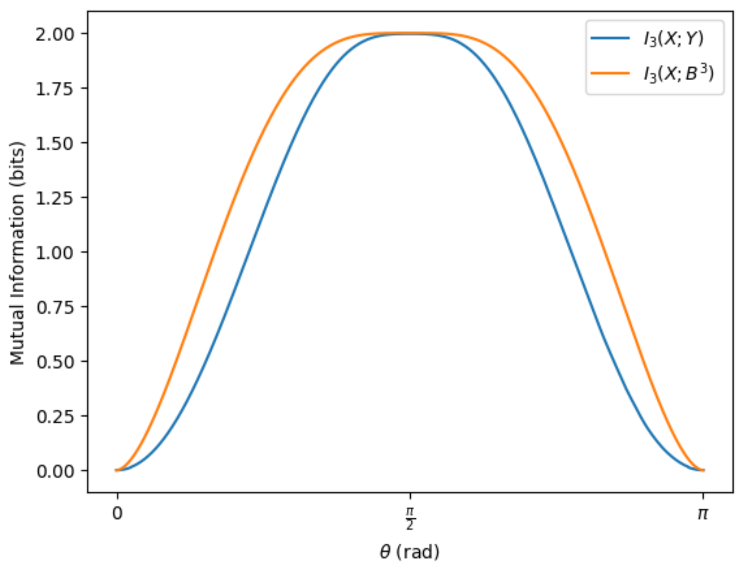

Figure 2: Mutual information $I_3(X;Y)$ and $I_3(X;B^3)$, as a function of the angle $\theta \in [0,\pi]$. Both are maximal for $\theta = \pi/2$ and minimal in the extremes.

to our satisfaction, we can see that $I_3(X,B^3)_\rho \geq I_3(X;Y)$ which is once again in accordance with the processing inequality.

And again, for completely indistinguishable states ($\theta = 0$ or π) the mutual information is zero for both as one would expect, and when the states are orthogonal ($\theta = \pi/2$) reaches the maximum value as expected too, this time of 2 bits. This works out because if both $|\theta_i\rangle$ states are the same, implies that the four $|\psi_i\rangle$ states will also be the same, and on the other hand, if both $|\theta_i\rangle$ states are orthogonal that also implies that the four $|\psi_i\rangle$ states are orthogonal.

Part III

Finally plot $I_3(X;Y) - 3I(X;Y)$ and $I_3(X;B^3)_\rho - 3I(X;B)_\rho$ for $\theta \in (0, \pi]$. Analyze and discuss the results.

Answer

Last but not least, we look at the additivity of these two different settings. The graph for the relationship of $I_3(X;B^3)_\rho - I(X,B)_\rho$ shows negative values for the full range of theta. The quantum mutual information is thereby subadditive. As can be seen in the corresponding graph, the maximal value of $I_3(X;B^3)_\rho$ is only 2 bits. This leads to a difference of one bit in comparison to $3I(X,B)_\rho$ (maximal of 1 bit, times 3) at $\theta = \frac{\pi}{2}$.

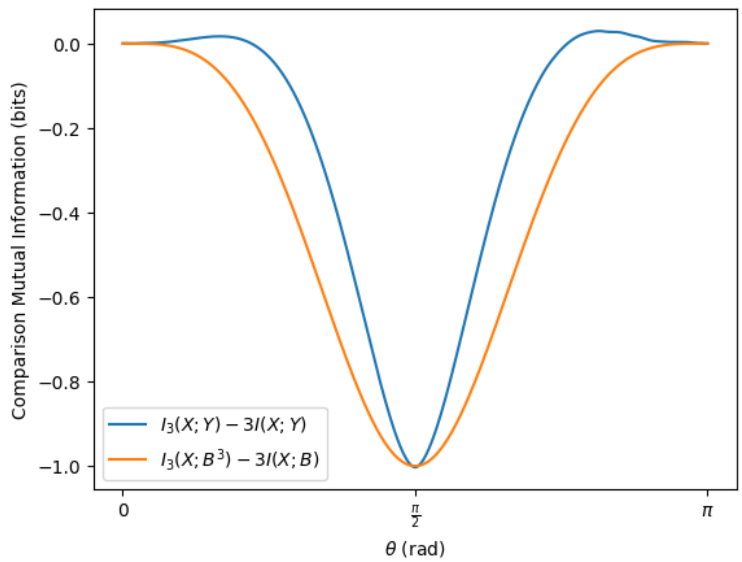

Figure 3: Comparison of the mutual information between the use of three parallel channels and a single channel thrice, as a function of the angle $\theta \in [0, \pi]$. The difference again is maximal for $\theta = \pi/2$ and minimal in the extremes.

This relationship is not preserved after the application of the POVM and the accessible information $I_3(X;Y) - 3I(X;Y)$, for a small range of input values, is greater than zero implying an advantage in choosing the three parallel channels over the repeated usage of one channel three times in those small ranges.

To end, notice that as expected in the extremes (where θ is 0 or π), we don't have any difference, since all cases have 0 mutual information being the states parallel. And likewise, the maximum we commented on before, is located in the middle (for $\theta = \pi/2$) as expected too, where the states are orthogonal.