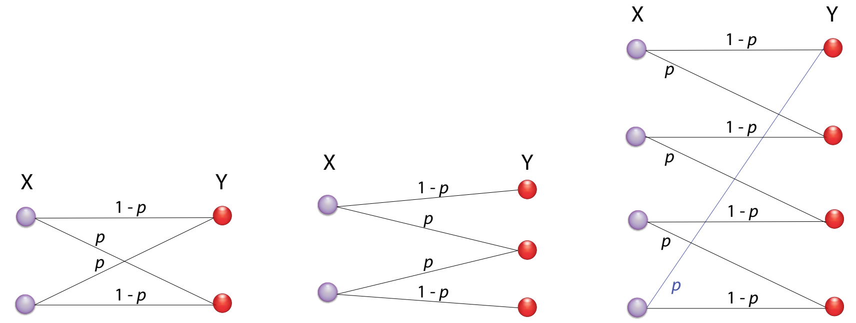
Quantum Information Theory
Solutions 3

HS 12
Prof. R. Renner

Exercise 3.1 Channel capacity



(a) Binary Symmetric Channel

(b) Symmetric Erasure Channel

(c) Yet Another Channel

a) *The asymptotic channel capacity is given by $C = \max_{P_X} I(X:Y)$. Calculate the asymptotic capacities of the first two channels depicted above.*

The capacity of the binary symmetric channel evaluates to

$$\begin{aligned}
C &= \max_{P_X} I(X:Y) = \max_{P_X} H(Y) - H(Y|X) \\
&\qquad\qquad\quad = \max_{P_X} H(Y) + \sum_{x,y} P_X(x) P_{Y|X=x}(y) \log P_{Y|X=x}(y) \\
&\qquad\qquad\quad = \max_{P_X} H(Y) - \sum_{x} P_X(x)\ H_{\text{bin}}(p) \\
&\qquad\qquad\quad = \max_{P_X} H(Y) - \ H_{\text{bin}}(p) \\
&\qquad\qquad\quad = 1 - H_{\text{bin}}(p),
\end{aligned}$$

where $H_{\text{bin}}(p)$ is the *binary entropy*, i.e. the entropy of the probability distribution $(p, 1-p)$,

$$H_{\text{bin}}(p) = p \log \frac{1}{p} + (1-p) \log \frac{1}{1-p}.$$

To maximise $H(Y)$ we chose the uniform distribution on the input, $P_X^u = (\frac{1}{2}, \frac{1}{2})$ (see part *b*).

Similarly, for the symmetric erasure channel, we have

$$\begin{aligned}
C &= \max_{P_X} H(Y) - H(Y|X) \\
&= \max_{P_X} H(Y) - H_{\text{bin}}(p) \\
&= 2\frac{1-p}{2} \log \frac{2}{1-p} + p \log \frac{1}{p} - H_{\text{bin}}(p) \\
&= 1 - p.
\end{aligned}$$

b) *Consider N possible probability distributions as input to a general channel, $\{P_X^i\}_i$, with the property that $I(X:Y)_{P^i} = I(X:Y)_{P^j}, \forall i,j$. Suppose you choose which distribution to use for the input by checking a random variable, B, with possible values $b = \{1, \ldots, N\}$. Show that $I(X:Y|B) \leq I(X:Y)$.*

We have

$$\begin{aligned} I(X:Y|B) &= H(Y|B) - H(Y|XB) \\ &= H(Y|B) - H(Y|X) \qquad {}^{(*)} \\ &\leq H(Y) - H(Y|X) \qquad {}^{(**)} \\ &= I(X:Y), \end{aligned}$$

where $^{(*)}$ stands because B is just a label that tells us which probability distribution P_X^i we used, so knowing X is as good as knowing X and B, in the sense that $H(Y|XB) = H(Y|X)$, and $^{(**)}$ comes from the data-processing inequality, $H(Y|B) \leq H(Y)$ (which in lay terms says that extra information cannot hurt).

How can you use that to find the probability distribution P_X that maximises the mutual information for symmetric channels? **Hint:** *consider $\{P_X^i\}_i$ permutations of P_X^1.*

For symmetric channels, the mutual information between input and output is invariant under permutation of the inputs (that's how they are defined). Look for instance at the symmetric erasure channel. The input distribution $P_X^1 = (0.75, 0.25)$ yields the same mutual information as $P_X^2 = (0.25, 0.75)$.

Not knowing which permutation of P_X was used in the input is equivalent to take a uniform mixture over all possible permutations of P_X. Conveniently, such mixture gives us the uniform distribution:

$$P_X = \begin{pmatrix} P_X(x_1) \\ P_X(x_2) \\ \vdots \\ P_X(x_N) \end{pmatrix}, \qquad \{P_X^i\}_{i=1,\dots,N!} \text{ permutations of } P_X,$$

$$\sum_i^{N!} \frac{1}{N!} P_X^i = \frac{1}{N!} \begin{pmatrix} (N-1)!\ P_X(x_1) + (N-1)!\ P_X(x_2) + \cdots + (N-1)!\ P_X(x_N) \\ (N-1)!\ P_X(x_1) + (N-1)!\ P_X(x_2) + \cdots + (N-1)!\ P_X(x_N) \\ \vdots \\ (N-1)!\ P_X(x_1) + (N-1)!\ P_X(x_2) + \cdots + (N-1)!\ P_X(x_N) \end{pmatrix}$$

$$= \frac{1}{N} \begin{pmatrix} \sum_i P_X(x_i) \\ \sum_i P_X(x_i) \\ \vdots \\ \sum_i P_X(x_i) \end{pmatrix} = \begin{pmatrix} 1/N \\ 1/N \\ \vdots \\ 1/N \end{pmatrix}.$$

This means that for any input distribution P_X the mutual information always increases if instead you use the uniform distribution. Here, $I(X:Y|B)$ is the mutual information knowing you used P_X (take $B=1$), and $I(X:Y)$ is the mutual information for the uniform distribution. Conclusion: for symmetric channels the mutual information is maximised if one takes the uniform distribution as input.

c) Using part *b*), we choose the uniform distribution on X and calculate the capacity:

$$\begin{aligned} C &= \max_{P_X} I(X:Y) = I(X:Y)_{P_X^u} = H(Y) - H(Y|X) \\ &= -\sum_y P_Y(y) \log P_Y(y) + \sum_{x,y} P_{XY}(x,y) \log P_{Y|X}(y) \\ &= -\sum_y \left(\sum_x P_X(x) P_{Y|X=x}(y) \right) \log \left(\sum_x P_X(x) P_{Y|X=x}(y) \right) + \sum_{x,y} P_X(x) P_{Y|X=x}(y) \log P_{Y|X=x}(y) \\ &= -4 \cdot \frac{1}{4} \log \left(\frac{1}{4} \right) + 4 \cdot \frac{1}{4} H_{\text{bin}}(p) = 2 - H_{\text{bin}}(p). \end{aligned}$$

Exercise 3.2 Smooth min-entropy in the i.i.d. limit

Use the weak law of large numbers to show that the smooth min-entropy converges to the Shannon entropy $H(X)$ in the i.i.d. limit:

$$\lim_{\epsilon\to 0}\lim_{n\to\infty}\frac{1}{n}H^{\epsilon}_{\min}(\vec{X})_{P^n} = H(X)_P. \tag{1}$$

The weak law of large numbers as shown in exercise series 1 is given by:

$$\lim_{n\to\infty} P\left[\left(\frac{1}{n}\sum_i S_i - \mu\right)^2 \geq \eta\right] = 0 \quad \text{for any } \eta > 0,\ \mu = \mathcal{E}[S].$$

Setting $S_i = h_P(x_i) = -\log P_X(x_i)$ we get $\mu = H(X)$ and thus

$$\lim_{n\to\infty} P\left[\left|\frac{1}{n}\sum_i h_P(x_i) - H(X)\right| < \nu\right] = 1 \quad \text{for any } \nu > 0.$$

This knowledge allows us to restrict the set of vectors $\vec{x}$ to typical outcomes, namely we introduce a subset $\mathcal{G}_\nu$ of $\mathcal{X}^{\times n}$:

$$\mathcal{G}_\nu = \left\{\vec{x}\in\mathcal{X}^{\times n} : \left|\frac{1}{n}\sum_i h_P(x_i) - H(X)\right| < \nu\right\}.$$

The weak law of large numbers can now be restated simply as

$$\lim_{n\to\infty} P_{\vec{X}}[\mathcal{G}_\nu] = \lim_{n\to\infty} P_{\vec{X}}[\vec{x}\in\mathcal{G}_\nu] = 1.$$

Furthermore, let $\mathcal{G}^c_\nu$ denote the complement of $\mathcal{G}_\nu$ in $\mathcal{X}^{\times n}$. As a next step we choose

$$Q_{\vec{X}}(\vec{x}) = \begin{cases} P_{\vec{X}}(\vec{x})/P_{\vec{X}}[\mathcal{G}_\nu] & \text{if } \vec{x}\in\mathcal{G}_\nu \\ 0 & \text{if } \vec{x}\in\mathcal{G}^c_\nu \end{cases}.$$

This distribution has the property that for any fixed ν, the trace distance $\delta(P_{\vec{X}}, Q_{\vec{X}})$ vanishes in the limit of $n\to\infty$. This can be seen when we use the alternative definition of δ introduced in problem set 1 and when we take the probabilities over the set $\mathcal{G}^c_\nu$ of all events where $Q_{\vec{X}} < P_{\vec{X}}$:

$$\lim_{n\to\infty}\delta(P_{\vec{X}}, Q_{\vec{X}}) = \lim_{n\to\infty} P_{\vec{X}}[\mathcal{G}^c_\nu] - Q_{\vec{X}}[\mathcal{G}^c_\nu] = 0.$$

In particular, we can now evaluate the "smooth" min-entropy for any fixed $\epsilon > 0$ and $\nu > 0$:

$$\begin{aligned}
\lim_{n\to\infty}\frac{1}{n}H^{\epsilon}_{\min}(\vec{X}) &\geq \lim_{n\to\infty}\min_{\vec{x}\in\mathcal{X}^{\times n}}\frac{1}{n}h_Q(\vec{x}) \\
&= \lim_{n\to\infty}\min_{\vec{x}\in\mathcal{G}_\nu}\frac{1}{n}h_P(\vec{x}) + \lim_{n\to\infty}\frac{1}{n}\log P_{\vec{X}}[\mathcal{G}_\nu] \\
&= \lim_{n\to\infty}\min_{\vec{x}\in\mathcal{G}_\nu}\frac{1}{n}\sum_i h_P(x_i) \\
&\geq H(X) - \nu
\end{aligned}$$

The first inequality is a consequence of the fact that our $Q_{\vec{X}}$ is not necessarily optimal (as a matter of fact, it could be shown that it actually is). It follows that this construction only gives us a lower bound on the i.i.d. limit once we set ν arbitrarily close to zero and let $\epsilon\to 0$. However, from the definition of Shannon and min-entropy it follows directly that the min-entropy can never exceed Shannon entropy, since the information gain in the worst-case can never be higher than the average information gain.

Exercise 3.3 Quantum-Telepathy Game: Introduction

a) *i)* *Find projective measurements that the players can perform so that they always get opposite outcomes x_1 and x_2, and therefore can use their outcomes to win the game.*

We want to probability of getting the same outcome to be zero. From this observation we see that each player using a measurement in the basis $\{|+\rangle, |-\rangle\}$ gives

$$p(++) = |\langle ++|\phi\rangle|^2 = 0, \quad p(+-) = |\langle +-|\phi\rangle|^2 = 1/2$$
$$p(-+) = |\langle -+|\phi\rangle|^2 = 1/2, \quad p(--) = |\langle --|\phi\rangle|^2 = 0,$$

as desired.

ii) *Explain how this game can be won without using $|\phi\rangle$.*

The players could agree in the beginning to set $x_1 = 0$ and $x_2 = 1$, and just output these values when they get their qubits.

b) *i)* *First, rewrite the state $|\phi\rangle$ in the computational basis ($\{|0\rangle, |1\rangle\}$ for each qubit).*

By substitution we find that

$$|\phi\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle).$$

ii) *What projective measurement should P_3 do so that after one of the outcomes of the measurement (he chooses $b = 0$ for this outcome), the other two players are left with the state $|\phi\rangle$ from part (a)?*

Note that

$$\sqrt{2}\mathbb{1}^{\otimes 2} \otimes \langle +|_3|\Psi^3_-\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle) = |\phi\rangle.$$

Therefore P_3 should measure in the basis $\{|+\rangle, |+^\perp\rangle = |-\rangle\}$.

iii) *What is the state $|\psi\rangle$ that P_1 and P_2 share after P_3 gets the other outcome ($b = 1$)? Write $|\psi\rangle$ in the basis $\{|\circlearrowright\rangle = (|0\rangle + i|1\rangle)/\sqrt{2}, |\circlearrowleft\rangle = (|0\rangle - i|1\rangle)/\sqrt{2}\}$.*

We just need to calculate

$$\sqrt{2}(\mathbb{1}^{\otimes 2} \otimes \langle -|_3)|\Psi^3_-\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) = |\psi\rangle.$$

To write this in the circular basis we first can write

$$|0\rangle = \frac{1}{\sqrt{2}}(|\circlearrowright\rangle + |\circlearrowleft\rangle), \quad |1\rangle = \frac{1}{\sqrt{2}i}(|\circlearrowright\rangle - |\circlearrowleft\rangle).$$

Substituting this into $|\psi\rangle$, we get

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|\circlearrowright\circlearrowleft\rangle + |\circlearrowleft\circlearrowright\rangle).$$

iv) *What projective measurements do P_1 and P_2 do in order to get different results from the state $|\psi\rangle$?*

This is the same situation as in part (a) (i), except in a different basis. Therefore measuring in the basis $\{|\circlearrowright\rangle, |\circlearrowleft\rangle\}$ will never give P_1 and P_2 the same result.

Exercise 3.4 Quantum-Telepathy Game: The Full Story

a) *Use the same measurement you found in 3.2 (*b*) (*ii*) to find the possible results of $\mathcal{M}_n(|\Psi^n_\pm\rangle)$. Specifically, find $M^0_n(|\Psi^n_+\rangle)$, $M^1_n(|\Psi^n_+\rangle)$, $M^0_n(|\Psi^n_-\rangle)$, and $M^1_n(|\Psi^n_-\rangle)$, where $\mathcal{M}^{0,1}$ denote the different measurements performed.*

First, let the superscript 0 and 1 on the measurements correspond to the projection onto $|+\rangle$ and $|-\rangle$ respectively. Then

$$\begin{aligned}
M_n^0(|\Psi_+^n\rangle) &= \sqrt{2}(\mathbb{1}^{\otimes n-1} \otimes \langle +|_n)|\Psi_+^n\rangle = |\Psi_+^{n-1}\rangle,\\
M_n^1(|\Psi_+^n\rangle) &= \sqrt{2}(\mathbb{1}^{\otimes n-1} \otimes \langle -|_n)|\Psi_+^n\rangle = |\Psi_-^{n-1}\rangle,\\
M_n^0(|\Psi_-^n\rangle) &= \sqrt{2}(\mathbb{1}^{\otimes n-1} \otimes \langle +|_n)|\Psi_-^n\rangle = |\Psi_-^{n-1}\rangle,\\
M_n^1(|\Psi_-^n\rangle) &= \sqrt{2}(\mathbb{1}^{\otimes n-1} \otimes \langle -|_n)|\Psi_-^n\rangle = |\Psi_+^{n-1}\rangle.
\end{aligned}$$

b) *Given the above results, work out a detailed quantum strategy that always wins this game.*

All n players start with the state $|\Psi\rangle = |\Psi_-^n\rangle$. The $n-2$ players keep track of the number of negative $|-\rangle$ measurement outcomes. The parity of this number is then sent to the two selected players ($b = 0$ for even and $b = 1$ for odd). If they receive $b = 0$, they know that they share the state $|\Psi_-^2\rangle$ among each other and a measurement in the $\{|+\rangle, |-\rangle\}$ basis is guaranteed to give them different results. They agree beforehand that the $|+\rangle$ outcome corresponds to a 0, and the $|-\rangle$ outcome corresponds to a 1. If they receive $b = 1$, they share the state $|\Psi_+^2\rangle$ and the measurement will be done in the $\{|\circlearrowleft\rangle, |\circlearrowright\rangle\}$ basis. In this case, the $|\circlearrowright\rangle$ outcome corresponds to a 0, and the $|\circlearrowleft\rangle$ outcome corresponds to a 1. The game is won in both cases.

Note that you could equivalently start with $|\Psi_+^n\rangle$ instead, in which case the parity of the number of negative outcomes $|-\rangle$ determines b instead by: $b = 0$ for odd and $b = 1$ for even. The rest of the strategy is the same as above.