1. **Classical capacities of quantum channels**

Although this exercise might look very long, it isn't. In the next paragraphs we just want to give you an overview on the formalism introduced in the lecture and needed for this exercise in a compressed fashion. No need to be intimidated ;)

In the lecture, we saw two alternative characterisations of the classical channel capacity of a quantum channel $\mathcal{E}$, which is given by the its Holevo-information $\chi(\mathcal{E})$. The task here is to establish the equivalence of these expressions.

To this end, recall the definition of the quantum mutual information of a bi-partite quantum system in a state ρ_{AB}

$$I(A:B)_{\rho_{AB}} := S(\rho_A) + S(\rho_B) - S(\rho_{AB}). \tag{1}$$

The Holevo information of channel can be defined using the following scheme: Alices encodes the information of a classical random variable X taking values in $\mathcal{X}$ with probability distribution p_X into a quantum state using a set of states $\{\rho_x\}_{x\in\mathcal{X}}$. To keep track of the classical random variable but formulating everything quantum mechanically, we think of Alice encoding the result in another faithfully register N using orthogonal basis $\{|x\rangle\}_{x\in\mathcal{X}}$. From this notebook register N the classical information of X can be completely recovered. Altogether, Alice prepares the bi-partite state

$$\rho_{NA} = \sum_x p_X(x)\, |x\rangle\langle x|_N \otimes \rho_A^x. \tag{2}$$

Then, the state in system A is sent to Bob using the channel $\mathcal{E}$. Thus, we end up with a final state shared between Alice's notebook and Bob

$$\rho_{NB} = \sum_x p_X(x)\, |x\rangle\langle x|_N \otimes \mathcal{E}(\rho_A^x)_B. \tag{3}$$

We can now ask for the mutual information between the variable X encoded in N and Bob's output of the channel. Analogously to the classical result, maximizing the mutual information over all possible input variables X and encodings yields the capacity of the quantum channel to transmit classical informations, i.e.

$$\chi(\mathcal{E}) = \max_{(X,p_X,\{\rho^x\})} I(N,B)_{\rho_{NA}}. \tag{4}$$

a) Show that

$$\chi(\mathcal{E}) = \max_{(X,p_X,\{\rho^x\})} \left\{ S(\mathcal{E}(\sum_x p_X(x)\rho^x)) - \sum_x p_X(x) S(\mathcal{E}(\rho^x)) \right\}. \tag{5}$$

$$I(N,B) = S(\underbrace{\mathrm{tr}_B(\rho_{NB})}_{\overset{!}{=}\,\rho_N}) + S(\underbrace{\mathrm{tr}_N(\rho_{NB})}_{\overset{!}{=}\,\rho_B}) - S(\rho_{NB})$$

with $\rho_{NB} := \sum_{x\in X} p_X(x)\, |x\rangle\langle x|_N \otimes \mathcal{E}(\rho_A^x)$

- $$\rho_N = \sum_{x \in X} p_x(x) \, |x\rangle\langle x|_N \, \mathrm{tr}\left(\mathcal{E}(\rho_A^x)\right)$$

$$= \sum_{x \in X} p_x(x) \, |x\rangle\langle x|_N$$

$\mathrm{tr}\left(\mathcal{E}(\rho_A^x)\right) \underset{\uparrow}{=} \mathrm{tr}\left(\rho_A^x\right) = 1$

$\mathcal{E}$ - trace preserving

$$S(\rho_N) \underset{\uparrow}{=} H(X)$$

$p_x(x)$ eigenvalues of ρ_N.

- $$\rho_B = \sum_{x \in X} p_x(x) \, \underbrace{\mathrm{tr}\left(|x\rangle\langle x|_N\right)}_{=1} \, \mathcal{E}(\rho_A^x)$$

$$= \sum_{x \in X} p_x(x) \, \mathcal{E}(\rho_A^x) \underset{\uparrow}{=} \mathcal{E}\left(\sum_x p_x(x) \, \rho_A^x\right)$$

$\mathcal{E}(\cdot)$ linear

$$S(\rho_B) = S\left(\mathcal{E}\left(\sum_x p_x(x) \, \rho_A^x\right)\right)$$

- $$\rho_{NB} := \sum_{x \in X} p_x(x) \, |x\rangle\langle x|_N \otimes \mathcal{E}(\rho_A^x) \underset{\uparrow}{=} \sum_{x,s} p_x(x) \, \lambda_s^{(x)} \, |x\rangle\langle x| \otimes |\lambda_s^{(x)}\rangle\langle \lambda_s^{(x)}|$$

$\mathcal{E}(\rho_A^x) = \sum_s \lambda_s^{(x)} |\lambda_s^{(x)}\rangle\langle\lambda_s^{(x)}|$

eigendec.

$$S(\rho_{NB}) \underset{\uparrow}{=} H\left(\left\{p_x(x) \, \lambda_s^{(x)}\right\}\right) = H(X) + \sum_x p_x(x) \, S\left(\mathcal{E}(\rho_A^x)\right)$$

Seen in Ex. sheet 5.

$$\Rightarrow I(N,B) = S(\rho_B) + S(\rho_N) - S(\rho_{NB})$$

$$= S\left(\mathcal{E}\left(\sum_x p_x(x)\, \rho_A^x\right)\right) + \cancel{H(X)} - \left(\cancel{H(X)} + \sum_x p_x(x)\, S\left(\mathcal{E}(\rho_A^x)\right)\right)$$

b) Determine the channel capacity of the binary symmetric channel defined by

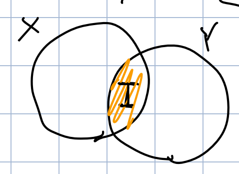
$$\Pr(0|0) = \Pr(1|1) = 1 - p$$
$$\Pr(1|0) = \Pr(0|1) = p.$$

Hint: It may be useful to expand $H(Y|X)$ *as* $\sum_x p(x) H(Y|X = x)$.

$$C\left(P(Y|X)\right) := \sup_{P_X(x)} I(X:Y)$$

where $I(X:Y) = H(Y) - H(Y|X)$

QUICK REMAINDER:



where Y is defined by:

$$P_Y(y) = \sum_x P(x,y) \underset{\uparrow}{=} \sum_x P(y|x)\, P_X(x)$$
$$P(y|x)\, P_X(x) = P(x,y)$$

- $$H(Y) = -\sum_y P_Y(y) \log\left(P_Y(y)\right)$$

- $$H(Y|X) = \mathbb{E}_X\left[H(Y|X=x)\right] = \sum_x P_X(x)\, H(Y|X=x)$$

$$= -\sum_{x,y} P_X(x)\, P(y|x) \log\left(P(y|x)\right) =$$

$$= -P_X(0)\left(P(0|0) \log\left(P(0|0)\right) + P(1|0) \log\left(P(1|0)\right)\right)$$

$$- P_X(1)\left(P(0|1) \log\left(P(0|1)\right) + P(1|1) \log\left(P(1|1)\right)\right)$$

$$= - P_X(0)\Big((1-p)\log(1-p) + p\log(p)\Big)$$

$$- P_X(1)\Big(p\log(p) + (1-p)\log(1-p)\Big)$$

$$= -\Big(p\log(p) + (1-p)\log(1-p)\Big) =: H_2(p)$$

$$I(X:Y) = H(Y) - H(Y|X) = H(Y) - H_2(p)$$

$$\leq \log_2(2) - H_2(p) = 1 - H_2(p)$$

$$C(P(Y|X)) := \sup_{P_X(x)} I(X:Y) = \sup_{P_X(x)}\big(H(Y)\big) - H_2(p) \overset{H(Y)\leq \log(d)}{\leq} \underbrace{\log(2)}_{=1} - H_2(p)$$

- This is saturated if $P_Y(Y) = \frac{1}{2}$.

If $P_X(X) = \frac{1}{2} \Rightarrow P_Y(Y) = \sum_x P(Y|X)\, P_X(x) = \frac{1}{2}\Big(\underbrace{\sum_x P(Y|x)}_{=1}\Big) = \frac{1}{2}$

$$\Rightarrow C(P(Y|X)) = 1 - H_2(p)$$

We now want to determine the channel capacity of the binary erasure channel as defined by

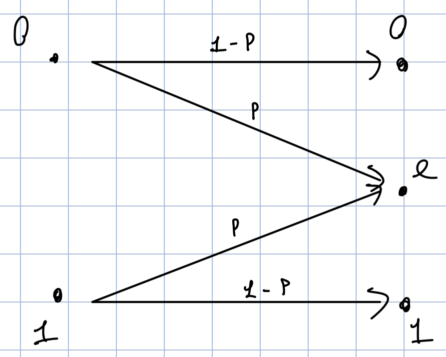
$\Pr(1|0) = \Pr(0|1) = 0$

$\Pr(0|0) = \Pr(1|1) = 1 - p$

$\Pr(e|0) = \Pr(e|1) = p.$

c) First, use the expansion $H(Y) = H(Y,Z) = H(E) + H(Y|Z)$ to show that $H(Y) = H(p) + (1-p)H(\pi)$. Here, we let Z be the random variable distinguishing between the event $E = \{Y = e\}$ and $\neg E = \{Y \neq e\}$. We have that $\Pr(Z = E) = p$. Furthermore we call the propability defining the distribution of the input variable $\pi = \Pr(X = 1)$.

Hint: Use that $\Pr(Y = y|Y \neq e) = \Pr(X = y)$.



- $$H(Y) = -\sum_{y} P_Y(y)\log\big(P_Y(y)\big) \quad \text{with} \quad P_Y(y) = \sum_{x} P(y|x)\,P(x).$$

- $$P_Y(0) = P(0|0)P_X(0) + \underbrace{P(0|1)}_{0} P_X(1) = (1-p)\,P_X(0)$$

$$P_Y(1) = \underbrace{P(1|0)}_{0} P_X(0) + P(1|1)P_X(1) = (1-p)\,P_X(1)$$

$$P_Y(e) = P(e|0)P_X(0) + P(e|1)P_X(1) = p$$

- $$H(Y) = -\sum_{y} P_Y(y)\log\big(P_Y(y)\big) =$$

$$= -(1-p)P_X(0)\log\big((1-p)P_X(0)\big)$$
$$-(1-p)P_X(1)\log\big((1-p)P_X(1)\big)$$
$$-p\log p$$
$$= -(1-p)P_X(0)\log\big((1-p)\big) - (1-p)P_X(0)\log\big(P_X(0)\big)$$
$$-(1-p)P_X(1)\log\big((1-p)\big) - (1-p)P_X(1)\log\big(P_X(1)\big)$$
$$-p\log p$$
$$= -(1-p)\log(1-p) + (1-p)H(P_X) - p\log p$$
$$\underset{\pi = P_X}{=} H(p) + (1-p)H(\pi)$$

d) Use this result and proceed analogously to the binary symmetric channel to determine the channel capacity of the erasure channel.

$$C\left(P(Y|X)\right) = \sup_{P_X(x)} I(X:Y) =$$

$$= \sup_{P_X(x)} \left(H(Y) - H(Y|X) \right)$$

$$H(Y|X) = \sum_x P_X(x)\, H(Y|X=x)$$

$$= P_X(0)\, H(Y|X=0) + P_X(1)\, H(Y|X=1) =$$

$$H(Y|X=1) = H(Y|X=0)$$

$\{P(Y|0)$ with $y=0,1,e\} = \{P(Y|1)$ with $y=0,1,e\}$

$\{0, p, 1-p\} = \{0, p, 1-p\}$

$$\underset{P_X(0)+P_X(1)=1}{=} H(Y|X=0) =$$

$$= -\sum_Y P(Y|0) \log(P(Y|0)) =$$

$$= -\underbrace{(1-p)}_{Y=0} \log(1-p) - \underbrace{p}_{Y=1} \log p + \underbrace{0}_{e} =$$

$$= H(p)$$

$$C\left(P(Y|X)\right) = \sup_{P_X(x)} I(X:Y) = \sup_{P_X(x)} \left(H(Y) - H(Y|X) \right)$$

$$= \sup_{P_X(x)} \left(H(Y) - H(p) \right)$$

$$= \sup_{P_X(x)} \left(\cancel{H(p)} + (1-p)\, H(P_x) - \cancel{H(p)} \right) =$$

$$= (1-p) \sup_{P_X(x)} H(P_x) \underset{\uparrow}{=} (1-p) \cdot 1$$

$$\sup_{P_X(x)} H(P_x) = \log(2) = 1$$

- The sup is achieved for $P_x(x) = \frac{1}{2}$.

2. **Majorisation and transforming quantum states by local unitaries.** (8 Points: 2+2+2+2)

In this problem we will look at the task of transforming a given copy of a pure bipartite quantum state $|\psi\rangle$ to another quantum state $|\phi\rangle$ using LOCC. The question is: Under which conditions is the transition $|\psi\rangle \xrightarrow{\text{LOCC}} |\phi\rangle$ possible?

The key to the answer of this question is the concept of majorisation. We say that a vector $x \in R^n$ majorises $y \in R^n$ (denoted by $x \succ y$) if for all $k = 1, \ldots, n$, $\sum_{j=1}^{k} x_j^{\downarrow} \geq \sum_{j=1}^{k} y_j^{\downarrow}$ and $\sum_{j=1}^{n} x_j^{\downarrow} = \sum_{j=1}^{n} y_j^{\downarrow}$. Here, $x^{\downarrow}$ denotes the sorted version of x, i.e., a permutation of the elements of x such that $x_1^{\downarrow} \geq x_2^{\downarrow} \geq \ldots \geq x_n^{\downarrow}$. From now on, let x and y be non-negative vectors.

a) Show that $x = (\frac{2}{3}, \frac{1}{3}, 0)^T$ majorises $y = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})^T$.

- $x^{\downarrow} := \text{sort}(x, \text{''}>\text{''})$
- $X \succ Y \overset{\text{DEF.}}{\iff} \sum_{j=1}^{k} x_j^{\downarrow} \geq \sum_{j=1}^{k} y_j^{\downarrow}$ for $k=1,\ldots,n$ and $\sum_{j=1}^{n} x_j = \sum_{j=1}^{n} y_j$
- $x^{\downarrow} = \left(\frac{2}{3}, \frac{1}{3}, 0\right)^T$
- $y^{\downarrow} = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)^T$

$$\sum_{j=1}^{k} x_j^{\downarrow} \geq \sum_{j=1}^{k} y_j^{\downarrow} \quad \text{for } k = 1,2,3 \quad \text{and} \quad \underset{\overset{\shortparallel}{1}}{\sum_{j=1}^{n} x_j} = \underset{\overset{\shortparallel}{1}}{\sum_{j=1}^{n} y_j}$$

One can show that $x \prec y$ if and only if $x = \sum_j p_j \Pi_j y$ for a probability distribution p and permutation matrices Π_j (Please accept this equation.). By Birkhoff's theorem, which lies at the heart of majorisation theory, that statement is equivalent to saying that $x \prec y$ if and only if $x = Dy$ for some doubly stochastic matrix D[1].

For two Hermitian operators $X, Y \in L(C^d)$ we say that $X \prec Y$ if $\lambda(X) \prec \lambda(Y)$, where $\lambda(A)$ is the spectrum of a matrix A.

b) Show that $X \prec Y$ if and only if there exists a probability distribution p and unitary matrices U_j such that

$$X = \sum_j p_j U_j Y U_j^\dagger.$$

Hint: For "only if" direction, do eigenvalue decomposition as $X = U\Lambda_X U^\dagger$ *and* $Y = V\Lambda_Y V^\dagger$, *and use the fact "*$\lambda(X) \prec \lambda(Y)$ *if and only if* $\lambda(X) = \sum_j p_j \Pi_j \lambda(Y)$*". For "if" direction, use again eigenvalue decomposition and the fact "*$\lambda(X) \prec \lambda(Y)$ *if and only if* $\lambda(X) = D\lambda(Y)$ *for some doubly stochastic matrix D".*

[1] A matrix D is called doubly stochastic if $\forall i, j\, D_{ij} \geq 0$ and $\forall i \sum_j D_{ij} = \sum_j D_{ji} = 1$, i.e., all rows and columns are probability distributions.

"$\Leftarrow$"

$$X = \sum_j p_j U_j Y U_j^+ \Rightarrow \tilde{U}\Lambda_x\tilde{U}^+ = \sum_j p_j U_j V \Lambda_Y V^+ U_j^+$$

$$X = \tilde{U}\Lambda_x\tilde{U}^+$$
$$Y = V\Lambda_Y V^+$$

$$\Rightarrow \Lambda_x = \sum_j p_j \underbrace{(\tilde{U}^+ U_j V)}_{=V^{(j)}} \Lambda_Y \underbrace{(V^+ U_j^+ \tilde{U})}_{=V^{(j)+}}$$

$$(\lambda(X))_i = (\Lambda_x)_{ii} = \sum_j p_j (V^{(j)} \Lambda_Y V^{(j)+})_{ii} =$$

$$= \sum_j p_j \sum_l V^{(j)}_{il} \Lambda_{Y_{ll}} V^{(j)+}_{li} =$$

$$= \sum_l \underbrace{\left(\sum_j p_j V^{(j)}_{il} V^{(j)+}_{li}\right)}_{=D_{i,l}} \underbrace{(\Lambda_Y)_{ll}}_{=(\lambda(Y))_l} = (D\,\lambda(Y))_i$$

- $D_{i,l} = \sum_s p_s V^{(s)}_{il} V^{(s)*}_{il} = \sum_s p_s |V^{(s)}_{il}|^2 \geq 0$

- $\sum_l D_{i,l} = \sum_s p_s \underbrace{\sum_l |V^{(s)}_{il}|^2}_{\overset{V \text{ unitary}}{=} 1} = 1$

- $\sum_i D_{i,l} = \sum_s p_s \underbrace{\sum_i |V^{(s)}_{il}|^2}_{\overset{V \text{ unitary}}{=} 1} = 1$

$\Rightarrow D$ stoc. $\Rightarrow X \lesssim Y$

"$\Rightarrow$" $X \prec Y \Leftrightarrow \lambda(X) = \sum_s p_s \Pi_s \lambda(Y)$

$$\Rightarrow \quad \big(\lambda(X)\big)_i = \sum_s p_s \big(\Pi_s \lambda(Y)\big)_i$$

$$\Lambda_X = \sum_i \big(\lambda(X)\big)_i |i\rangle\langle i| = \sum_i \sum_s p_s \big(\Pi_s \lambda(Y)\big)_i |i\rangle\langle i|$$

$$= \sum_i \sum_s p_s \sum_l (\Pi_s)_{i,l} \big(\lambda(Y)\big)_l |i\rangle\langle i|$$

$$= \sum_s p_s \sum_l \big(\lambda(Y)\big)_l \underbrace{\sum_i (\Pi_s)_{i,l} |i\rangle\langle i|}_{\overset{\Pi_s \text{ is a permutation matrix}}{=} \Pi_s |l\rangle\langle l| \Pi_s^+}$$

$$= \sum_s p_s \Pi_s \Big(\sum_l \lambda(Y)_l |l\rangle\langle l|\Big) \Pi_s^+ = \sum_s p_s \Pi_s \Lambda_Y \Pi_s^+$$

$$\Rightarrow X = U \Lambda_X U^\dagger = \sum_j p_j U \Pi_j \Lambda_Y \Pi_j^\dagger U^\dagger \underset{\Lambda_Y = V^\dagger Y V}{=}$$

$$= \sum_j p_j (U \Pi_j V^\dagger) Y (V \Pi_j^\dagger U^\dagger) \implies \square$$

We are now ready to prove the (surprising!) theorem: $|\psi\rangle \xrightarrow{\text{LOCC}} |\phi\rangle$ if and only if $\mathrm{Tr}_B[|\psi\rangle\langle\psi|] \prec \mathrm{Tr}_B[|\phi\rangle\langle\phi|]$. (We encourage you to have a look into `https://arxiv.org/pdf/quant-ph/9811053.pdf`, which is the original paper of the thorem.)

c) Show the "only if" direction using the previous result. You can suppose that LOCC is realised by a measurement on Alice's side and a corresponding unitary on Bob's side. In other words, from Alice's point of view it must be the case that [2]

$$M_j \mathrm{Tr}_B[|\psi\rangle\langle\psi|] M_j^\dagger = p_j \mathrm{Tr}_B[|\phi\rangle\langle\phi|].$$

Hint: Use the polar decomposition of $M_j \sqrt{\mathrm{Tr}_B[|\psi\rangle\langle\psi|]}$.

POLAR DECOMPOSITION (NIELSEN & CHUANG PAG. 79)

HP: • A linear operator

TH: • $A = UJ = KU^\dagger$ with U unitary and $J \geq 0$, $K \geq 0$.

J and K are unique: $J = \sqrt{A^\dagger A}$ and $K = \sqrt{AA^\dagger}$

PROOF:

$A = U\sqrt{A^\dagger A}$

SUBPROOF:

• $J := \sqrt{A^\dagger A} \underset{\sqrt{A^\dagger A} \geq 0 \Rightarrow \text{Hermitian} \Rightarrow \text{diagonalizable.}}{=} \sum_{i=1}^N \underset{\geq 0}{\lambda_i} |i\rangle\langle i|$

• $\forall \lambda_i \neq 0$, $|e_i\rangle := \dfrac{A|i\rangle}{\lambda_i}$ with $\langle e_k|e_i\rangle = \dfrac{\langle k|A^\dagger A|i\rangle}{\lambda_k \lambda_i} = \dfrac{\langle k|J^2|i\rangle}{\lambda_k\lambda_i} = \dfrac{\lambda_i^2}{\lambda_k\lambda_i}\underset{\delta_{ik}}{\langle k|i\rangle} = \delta_{i,k}$

• Gram-Schmidt procedure extends $\{|e_i\rangle\}_{i:\lambda_i \neq 0}$ to an ort. basis. $\{|e_i\rangle\}_{i=1,\dots,N}$

• $U := \sum_{i=1}^N |e_i\rangle\langle i|$

- $U S |i\rangle = A|i\rangle \quad \forall i = 1, \dots, N \Rightarrow US = A$

  $\cdot\, US|i\rangle = \lambda_i U|i\rangle = \lambda_i |e_i\rangle = A|i\rangle$ $\Rightarrow \square$ SUB.

- UNIQUENESS:

  SUBPROOF:

  $A = US \Rightarrow A^+A = S^+U^+US = S^+S \underset{\substack{S^+=S \\ \uparrow \\ S\geq 0}}{=} S^2 \Rightarrow S = \sqrt{A^+A}$ $\Rightarrow \square$ SUB.

- $A = \sqrt{AA^+}\, U$

  SUBPROOF:

  - $A = US = USU^+U = KU$ with $K := USU^+ \geq 0$
  - $A = KU \Rightarrow AA^+ = KUU^+K^+ = KK^+ \underset{K^+=K}{=} K^2 \Rightarrow K = \sqrt{AA^+} \geq 0$ $\Rightarrow \square$

- $\boxed{SVD}$

  HP: A linear operator.

  TH: $A = UDV^+$ with U, V unitary, D diagonal with $\sum_{i=1}^{N} D_{ii}^2 = tr(A^+A) = \sum_{i,j} |A_{ij}|^2$

  PROOF:

  - $\underset{\substack{\uparrow \\ \text{POLAR}}}{A =} \tilde{U}S \underset{\substack{\uparrow \\ (S\geq 0 \Rightarrow \text{diagonaliz.} \\ \Rightarrow S = VDV^+)}}{=} \tilde{U}VDV^+ \underset{\substack{\uparrow \\ U = \tilde{U}V}}{=} UDV^+$ with D diagonal and $D \geq 0$.
  - $A^+A = VD^2V^+ \Rightarrow \underset{\substack{\| \\ \sum_{i,j} A_{ij}^2}}{tr(A^+A)} = tr(D^2) = \sum_{i=1}^{N} D_{ii}^2$

We are now ready to prove the (surprising!) theorem: $|\psi\rangle \xrightarrow{\text{LOCC}} |\phi\rangle$ if and only if $\mathrm{Tr}_B[|\psi\rangle\langle\psi|] \prec \mathrm{Tr}_B[|\phi\rangle\langle\phi|]$. (We encourage you to have a look into `https://arxiv.org/pdf/quant-ph/9811053.pdf`, which is the original paper of the thorem.)

c) Show the "only if" direction using the previous result. You can suppose that LOCC is realised by a measurement on Alice's side and a corresponding unitary on Bob's side. In other words, from Alice's point of view it must be the case that [2]

$$M_j \mathrm{Tr}_B[|\psi\rangle\langle\psi|] M_j^\dagger = p_j \mathrm{Tr}_B[|\phi\rangle\langle\phi|].$$

Hint: Use the polar decomposition of $M_j\sqrt{\mathrm{Tr}_B[|\psi\rangle\langle\psi|]}$.

$$|\psi\rangle \xrightarrow{LOCC} |\phi\rangle \implies \mathrm{Tr}_B[|\psi\rangle\langle\psi|] \prec \mathrm{Tr}_B[|\phi\rangle\langle\phi|]$$

PROOF:

- $\mathrm{Tr}_B[|\psi\rangle\langle\psi|] \prec \mathrm{Tr}_B[|\phi\rangle\langle\phi|] \iff \exists\, U_j$ and p_j such that $\mathrm{Tr}_B[|\psi\rangle\langle\psi|] = \sum_j p_j U_j^\dagger \mathrm{Tr}_B[|\phi\rangle\langle\phi|] U_j$

- $\mathrm{Tr}_B[|\psi\rangle\langle\psi|] = \sqrt{\mathrm{Tr}_B[|\psi\rangle\langle\psi|]}\sqrt{\mathrm{Tr}_B[|\psi\rangle\langle\psi|]} =$

$$\underset{\uparrow}{=} \sum_i \sqrt{\mathrm{Tr}_B[|\psi\rangle\langle\psi|]}\, M_i^\dagger M_i \sqrt{\mathrm{Tr}_B[|\psi\rangle\langle\psi|]} =$$

$\sum_i M_i^\dagger M_i = \mathbb{1}$

$\{M_i^\dagger M_i\}_{i=1}^{d}$ POVM: $\sum_{i=1}^{N} M_i^\dagger M_i = \mathbb{1}$ and $\mathrm{Prob}(i) = \mathrm{tr}(\rho M_i^\dagger M_i)$

POST MEAS. STATE: $\rho_{POST} = \sum_i M_i \rho M_i^\dagger$

$$= \sum_i U_i^\dagger \left(M_i \,\mathrm{Tr}_B[|\psi\rangle\langle\psi|] M_i^\dagger\right) U_i = (\star)$$

$$M_i\sqrt{\mathrm{Tr}_B[|\psi\rangle\langle\psi|]} \underset{\text{POLAR}}{=} \sqrt{M_i \,\mathrm{Tr}_B[|\psi\rangle\langle\psi|] M_i^\dagger}\; U_i$$

$|\psi\rangle \xrightarrow{LOCC} |\phi\rangle$ $\Rightarrow$ We proved in previous ex. that a Bob's measurement can be simulated by Alice's measurement + classical communication and local unitaries.

ALICE BOB

$$\Rightarrow \quad \left(\mathbb{1} \otimes \mathcal{U}_B^{(i)}\right)\left[\left(\frac{M_i \otimes \mathbb{1}\ |\psi\rangle\langle\psi|\ M_i^\dagger \otimes \mathbb{1}}{\mathrm{Tr}[\ M_i^\dagger M_i \otimes \mathbb{1}\ |\psi\rangle\langle\psi|]}\right)\right] = |\phi\rangle\langle\phi|$$

Quantum Channel by Bob.

$$\Rightarrow \quad \mathrm{Tr}_B\left[\left(\mathbb{1} \otimes \mathcal{U}_B^{(i)}\right)\left[\left(\frac{M_i \otimes \mathbb{1}\ |\psi\rangle\langle\psi|\ M_i^\dagger \otimes \mathbb{1}}{\mathrm{Tr}[\ M_i^\dagger M_i \otimes \mathbb{1}\ |\psi\rangle\langle\psi|]}\right)\right]\right] = \mathrm{Tr}_B\left[|\phi\rangle\langle\phi|\right]$$

$\mathcal{U}_B$ trace preserving $\rightarrow$ $\|$

$$\mathrm{Tr}_B\left[\left(\frac{M_i \otimes \mathbb{1}\ |\psi\rangle\langle\psi|\ M_i^\dagger \otimes \mathbb{1}}{\mathrm{Tr}[\ M_i^\dagger M_i \otimes \mathbb{1}\ |\psi\rangle\langle\psi|]}\right)\right]$$

$$\| $$

$$\frac{M_i\ \mathrm{Tr}_B[|\psi\rangle\langle\psi|]\ M_i^\dagger}{\left(\mathrm{Tr}[\ M_i^\dagger M_i\ \mathrm{Tr}_B[|\psi\rangle\langle\psi|]]\right) =: p_i}$$

$$\text{(★)} \Rightarrow \mathrm{Tr}_B[|\psi\rangle\langle\psi|] = \sum_i U_i^\dagger \left(M_i\ \mathrm{Tr}_B[|\psi\rangle\langle\psi|]\ M_i^\dagger\right) U_i$$

$$= \sum_i p_i\ U_i^\dagger\ \mathrm{Tr}_B[|\phi\rangle\langle\phi|]\ U_i$$

$$\Rightarrow \quad \mathrm{Tr}_B[|\psi\rangle\langle\psi|] \prec \mathrm{Tr}_B[|\phi\rangle\langle\phi|]$$

Freie Universität Berlin
Tutorials on Quantum Information Theory
Winter term 2022/23

Problem Sheet 7
Capacities and Majorization

J. Eisert, A. Townsend-Teague, A. Mele, A. Burchards, J. Denzler

1. **Classical capacities of quantum channels**
Although this exercise might look very long, it isn't. In the next paragraphs we just want to give you an overview on the formalism introduced in the lecture and needed for this exercise in a compressed fashion. No need to be intimidated ;)
In the lecture, we saw two alternative characterisations of the classical channel capacity of a quantum channel $\mathcal{E}$, which is given by the its Holevo-information $\chi(\mathcal{E})$. The task here is to establish the equivalence of these expressions.

To this end, recall the definition of the quantum mutual information of a bi-partite quantum system in a state ρ_{AB}

$$I(A:B)_{\rho_{AB}} \coloneqq S(\rho_A) + S(\rho_B) - S(\rho_{AB}). \tag{1}$$

The Holevo information of channel can be defined using the following scheme: Alices encodes the information of a classical random variable X taking values in $\mathcal{X}$ with probability distribution p_X into a quantum state using a set of states $\{\rho_x\}_{x\in\mathcal{X}}$. To keep track of the classical random variable but formulating everything quantum mechanically, we think of Alice encoding the result in another faithfully register N using orthogonal basis $\{|x\rangle\}_{x\in\mathcal{X}}$. From this notebook register N the classical information of X can be completely recovered. Altogether, Alice prepares the bi-partite state

$$\rho_{NA} = \sum_x p_X(x)\, |x\rangle\langle x|_N \otimes \rho_A^x. \tag{2}$$

Then, the state in system A is sent to Bob using the channel $\mathcal{E}$. Thus, we end up with a final state shared between Alice's notebook and Bob

$$\rho_{NB} = \sum_x p_X(x)\, |x\rangle\langle x|_N \otimes \mathcal{E}(\rho_A^x)_B. \tag{3}$$

We can now ask for the mutual information between the variable X encoded in N and Bob's output of the channel. Analogously to the classical result, maximizing the mutual information over all possible input variables X and encodings yields the capacity of the quantum channel to transmit classical informations, i.e.

$$\chi(\mathcal{E}) = \max_{(X,p_X,\{\rho^x\})} I(N,B)_{\rho_{NA}}. \tag{4}$$

a) Show that

$$\chi(\mathcal{E}) = \max_{(X,p_X,\{\rho^x\})} \left\{ S(\mathcal{E}(\sum_x p_X(x)\rho^x)) - \sum_x p_X(x) S(\mathcal{E}(\rho^x)) \right\}. \tag{5}$$

Remember that Shannon's noisy channel coding theorem states that the capacity of a noisy channel T is given by the maximum over all inputs of the mutual information:

$$C(T) = \max_{X,p_X} I(X:Y),$$

where Y is the random variable describing the output of the channel T with input X.

b) Determine the channel capacity of the binary symmetric channel defined by

$$\Pr(0|0) = \Pr(1|1) = 1 - p$$
$$\Pr(1|0) = \Pr(0|1) = p.$$

Hint: It may be useful to expand $H(Y|X)$ as $\sum_x p(x)H(Y|X = x)$.

We now want to determine the channel capacity of the binary erasure channel as defined by

$$\Pr(0|0) = \Pr(1|1) = 1 - p$$
$$\Pr(e|0) = \Pr(e|1) = p.$$

c) First, use the expansion $H(Y) = H(Y, Z) = H(E)+H(Y|Z)$ to show that $H(Y) = H(p)+(1-p)H(\pi)$. Here, we let Z be the random variable distinguishing between the event $E = \{Y = e\}$ and $\neg E = \{Y \neq e\}$. We have that $\Pr(Z = E) = p$. Furthermore we call the propability defining the distribution of the input variable $\pi = \Pr(X = 1)$.

Hint: Use that $\Pr(Y = y|Y \neq e) = \Pr(X = y)$.

d) Use this result and proceed analogously to the binary symmetric channel to determine the channel capacity of the erasure channel.

2. **Majorisation and transforming quantum states by local unitaries.** (8 Points: 2+2+2+2)

In this problem we will look at the task of transforming a given copy of a pure bipartite quantum state $|\psi\rangle$ to another quantum state $|\phi\rangle$ using LOCC. The question is: Under which conditions is the transition $|\psi\rangle \xrightarrow{\text{LOCC}} |\phi\rangle$ possible?

The key to the answer of this question is the concept of majorisation. We say that a vector $x \in R^n$ majorises $y \in R^n$ (denoted by $x \succ y$) if for all $k = 1, \ldots, n$, $\sum_{j=1}^{k} x_j^{\downarrow} \geq \sum_{j=1}^{k} y_j^{\downarrow}$ and $\sum_{j=1}^{n} x_j^{\downarrow} = \sum_{j=1}^{n} y_j^{\downarrow}$. Here, $x^{\downarrow}$ denotes the sorted version of x, i.e., a permutation of the elements of x such that $x_1^{\downarrow} \geq x_2^{\downarrow} \geq \ldots \geq x_n^{\downarrow}$. From now on, let x and y be non-negative vectors.

a) Show that $x = (\frac{2}{3}, \frac{1}{3}, 0)^T$ majorises $y = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})^T$.

One can show that $x \prec y$ if and only if $x = \sum_j p_j \Pi_j y$ for a probability distribution p and permutation matrices Π_j (Please accept this equation.). By Birkhoff's theorem, which lies at the heart of majorisation theory, that statement is equivalent to saying that $x \prec y$ if and only if $x = Dy$ for some doubly stochastic matrix D[1].

For two Hermitian operators $X, Y \in L(C^d)$ we say that $X \prec Y$ if $\lambda(X) \prec \lambda(Y)$, where $\lambda(A)$ is the spectrum of a matrix A.

b) Show that $X \prec Y$ if and only if there exists a probability distribution p and unitary matrices U_j such that

$$X = \sum_j p_j U_j Y U_j^{\dagger}.$$

Hint: For "only if" direction, do eigenvalue decomposition as $X = U\Lambda_X U^{\dagger}$ and $Y = V\Lambda_Y V^{\dagger}$, and use the fact "$\lambda(X) \prec \lambda(Y)$ if and only if $\lambda(X) = \sum_j p_j \Pi_j \lambda(Y)$". For "if" direction, use again eigenvalue decomposition and the fact "$\lambda(X) \prec \lambda(Y)$ if and only if $\lambda(X) = D\lambda(Y)$ for some doubly stochastic matrix D".

[1] A matrix D is called doubly stochastic if $\forall i, j D_{ij} \geq 0$ and $\forall i \sum_j D_{ij} = \sum_j D_{ji} = 1$, i.e., all rows and columns are probability distributions.

We are now ready to prove the (surprising!) theorem: $|\psi\rangle \xrightarrow{\text{LOCC}} |\phi\rangle$ if and only if $\mathrm{Tr}_B[|\psi\rangle\langle\psi|] \prec \mathrm{Tr}_B[|\phi\rangle\langle\phi|]$. (We encourage you to have a look into `https://arxiv.org/pdf/quant-ph/9811053.pdf`, which is the original paper of the thorem.)

c) Show the "only if" direction using the previous result. You can suppose that LOCC is realised by a measurement on Alice's side and a corresponding unitary on Bob's side. In other words, from Alice's point of view it must be the case that [2]

$$M_j \mathrm{Tr}_B[|\psi\rangle\langle\psi|] M_j^\dagger = p_j \mathrm{Tr}_B[|\phi\rangle\langle\phi|].$$

Hint: Use the polar decomposition of $M_j \sqrt{\mathrm{Tr}_B[|\psi\rangle\langle\psi|]}$.

d) Now show the "if" direction by proceeding analogously.

[2]This is because the transition from $|\psi\rangle$ to $|\phi\rangle$ comes about as a post-measurement state with probability p_j.