Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Quantum Information Theory
Solutions 6

HS 12
Prof. R. Renner

Exercise 6.1 Data hiding

Consider a 2d-qubit Hilbert space, $\mathcal{H}_A \otimes \mathcal{H}_B$, and the computational basis of both spaces. Consider the projectors onto the symmetric and antisymmetric subspaces of $\mathcal{H}_A \otimes \mathcal{H}_B$,

$$\Pi^S = \frac{1}{2}\sum_{i<j}\Big(|i\rangle_A|j\rangle_B + |j\rangle_A|i\rangle_B\Big)\Big(\langle i|_A\langle j|_B + \langle j|_A\langle i|_B\Big) + \sum_i |i\rangle_A|i\rangle_B\langle i|_A\langle i|_B,$$
$$\Pi^A = \frac{1}{2}\sum_{i<j}\Big(|i\rangle_A|j\rangle_B - |j\rangle_A|i\rangle_B\Big)\Big(\langle i|_A\langle j|_B - \langle j|_A\langle i|_B\Big).$$

You will encode only one bit of information, b, giving Alice and Bond each their $d-$qubit part of ρ^b_{AB}, with

$$\rho^{b=0} = \frac{2}{d(d+1)}\Pi^S, \qquad \rho^{b=1} = \frac{2}{d(d-1)}\Pi^A.$$

a) *Show that $\rho^{b=0}$ and $\rho^{b=1}$ are valid density operators and explain how you would proceed to recover b if you had access to Alice and Bond's systems (together).*

Both Π^S and Π^A are projectors (they have the form $\sum_i |\phi_i\rangle\langle\phi_i|$, for orthonormal $\{|\phi_i\rangle\}_i$), so $\rho^{b=0}$ and $\rho^{b=1}$ are Hermitian and positive semi-definite. As for normalization, we have

$$\rho^{b=0} = \frac{2}{d(d+1)}\left(\overbrace{\sum_i}^{d\text{ terms}} |ii\rangle\langle ii| + \frac{1}{2}\overbrace{\sum_j\sum_{i<j}}^{\frac{d(d-1)}{2}\text{ terms}} |ij\rangle\langle ij| + |ji\rangle\langle ji| + |ij\rangle\langle ji| + |ji\rangle\langle ij|\right)$$
$$\mathrm{Tr}(\rho^{b=0}) = \frac{2}{d(d+1)}\left(d + \frac{1}{2}\left[\frac{d(d-1)}{2} + \frac{d(d-1)}{2}\right]\right) = 1;$$

and

$$\rho^{b=1} = \frac{2}{d(d-1)}\left(\frac{1}{2}\overbrace{\sum_j\sum_{i<j}}^{\frac{d(d-1)}{2}\text{ terms}} |ij\rangle\langle ij| + |ji\rangle\langle ji| - |ij\rangle\langle ji| - |ji\rangle\langle ij|\right)$$
$$\mathrm{Tr}(\rho^{b=1}) = \frac{2}{d(d-1)}\left(\frac{1}{2}\left[\frac{d(d-1)}{2} + \frac{d(d-1)}{2}\right]\right) = 1.$$

If we had access to both systems, we could perform the global measurement described by the POVM $\{\Pi^S, \Pi^A, \mathbb{1} - \Pi^S - \Pi^A\}$. The probabilities of the three possible outcomes are $(1,0,0)$ if the state is $\rho^{b=10}$ and $(0,1,0)$ if the state is $\rho^{b=1}$, so we could recover the value of b with certainty.

b) *Consider the flip operator in basis $\{|i\rangle_A|j\rangle_B\}_{ij}$,*

$$F = \Pi^S - \Pi^A = \sum_{i,j}|i\rangle_A|j\rangle_B\langle j|_A\langle i|_B.$$

Show that, for all operators $M_A \in End(\mathcal{H}_A), N_B \in End(\mathcal{H}_B)$,

$$Tr[F(M_A \otimes N_B)] = Tr(M_A N_B).$$

In particular, for all pure states $|x\rangle_A, |y\rangle_B$, $Tr[F|xy\rangle\langle xy|] = |\langle x|y\rangle|^2$.

We expand the operators in the basis of the flip operator,

$$M = \sum_{i,j} x_{ij} |i\rangle\langle j|, \qquad N = \sum_{k,\ell} y_{k\ell} |k\rangle\langle \ell|.$$

Applying the flip operator, we have

$$\begin{aligned} F(M_A \otimes N_B) &= \sum_{i,j'} |i'j'\rangle\langle j'i'| \left(\sum_{i,j,k,\ell} x_{ij}\ y_{k\ell}\ |ik\rangle\langle j\ell| \right) \\ &= \sum_{i,j,k,\ell} x_{ij}\ y_{k\ell}\ |ki\rangle\langle j\ell|. \end{aligned}$$

Now we take the trace,

$$\begin{aligned} \mathrm{Tr}[F(M_A \otimes N_B)] &= \sum_{i',j'} \langle i', j'| \left(\sum_{i,j,k,\ell} x_{ij}\ y_{k\ell}\ |ki\rangle\langle j\ell| \right) |i', j'\rangle \\ &= \sum_{i,j} x_{ij} y_{ji}. \end{aligned}$$

On the other hand,

$$\begin{aligned} M_A N_B &= \left(\sum_{i,j} x_{ij} |i\rangle\langle j| \right) \left(\sum_{k,\ell} y_{k\ell} |k\rangle\langle \ell| \right) = \sum_{i,j,\ell} x_{ij} y_{j\ell} |i\rangle\langle \ell|, \\ \mathrm{Tr}(M_A N_B) &= \sum_{i',j'} \langle i'| \left(\sum_{i,j,\ell} x_{ij} y_{j\ell} |i\rangle\langle \ell| \right) |i'\rangle = \sum_{i,j} x_{ij} y_{ji}, \end{aligned}$$

which proves our claim. In the particular case of pure states, $M = |x\rangle\langle x|, N = |y\rangle\langle y|$, we can take the trace using an on. basis $\{|x_i\rangle\}_i$, such that $|x_0\rangle = |x\rangle$,

$$\mathrm{Tr}(MN) = \sum_i \langle x_i|x\rangle\langle x|y\rangle\langle y|x_i\rangle = \langle x|y\rangle\langle y|x\rangle = |\langle x|y\rangle|^2 .$$

c) *Suppose that Alice and Bond perform local projective measurements in arbitrary bases $\{|x\rangle_A\}$ and $\{|y\rangle_B\}$ respectively. We call the joint probability distribution of the outcomes P_{XY} when they measure state $\rho^{b=0}$ and Q_{XY} when they measure $\rho^{b=1}$. We want them to be unable to determine which state they measured, i.e., to distinguish the two distributions, so we want to show that $\delta(P_{XY}, Q_{XY})$ is small. Remember that*

$$P_{XY}(x,y) = Tr(|xy\rangle\langle xy|\rho^{b=0}), \qquad Q_{XY}(x,y) = Tr(|xy\rangle\langle xy|\rho^{b=1}).$$

Use the results from b) to show that $\delta(P_{XY}, Q_{XY}) \leq \frac{2}{d+1}$.

Hint: start from the trace distance as

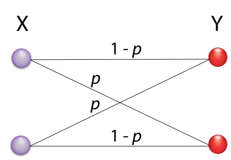
$$\delta(P_{XY}, Q_{XY}) = \sum_{x,y\in\mathcal{S}} P_{XY}(x,y) - Q_{XY}(x,y),$$

with $\mathcal{S} = \{(x,y) : P_{XY}(x,y) > Q_{XY}(x,y)\}$.

$$
\begin{aligned}
\delta(P_{XY}, Q_{XY}) &= \sum_{x,y\in\mathcal{S}} P_{XY}(x,y) - Q_{XY}(x,y) \\
&= \sum_{x,y\in\mathcal{S}} \mathrm{Tr}(|xy\rangle\langle xy|\rho^{b=0}) - \mathrm{Tr}(|xy\rangle\langle xy|\rho^{b=1}) \\
&= \sum_{x,y\in\mathcal{S}} \mathrm{Tr}(|xy\rangle\langle xy|[\rho^{b=0} - \rho^{b=1}]) \\
&= \sum_{x,y\in\mathcal{S}} \mathrm{Tr}\left(|xy\rangle\langle xy|\left[\frac{2}{d(d+1)}\Pi^S - \frac{2}{d(d-1)}\Pi^A\right]\right) \qquad \text{Note: } \frac{2}{d(d-1)} = \frac{2}{d(d+1)} + \frac{4}{d(d-1)(d+1)} \\
&= \frac{2}{d(d+1)} \sum_{x,y\in\mathcal{S}} \mathrm{Tr}\left(|xy\rangle\langle xy|\left[\Pi^S - \Pi^A\right]\right) - \frac{4}{d(d-1)(d+1)} \sum_{x,y\in\mathcal{S}} \mathrm{Tr}(|xy\rangle\langle xy|\Pi^A) \\
&\leq \frac{2}{d(d+1)} \sum_{x,y\in\mathcal{S}} \mathrm{Tr}\left(F|xy\rangle\langle xy|\right) \qquad \text{Because } \Pi^A \text{projector } \Rightarrow 0 \leq \mathrm{Tr}(|xy\rangle\langle xy|\Pi^A) \leq 1 \\
&\leq \frac{2}{d(d+1)} \sum_x^d \sum_y^d |\langle x|y\rangle|^2 \qquad \text{Note: } |x\rangle = \sum_y^d \langle y|x\rangle|y\rangle \Rightarrow |\langle x|x\rangle|^2 = \sum_y |\langle y|x\rangle|^2 \\
&= \frac{2}{d(d+1)} \sum_x^d \langle x|x\rangle^2 = \frac{2}{d+1}.
\end{aligned}
$$

Exercise 6.2 Classical channels as TPCPMs.

a) *Take the binary symmetric channel* $\mathbf{p}$,



Recall that we can represent the probability distributions on both ends of the channel as quantum states in a given basis: for instance, if $P_X(0) = q, P_X(1) = 1-q$, we may express this as the 1-qubit mixed state $\rho_X = q\ |0\rangle\langle 0| + (1-q)\ |1\rangle\langle 1|$.

What is the quantum state ρ_Y that represents the final probability distribution P_Y in the computational basis?

We have

$$
\begin{aligned}
P_Y(0) &= \sum_x P_X(x) P_{Y|X=x}(0) = q(1-p) + (1-q)p \\
P_Y(1) &= qp + (1-q)(1-p),
\end{aligned}
$$

which can be expressed as a quantum state $\rho_y = [q(1-p) + (1-q)p]\ |0\rangle\langle 0| + [qp + (1-q)(1-p)]\ |1\rangle\langle 1| \in \mathcal{L}(\mathcal{H}_Y)$.

b) *Now we want to represent the channel as a map*

$$
\begin{aligned}
\mathcal{E}_\mathbf{p} : \mathcal{S}(\mathcal{H}_X) &\mapsto \mathcal{S}(\mathcal{H}_Y) \\
\rho_X &\to \rho_Y.
\end{aligned}
$$

An operator-sum representation (also called the Kraus-operator representation) of a CPTP map $\mathcal{E} : \mathcal{S}(\mathcal{H}_X) \to \mathcal{S}(\mathcal{H}_Y)$ is a decomposition $\{E_k\}_k$ of operators $E_k \in Hom(\mathcal{H}_X, \mathcal{H}_Y)$, $\sum_k E_k {E_k}^\dagger = \mathbb{1}$, such that

$$
\mathcal{E}(\rho_X) = \sum_k E_k \rho_X {E_k}^\dagger.
$$

Find an operator-sum representation of $\mathcal{E}_\mathbf{p}$.

We take four operators, corresponding to the four different "branches" of the channel,

$$
\begin{aligned}
E_{0\to 0} &= \sqrt{1-p}|0\rangle\langle 0| \\
E_{0\to 1} &= \sqrt{p}|1\rangle\langle 0| \\
E_{1\to 0} &= \sqrt{p}|0\rangle\langle 1| \\
E_{1\to 1} &= \sqrt{1-p}|1\rangle\langle 1|.
\end{aligned}
$$

To check that this works for the classical state ρ_X, we do

$$
\begin{aligned}
\mathcal{E}(\rho_X) =& \sum_{xy} E_{x\to y}\ \rho_X\ E^\dagger_{x\to y} \\
=& \sum_{xy} E_{x\to y}\ \Big[q|0\rangle\langle 0| + (1-q)|1\rangle\langle 1|\Big]\ E^\dagger_{x\to y} \\
=&(1-p)\ |0\rangle\langle 0|\Big[q|0\rangle\langle 0| + (1-q)|1\rangle\langle 1|\Big]|0\rangle\langle 0| \\
&+ p\ |1\rangle\langle 0|\Big[q|0\rangle\langle 0| + (1-q)|1\rangle\langle 1|\Big]|0\rangle\langle 1| \\
&+ p\ |0\rangle\langle 1|\Big[q|0\rangle\langle 0| + (1-q)|1\rangle\langle 1|\Big]|1\rangle\langle 0| \\
&+ (1-p)\ |1\rangle\langle 1|\Big[q|0\rangle\langle 0| + (1-q)|1\rangle\langle 1|\Big]|1\rangle\langle 1| \\
=&q(1-p)\ |0\rangle\langle 0| \\
&+ qp\ |1\rangle\langle 1| \\
&+ (1-q)p\ |0\rangle\langle 0| \\
&+ (1-q)(1-p)\ |1\rangle\langle 1| = \rho_Y.
\end{aligned}
$$

c) *Now we have a representation of the classical channel in terms of the evolution of a quantum state. What happens if the initial state ρ_X is not diagonal in the computational basis?*

In general, we can express the state in the computational basis as $\rho_X = \sum_{ij} \alpha_{ij}|i\rangle\langle j|$, with the usual conditions (positivity, normalization). Applying the map gives us

$$
\begin{aligned}
\mathcal{E}(\rho_X) =& \sum_{xy} E_{x\to y}\ \Big[\sum_{ij}\alpha_{ij}|i\rangle\langle j|\Big]\ E^\dagger_{x\to y} \\
=&(1-p)\ |0\rangle\langle 0|\Big[\sum_{ij}\alpha_{ij}|i\rangle\langle j|\Big]|0\rangle\langle 0| \\
&+ p\ |1\rangle\langle 0|\Big[\sum_{ij}\alpha_{ij}|i\rangle\langle j|\Big]|0\rangle\langle 1| \\
&+ p\ |0\rangle\langle 1|\Big[\sum_{ij}\alpha_{ij}|i\rangle\langle j|\Big]|1\rangle\langle 0| \\
&+ (1-p)\ |1\rangle\langle 1|\Big[\sum_{ij}\alpha_{ij}|i\rangle\langle j|\Big]|1\rangle\langle 1| \\
=&\alpha_{11}(1-p)\ |0\rangle\langle 0| + \alpha_{11}p\ |1\rangle\langle 1| \\
&+ \alpha_{22}p\ |0\rangle\langle 0| + \alpha_{22}(1-p)\ |1\rangle\langle 1|.
\end{aligned}
$$

Using $\alpha_{11} := \alpha, \alpha_{22} = 1-\alpha$, we get $\mathcal{E}(\rho_X) = [\alpha(1-p) + (1-\alpha)p]\ |0\rangle\langle 0| + [\alpha p + (1-\alpha)(1-p)]\ |1\rangle\langle 1|$. The channel ignores the off-diagonal terms of ρ_X: it acts as a measurement on the computational basis followed by the classical binary symmetric channel.

d) *Consider an arbitrary classical channel* $\mathbf{p}$ *from an n-bit space X to an m-bit space Y, defined by the conditional probabilities* $\left\{P_{Y|X=x}(y)\right\}_{xy}$.

Express $\mathbf{p}$ *as a map* $\mathcal{E}_{\mathbf{p}} : \mathcal{S}(\mathcal{H}_X) \to \mathcal{S}(\mathcal{H}_Y)$ *in the operator-sum representation.*

We generalize the previous result as

$$\begin{aligned}
\mathcal{E}_{\mathbf{p}}(\rho_X) &= \sum_{x,y} P_{Y|X=x}(y)\ |y\rangle\langle x|\rho_X|x\rangle\langle y| \\
&= \sum_{x,y} E_{x\to y}\rho_X E^\dagger x \to y, \quad E_{x\to y} = \sqrt{P_{Y|X=x}(y)}\ |y\rangle\langle x|.
\end{aligned}$$

To see that this works, take a classical state $\rho_X = \sum_x P_X(x)\ |x\rangle\langle x|$ as input,

$$\begin{aligned}
\mathcal{E}_{\mathbf{p}}(\rho_X) &= \sum_{x,y} P_{Y|X=x}(y)\ |y\rangle\langle x| \Big(\sum_{x'} P_X(x')\ |x'\rangle\langle x'| \Big) |x\rangle\langle y| \\
&= \sum_{x,y} P_{Y|X=x}(y)\ P_X(x)\ |y\rangle\langle y| \\
&= \sum_{y} P_y(y)\ |y\rangle\langle y|.
\end{aligned}$$

Exercise 6.3 TPCPMs as channels

Consider two single-qubit Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ and a TPCPM

$$\begin{aligned}
\mathcal{E}_p : \mathcal{S}(\mathcal{H}_X) &\mapsto \mathcal{S}(\mathcal{H}_Y) \\
\rho &\to p\frac{\mathbb{1}}{2} + (1-p)\rho.
\end{aligned}$$

a) *Find an operator-sum representation for $\mathcal{E}_p$.*

For simplicity of notation, we denote the Pauli matrices by X, Y, Z.

Remembering that $X^2 = Y^2 = Z^2 = \mathbb{1}$, $XY = -YX = Z$, $YZ = -ZY = X$ and $ZX = -XZ = Y$, you can verify that

$$\mathbb{1} = \frac{1}{2}(\rho + X\rho X + Y\rho Y + Z\rho Z).$$

From this follows the operator sum representation $\{M_x\}_x$,

$$M_1 = \sqrt{1 - \frac{3p}{4}}\ \mathbb{1}, \quad M_2 = \frac{\sqrt{p}}{2}X, \quad M_3 = \frac{\sqrt{p}}{2}Y, \quad M_4 = \frac{\sqrt{p}}{2}Z.$$

b) *What happens to the radius $\vec{r}$ when we apply $\mathcal{E}_p$? What is the physical interpretation of this?*

Using the result from part *a*) we have

$$\begin{aligned}
\mathcal{E}(\rho) &= \frac{p}{2}\mathbb{1} + (1-p)\ \rho \\
&= \frac{1}{2}(\mathbb{1} + (1-p)\ \vec{r}\cdot\vec{X})
\end{aligned}$$

Thus, points on a sphere with radius r are mapped to a smaller sphere with radius $(1-p)r$ — they get more mixed. In particular, pure states will not remain pure under this CPM.

c) *Now we will see what happens when we use this quantum channel to send classical information. We start with an arbitrary input probability distribution $P_X(0) = q, P_X(1) = 1-q$. We encode this distribution in a state $\rho_X = q\ |0\rangle\langle 0| + (1-q)|1\rangle\langle 1|$. Now we send ρ_X over the quantum channel, i.e., we let it evolve under $\mathcal{E}_{\mathbf{p}}$. Finally, we measure the output state, $\rho_Y = \mathcal{E}_{\mathbf{p}}(\rho_X)$ in the computational basis. Compute the conditional probabilities $\{P_{Y|X=x}(y)\}_{xy}$.*

Applying the map to this state results in

$$\begin{aligned}
\mathcal{E}(\rho_X) &= \Big(\frac{p}{2} + (1-p)q\Big)\ |0\rangle\langle 0| + \Big(\frac{p}{2} + (1-p)(1-q)\Big)\ |1\rangle\langle 1| \\
&= P_Y(0)\ |0\rangle\langle 0| + P_Y(1)\ |1\rangle\langle 1|,
\end{aligned}$$

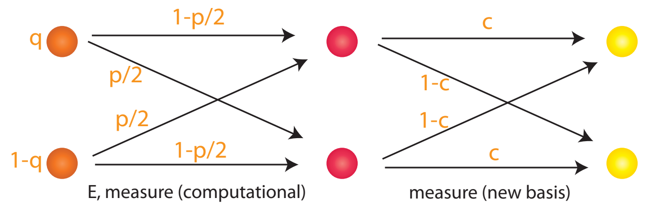


Figure 1: The result is a binary symmetric channel with $p' = 1 - c - p/2 + pc$.

so $P_Y(0) = \frac{p}{2} + (1-p)q, P_Y(1) = \frac{p}{2} + (1-p)(1-q)$. The conditional probabilities can be arranged in a transition matrix $(T)_{xy} = P_{Y|X=x}(y)$ as follows:

$$T = \begin{pmatrix} \frac{p}{2} + (1-p) & \frac{p}{2} \\ \frac{p}{2} & \frac{p}{2} + (1-p) \end{pmatrix} = \begin{pmatrix} 1 - \frac{p}{2} & \frac{p}{2} \\ \frac{p}{2} & 1 - \frac{p}{2} \end{pmatrix}.$$

We obtained the binary symmetric channel, with $p' = p/2$.

d) *Maximize the mutual information over q to find the classical channel capacity of the depolarizing channel.*

The channel capacity of the binary symmetric channel, as has been shown in a previous exercise, is given by

$$C = 1 - H_{\text{bin}}(p/2), \quad H_{\text{bin}}(r) = -\left(r \log r + (1-r)\log(1-r)\right), \quad r \in [0,1].$$

e) *What happens to the channel capacity if we measure the final state in a different basis?*

Take an arbitrary basis $\{|\alpha\rangle, |\alpha^\perp\rangle\}$, where

$$|\alpha\rangle = \cos(\alpha)|0\rangle + \sin(\alpha)|1\rangle, \quad |\alpha^\perp\rangle = \cos\left(\alpha + \frac{\pi}{2}\right)|0\rangle + \sin\left(\alpha + \frac{\pi}{2}\right)|1\rangle = -\sin\alpha|0\rangle + \cos\alpha|1\rangle.$$

Then

$$\begin{aligned} P_Y(\alpha) &= \operatorname{Tr}\left[|\alpha\rangle\langle\alpha|\ \mathcal{E}(\rho_X)\right] = \operatorname{Tr}\left[\begin{pmatrix} \cos^2\alpha & \cos\alpha\sin\alpha \\ \cos\alpha\sin\alpha & \sin^2\alpha \end{pmatrix}\begin{pmatrix} P_Y(0) & 0 \\ 0 & P_Y(1) \end{pmatrix}\right] \\ &= \cos^2(\alpha)P_Y(0) + \sin^2(\alpha)P_Y(1), \\ P_Y(\alpha^\perp) &= \operatorname{Tr}\left[|\alpha^\perp\rangle\langle\alpha^\perp|\ \mathcal{E}(\rho_X)\right] = \operatorname{Tr}\left[\begin{pmatrix} \sin^2\alpha & -\cos\alpha\sin\alpha \\ -\cos\alpha\sin\alpha & \cos^2\alpha \end{pmatrix}\begin{pmatrix} P_Y(0) & 0 \\ 0 & P_Y(1) \end{pmatrix}\right] \\ &= \sin^2(\alpha)P_Y(0) + \cos^2(\alpha)P_Y(1). \end{aligned}$$

We can see this result in the following way: take $c = \cos^2(\alpha)$. Then " preparing $q|0\rangle\langle 0| + (1-q)|1\rangle\langle 1|$, applying $\mathcal{E}_p$ and measuring in basis $\{|\alpha\rangle, |\alpha^\perp\rangle\}$" is equivalent to the concatenation of two binary symmetric channels (Fig. 1).

The final probability distributions are the same if we apply $\mathcal{E}_p$, measure in the computational basis, and then measure again in the new basis. This holds because $\mathcal{E}_p$ does not change the eigenbasis of the state, and is not necessarily true for a general TPCPM.

The capacity of the original channel is larger than the capacity of the concatenation of the two channels (because adding another channel just adds more noise, a fact otherwise known as the data processing inequality).