Exercise 1. ***POVMs are the Most General Quantum to Classical Evolutions.***

To motivate why we consider POVMs in quantum information theory, we will show in this exercise that they capture the most general evolution of a quantum system into a classical register. A classical system X, in the quantum information formalism, is a quantum system (with Hilbert space $\mathscr{H}_X$) which is in a state ρ_X known to be diagonal in a fixed basis $\{|x\rangle\}$.

Let $\mathcal{E}_{A\to X} : \mathscr{L}(\mathscr{H}_A) \to \mathscr{L}(\mathscr{H}_X)$ be a trace-preserving, completely positive map from a (finite dimensional) quantum system A into a classical register X.

Show that this evolution is described by a POVM $\{A_x\}$ with the required properties, i.e. the probability of the (classical) output state to be in $|x\rangle\langle x|$ is $\operatorname{tr}(A_x\rho)$.

The following steps might help you, but it is not mandatory to follow them.

(a) Argue that $\mathcal{E}$ has to take the following form:

$$\mathcal{E}_{A\to X}(\rho) = \sum_x |x\rangle\langle x| \, f_x(\rho) \ , \tag{1}$$

where $f_x : \mathscr{L}(\mathscr{H}_A) \to \mathbb{R}$ is a linear mapping of ρ onto real numbers. $f_x(\rho)$ are the eigenvalues of $\mathcal{E}(\rho)$ in the eigenbasis $\{|x\rangle\}$ (which is fixed because, remember, X is a classical register).

Solution. Since the output of the channel has to be diagonal in the basis $\{|x\rangle\}$, we can write the general expression (1), where f_x is a function of ρ giving the eigenvalue of the output of $\mathcal{E}$ corresponding to the eigenvector $|x\rangle$. Since $\mathcal{E}$ is linear, it follows that f_x has to be linear.

Also, since the output of $\mathcal{E}$ is a density operator, the output of f_x has to be a real number between 0 and 1 (those are the possible eigenvalues of density operators). In addition, by the condition that density operators have unit trace, the values of f_x for fixed ρ must sum up to one, $\sum_x f_x(\rho) = 1$.

(b) Argue that $f_x(\cdot)$ can be written in general as

$$f_x(\cdot) = \operatorname{tr}[\, A_x(\cdot)\,] \ , \tag{2}$$

for some hermitian operator A_x.

Hints. $\operatorname{tr}\left[A^\dagger B\right]$ *is the Hilbert-Schmidt scalar product in the space of linear operators* $\mathscr{L}(\mathscr{H}_A)$. *Also, what kind of object is* f_x*?*

Solution. The functional f_x is an element of the dual space of linear operators on $\mathscr{H}_A$ by definition (the dual space of the vector space V is the space of all functionals $V \to \mathbb{C}$ on V). Any element $\bar{v}$ of a (finite dimensional) dual space to the vector space V can be written as a scalar product operation $\langle v, \cdot\rangle$ with a fixed element v of the vector space.

Since f_x is an element of the dual vector space of $\mathscr{L}(\mathscr{H}_A)$, then in general there exists an element A_x of the vector space $\mathscr{L}(\mathscr{H}_A)$ such that

$$f_x(\cdot) = \langle A_x\,, (\cdot)\rangle = \operatorname{tr}(A_x(\cdot)) \ , \tag{S.1}$$

where we recall that the Hilbert-Schmidt product $\langle A, B\rangle = \operatorname{tr}\left(A^\dagger B\right)$ is a scalar product on $\mathscr{L}(\mathscr{H}_A)$.

A_x can be assumed to be hermitian because the input to f_x is always hermitian (density operators are hermitian), so if one had chosen A_x not hermitian, we could replace it by $A'_x = \frac{1}{2}\left(A_x + A_x^\dagger\right)$ which is obviously hermitian.

(c) Argue that for all ρ, f_x has to take positive values and that the values for all x have to sum up to 1, $\sum_x f_x(\rho) = 1$. Deduce that $A_x \geqslant 0$ and $\sum_x A_x = \mathbb{1}$.

Solution. We have argued in point (a) that the output of f_x is a real number between 0 and 1, with $\sum_x f_x(\rho) = 1$ for any fixed ρ.

The expression $\operatorname{tr}(A_x\rho)$ is positive for all ρ if and only if A_x is positive semidefinite (otherwise, if for a $|\psi\rangle$ we had $\langle\psi|A_x|\psi\rangle < 0$, then it would follow that $\operatorname{tr}(A_x|\psi\rangle\langle\psi|) < 0$). So A_x has to be positive semidefinite.

Likewise, $\sum_x \operatorname{tr}(A_x\rho) = 1$ for all ρ implies $\sum_x A_x = \mathbb{1}$. Indeed, for any basis $\{|\psi_k\rangle\}$, we have $1 = \operatorname{tr}\left(\left(\sum_x A_x\right)|\psi_k\rangle\langle\psi_k|\right) = \langle\psi_k|\sum_x A_x|\psi_k\rangle$, so that the operator $\sum_x A_x$ has 1's on its diagonal in any basis. The only operator satisfying this is the identity operator.

(d) Conclude from points (a)–(c).

Solution. We have shown that the operators A_x have all the necessary properties for forming a POVM (they are positive semidefinite and sum up to the identity). In addition, they correctly reproduce the outcome probabilities (diagonal elements of the output of $\mathcal{E}$) as $\operatorname{tr}(A_x\rho)$.

Exercise 2. *Distinguishing two quantum states*

Suppose you know the density operators of two quantum states $\rho, \sigma \in \mathscr{H}_A$. Then you are given one of the states at random—it may either be ρ or σ, with probability 1/2. The challenge is to perform a single measurement on your state and then guess which state that is.

(a) What is your best strategy? In which basis do you think you should perform the measurement? Can you express that measurement using a projector Q?

Hint. You can use the idea of exercise 1. What are you looking for? What should be the measurement outcome?

Solution. We are looking for a strategy to guess either if the state was ρ or σ, i.e. we need a mapping from the quantum system onto one classical bit of information, "guess ρ" or "guess σ". However, we have shown in Ex. 1 that the most general mapping of a quantum system to a classical register is precisely a POVM.

Denote the POVM elements by Q_ρ (for "guess ρ") and Q_σ (for "guess σ"). We need $Q_\rho + Q_\sigma = \mathbb{1}$, so $Q_\sigma = \mathbb{1} - Q_\rho$ (This is by definition of a POVM, or actually, it is needed in order to conserve probability).

We have reformulated the problem as follows: we are looking for a POVM, with elements Q_ρ and $\mathbb{1} - Q_\rho$, such that the total probability of guessing right is maximized.

The total probability of guessing right is given by

$$
\begin{aligned}
\Pr[\text{distinguish correctly}] &= \Pr[\rho \text{ is given}] \times \Pr[\text{measure outcome } Q_\rho \text{ from } \rho] \\
&\quad + \Pr[\sigma \text{ is given}] \times \Pr[\text{measure outcome } Q_\sigma \text{ from } \sigma] \\
&= \frac{1}{2}\operatorname{tr}(Q_\rho\rho) + \frac{1}{2}\operatorname{tr}((\mathbb{1} - Q_\rho)\sigma) \\
&= \frac{1}{2}\operatorname{tr}(Q_\rho\rho + \sigma - Q_\rho\sigma) = \frac{1}{2} + \frac{1}{2}\operatorname{tr}(Q_\rho(\rho - \sigma)) \ . \qquad \text{(S.2)}
\end{aligned}
$$

So we need to find the Q_ρ that maximizes the expression $\operatorname{tr}(Q_\rho(\rho - \sigma))$.

Choose a representation in terms of the eigenstates $|\eta_k\rangle$ of the operator $\rho - \sigma$,

$$(\rho - \sigma)|\eta_k\rangle = \eta_k|\eta_k\rangle \ . \qquad \text{(S.3)}$$

Note that the operator $\rho - \sigma$ is not a density operator. It is, however, hermitian and has trace zero.

We want to maximize the expression

$$\mathrm{tr}\left(Q_\rho\left(\rho-\sigma\right)\right) = \sum_k \eta_k \,\mathrm{tr}\left(Q_\rho|\eta_k\rangle\langle\eta_k|\right) = \sum_k \eta_k \langle\eta_k|Q_\rho|\eta_k\rangle\ . \tag{S.4}$$

The maximum value is then obtained with the choice (recall that we have the constraint $0 \leqslant Q_\rho \leqslant \mathbb{1}$)

$$\langle\eta_k|Q_\rho|\eta_k\rangle = 1 \quad \text{if } \eta_k \geqslant 0\ ; \tag{S.5}$$

$$\langle\eta_k|Q_\rho|\eta_k\rangle = 0 \quad \text{if } \eta_k < 0\ . \tag{S.6}$$

Thus our optimal Q_ρ is the projector onto the eigenspace for the positive eigenvalues of the operator $\rho-\sigma$.

The maximization is taken from [Helstrom, C. W., *Quantum Detection and Estimation Theory*, Journal of Statistial Physics, 1(2):231–252, 1969]. In the latter, discussion of quantum hypothesis testing is treated with more generality.

(b) Show that this optimal probability is directly related to the trace distance,

$$\Pr\left[\text{distinguish correctly}\right] = \frac{1}{2}\left[1+\delta\left(\rho,\sigma\right)\right]\ . \tag{3}$$

Solution. Remember that the trace distance between states ρ and σ is given by

$$\delta(\rho,\sigma) = \frac{1}{2}\|\rho-\sigma\|_1 = \frac{1}{2}\,\mathrm{tr}|\rho-\sigma|\ , \tag{S.7}$$

where $\|A\|_1 = \mathrm{tr}|A|$ is simply the sum of the absolute values of the eigenvalues of A (this norm is also called the *Shatten-1 norm*).

We just have to show that the optimal value from point (a) of this exercise satisfies (3).

We know from point (a) that

$$\Pr\left[\text{distinguish correctly}\right] = \frac{1}{2}+\frac{1}{2}\,\mathrm{tr}\left(Q_\rho\left(\rho-\sigma\right)\right) = \frac{1}{2}+\frac{1}{2}\sum_{k:\,\eta_k\geqslant 0}\eta_k\ . \tag{S.8}$$

On the other hand, we have

$$\delta\left(\rho,\sigma\right) = \frac{1}{2}\sum_k|\eta_k| = \frac{1}{2}\Big(\sum_{\eta_k\geqslant 0}\eta_k - \sum_{\eta_k<0}\eta_k\Big)\ . \tag{S.9}$$

However, since $\rho-\sigma$ has trace zero, we have $\sum\eta_k = 0$ and thus $\sum_{\eta_k<0}\eta_k = -\sum_{\eta_k\geqslant 0}\eta_k$. So

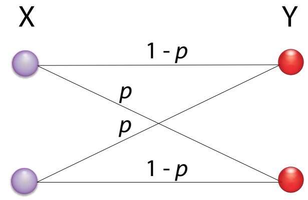
$$(\text{S.9}) = \frac{1}{2}\sum_{\eta_k\geqslant 0}\eta_k + \frac{1}{2}\sum_{\eta_k\geqslant 0}\eta_k = \sum_{\eta_k\geqslant 0}\eta_k\ . \tag{S.10}$$

Combining with (S.8), we eventually obtain (3).

Exercise 3. *Classical channels as trace-preserving completely positive maps.*

In this exercise we will see how to represent classical channels as trace-preserving completely positive maps (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?

Solution. 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 $\rho_y = [q(1-p)+(1-q)p]\ |0\rangle\langle 0| + [qp+(1-q)(1-p)]\ |1\rangle\langle 1|\ \in \mathcal{L}(\mathscr{H}_Y)$.

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

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

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

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

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

Hint. Think of each operator $E_k = E_{xy}$ as the representation of the branch that maps input x to output y.

Solution. 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?

Solution. 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) Now 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 $\mathcal{E}_\mathbf{p} : \mathcal{S}(\mathscr{H}_X) \to \mathcal{S}(\mathscr{H}_Y)$ in the operator-sum representation.

Solution. 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,

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