Exercise 1. *Trace distance and distinguishability*

Suppose you know the density operators of two quantum states $\rho, \sigma \in \mathcal{S}(\mathcal{H})$. Then you are given one of the states at random – it may either be ρ or σ with equal probability. The challenge is to perform a single projective measurement of an observable O 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 single projector P?*

(b) *Show that the probability of guessing correctly can be written as*

$$P_{\text{guess}}(\rho \text{ vs. } \sigma) = \frac{1}{2}(1 + \operatorname{tr}[P(\rho - \sigma)]), \tag{1}$$

where P is the appropriate projector from (a).

Just like in the classical case, that can be shown to be equivalent to

$$P_{\text{guess}}(\rho \text{ vs. } \sigma) = \frac{1}{2}[1 + \delta(\rho, \sigma)], \tag{2}$$

where $\delta(\rho, \sigma) := \frac{1}{2}\|\rho - \sigma\|_1$ is the trace distance between the two quantum states and $\|S\|_1 := \operatorname{tr}|S| \equiv \operatorname{tr}[\sqrt{S^\dagger S}]$ the 1-norm for matrices. (You do not have to show this here. Of course you can if you want.)

(c) *Given a trace-preserving quantum operation $\mathcal{E}$ (i.e. a CPTP map) and two states ρ and σ, show that*

$$\delta(\mathcal{E}(\sigma), \mathcal{E}(\rho)) \leq \delta(\sigma, \rho). \tag{3}$$

(d) *What does (3) imply about the task of distinguishing quantum states?*

Solution.

(a) We are looking for a measurement O that maximises our probability of guessing correctly. For each state (say e.g. ρ) the probabilities of obtaining any of the possible outcomes $\{y\}_y$ of the observable $O = \sum_y y P_y$ that represents the measurement define a classical probability distribution $\Pr_{O,\rho}(y) = \operatorname{tr}(P_y \rho)$, P_y being the mutually orthogonal projectors onto the eigenspaces of O.

Let $G = \{y : P_{O,\rho}(y) \geq P_{O,\sigma}(y)\}$ be the set of outcomes that are more likely to occur when we measure O on ρ than on σ. Naturally, if we obtain y after measuring our unknown state and obtain we should say it was ρ if $y \in G$ and vice-versa. The probability of guessing correctly is then

$$\begin{aligned} P_{\text{guess}} &= P(\rho)\cdot P(\text{say "}\rho\text{"}|\rho) + P(\sigma)\cdot P(\text{say "}\sigma\text{"}|\sigma) \\ &= \frac{1}{2}\cdot\sum_{y\in G} P_{O,\rho}(y) + \frac{1}{2}\cdot\sum_{y\in\bar{G}} P_{O,\sigma}(y) \\ &= \frac{1}{2}\sum_{y\in G}\operatorname{tr}(P_y\rho) + \frac{1}{2}\sum_{y\in\bar{G}}\operatorname{tr}(P_y\sigma) \\ &= \frac{1}{2}\operatorname{tr}\Big(\Big[\sum_{y\in G} P_y\Big]\rho\Big) + \frac{1}{2}\operatorname{tr}\Big(\Big[\sum_{y\in\bar{G}} P_y\Big]\sigma\Big) \\ &= \frac{1}{2}\operatorname{tr}\left(P_G\ \rho + P_{\bar{G}}\ \sigma\right), \end{aligned} \tag{S.1}$$

where $P_G := \sum_{y\in G} P_y$ and $P_{\bar{G}} := \sum_{y\in\bar{G}} P_y$ are projectors too, with $P_G + P_{\bar{G}} = \mathbb{1}$.

If we explore a little more, we obtain

$$\begin{aligned} 2P_{\text{guess}} &= \operatorname{tr}(P_G\ \rho + P_{\bar{G}}\ \sigma) \\ &= \operatorname{tr}(P_G\ \rho + [\mathbb{1} - P_G]\ \sigma) \\ &= \operatorname{tr}(P_G\ [\rho-\sigma]) + \operatorname{tr}(\mathbb{1}\sigma) \\ &= \operatorname{tr}(P_G\ [\rho-\sigma]) + 1, \qquad {}^{(*)} \end{aligned} \tag{S.2}$$

where $^{(*)}$ comes from the fact that σ is a density matrix and therefore $\operatorname{tr}(\sigma) = 1$.

Notice that we have only defined G depending on O so far. Hence, to maximise the guessing probability we need to find the optimal $\{P_y\}_y$ that maximise $\operatorname{tr}(P_G\ [\rho-\sigma])$.

First we express G in another way using linearity of the trace,

$$\begin{aligned} G &= \{y : P_{O,\rho}(y) \geq P_{O,\sigma}(y)\} \\ &= \{y : \operatorname{tr}(P_y\rho) \geq \operatorname{tr}(P_y\sigma)\} \\ &= \{y : \operatorname{tr}(P_y(\rho-\sigma)) \geq 0\}\,. \end{aligned} \tag{S.3}$$

Now we will try a clever choice of G. Let $\{y\}_y$ be the eigenbasis of $\rho - \sigma = \sum_y \lambda_y |y\rangle\langle y|$. Notice that $\rho - \sigma$ is *not* a density matrix – in particular it has trace zero. If we choose $\{P_y\}_y$ to be the projectors on that basis, $P_y = |y\rangle\langle y|$, we obtain

$$\begin{aligned} G &= \{y : \operatorname{tr}(P_y(\rho-\sigma)) \geq 0\} \\ &= \Big\{y : \operatorname{tr}\Big(|y\rangle\langle y| \sum_{y'} \lambda_{y'} |y'\rangle\langle y'|\Big) \geq 0\Big\} \\ &= \{y : \operatorname{tr}\left(|y\rangle\langle y|\lambda_y\right) \geq 0\} \\ &= \{y : \lambda_y \geq 0\}\ , \end{aligned} \tag{S.4}$$

i.e. G is the set of outcomes of O corresponding to projectors on states $|y\rangle\langle y|$ that correspond to non negative eigenvalues of $\rho - \sigma$. In this case, $\operatorname{tr}(P_G\ [\rho-\sigma])$ is the sum of all positive eigenvalues of $\rho - \sigma$.

This result is promising, but now we have to prove that is is indeed optimal, i.e. that no other choice of projector P could give better results. We can write $\rho - \sigma$ as $R - S$, where $R = \sum_{y\in G} \lambda_y |y\rangle\langle y|$ and $S = \sum_{y\in\bar{G}} -\lambda_y |y\rangle\langle y|$. Both operators R and S are positive and diagonal. Furthermore they are mutually orthogonal because $\{|y\rangle\}$ is an orthogonal basis.

We have that

$$\mathrm{tr}(P_G\ [\rho - \sigma]) = \sum_{y \in G} \lambda_y = \mathrm{tr}(R). \tag{S.5}$$

For any other projector P', however,

$$\begin{aligned}\mathrm{tr}(P'\ [\rho - \sigma]) &= \mathrm{tr}(P'\ [R - S]) \\ &= \mathrm{tr}(P'\ R) - \mathrm{tr}(P'\ S) \\ &\leq \mathrm{tr}(R) - \mathrm{tr}(P'\ S) \qquad {}^{(**)} \\ &\leq \mathrm{tr}(R), \qquad {}^{(***)}\end{aligned} \tag{S.6}$$

where $^{(**)}$ stands because projectors can only decrease the trace and $^{(***)}$ because $P'S$ is positive by assumption.

We have proved that a measurement represented by $O = \sum_y y|y\rangle\langle y|$, where $\{|y\rangle\}_y$ is the eigenbasis of $\rho - \sigma$ optimises the probability of guessing correctly which state we were given.

This solution corresponds to the following strategy. We measure our state (ρ or σ) in the eigenbasis of $\rho - \sigma$. If we obtain a state that corresponds to a positive eigenvalue of $\rho - \sigma$ (i.e. $y \in G$) then it is more likely that we have measured ρ. If we get a negative eigenvalue of $\rho - \sigma$ (i.e. $y \in \bar{G}$) we should say the state was σ.

In the particular case where the two density operators share the same eigenbasis, this corresponds to following the classical strategy for distinguishing two probability distributions after measuring the state in their common eigenbasis.

(b) We already proved that in the previous exercise, (S.2) and the following.

(c) In (a) we have shown constructively how to write the difference between two quantum states, $\rho - \sigma$, as $R - S$, where R and S are two positive operators with orthogonal support. We now use this fact to write $|\rho - \sigma| = R + S$ and obtain

$$\begin{aligned}\delta(\sigma, \rho) &= \frac{1}{2}\mathrm{tr}(|\rho - \sigma|) \\ &= \frac{1}{2}\left(\mathrm{tr}(R) + \mathrm{tr}(S)\right) \\ &= \mathrm{tr}(R) \qquad {}^{(*)} \\ &= \mathrm{tr}\left[\mathcal{E}(R)\right] \\ &\geq \max_P \left\{\mathrm{tr}\left[P\mathcal{E}(R)\right] - \mathrm{tr}\left[P\mathcal{E}(S)\right]\right\} \qquad {}^{(**)} \\ &= \max_P \mathrm{tr}\left[P(\mathcal{E}(R - S))\right] \qquad {}^{(***)} \\ &= \delta\left(\mathcal{E}(\sigma), \mathcal{E}(\rho)\right),\end{aligned} \tag{S.7}$$

where $^{(*)}$ stands because

$$\mathrm{tr}(R) - \mathrm{tr}(S) = \mathrm{tr}(R - S) = \mathrm{tr}(\rho - \sigma) = \mathrm{tr}(\rho) - \mathrm{tr}(\sigma) = 1 - 1 = 0,$$

and the inequality $^{(**)}$ follows from $\mathrm{tr}(P\mathcal{E}(R)) \leq \mathrm{tr}(\mathcal{E}(R))$ and $\mathrm{tr}(P\mathcal{E}(S)) \geq 0$ for any projector P, since projectors are positive operators and can only decrease the trace. Finally, linearity of TPMs allows us to perform step $^{(***)}$. We also used the characterization of the trace distance in terms of a maximization over projectors, see (2), in the very last step.

(d) This result implies that there is no experimental setup that allows us to distinguish non-orthogonal states with certainty (because whatever this setup is, its action on the quantum states can be described by some CPTPM). If there was such a setup, we could copy (clone) the states perfectly, hence the contradiction. In fact, the trace distance (as we have seen in the lecture) gives us an upper limit on our ability to distinguish them. If there were quantum operations that increase the distance between two states, we could design measurement devices such that this upper limit no longer holds.

Exercise 2. ***Fidelity and Uhlmann's Theorem***

Given two states ρ_A and σ_A on $\mathcal{H}_A$ with fixed basis $\{|i\rangle_A\}_i$ and a reference Hilbert space $\mathcal{H}_B$ with fixed basis $\{|i\rangle_B\}_i$, which is a copy of $\mathcal{H}_A$, Uhlmann's theorem claims that the fidelity can be written as

$$F(\rho_A, \sigma_A) = \max_{|\Psi\rangle_{AB}, |\Phi\rangle_{AB}} |\langle\Psi|\Phi\rangle| \,, \tag{4}$$

where the maximum is over all purifications $|\Psi\rangle_{AB}$ of ρ_A and $|\Phi\rangle_{AB}$ of σ_A on $\mathcal{H}_A \otimes \mathcal{H}_B$. Let us introduce the state $|\psi\rangle_{AB}$ as

$$|\psi\rangle = (\sqrt{\rho} \otimes U_B)\,|\Omega\rangle, \qquad |\Omega\rangle = \sum_i |i\rangle_A \otimes |i\rangle_B, \tag{5}$$

where U_B is any unitary on $\mathcal{H}_B$. We have seen in Exercise Sheet 6 that $|\psi\rangle_{AB}$ is a purification of ρ_A and that any purification of ρ_A can be written in this form.

(a) *Use the construction presented in the proof of Uhlmann's theorem to calculate the fidelity between $\sigma'_A = \mathbb{1}_2/2$ and $\rho'_A = p|0\rangle\langle 0|_A + (1-p)|1\rangle\langle 1|_A$ in the 2-dimensional Hilbert space.*

Hint: *Convince yourself that the vector $|\Omega\rangle$ has the property that $\mathbb{1} \otimes S\,|\Omega\rangle = S^T \otimes \mathbb{1}|\Omega\rangle$ for all linear operators S on $\mathcal{H}_A$.*

(b) *Give an expression for the fidelity between any pure state and the completely mixed state $\mathbb{1}_n/n$ in the n-dimensional Hilbert space.*

Hint: *You may want to use a different characterization of the fidelity than the one by Uhlmann for this exercise.*

Solution.

(a) Because of the special form in which we can write purifications, it is apparent that it is sufficient to maximise over one set of purifications only. We set

$$\begin{aligned} |\Psi\rangle &= (\sqrt{\rho'} \otimes V_B)|\Omega\rangle \,, \\ |\Phi\rangle &= \frac{1}{\sqrt{2}}(\mathbb{1}_A \otimes \mathbb{1}_B)|\Omega\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \end{aligned} \tag{S.8}$$

for some unitary V on B (which can also be seen as a unitary on A because A and B are isomorphic).

It follows that

$$
\begin{aligned}
|\langle\Psi|\Phi\rangle| &= \frac{1}{\sqrt{2}} \left|\langle\Omega|\sqrt{\rho'} \otimes V_B|\Omega\rangle\right| \\
&= \frac{1}{\sqrt{2}} \left|\mathrm{tr}\left[\sqrt{\rho'} \otimes V_B|\Omega\rangle\langle\Omega|\right]\right| \\
&= \frac{1}{\sqrt{2}} \left|\mathrm{tr}\left[\sqrt{\rho'} \cdot V_A^T \otimes \mathbb{1}_B|\Omega\rangle\langle\Omega|\right]\right| \qquad (*) \\
&= \frac{1}{\sqrt{2}} \left|\mathrm{tr}\left[\sqrt{\rho'} \cdot V_A^T\right]\right| \\
&\leq \frac{1}{\sqrt{2}} \mathrm{tr}\left[\left|\sqrt{\rho'}\right|\right] \\
&= \frac{1}{\sqrt{2}} \left(\sqrt{p} + \sqrt{1-p}\right).
\end{aligned} \tag{S.9}
$$

For $^{(*)}$ we used the fact that the state $|\Omega\rangle$ is such that applying V to its B part is equivalent to applying V^T to its A part. (Please check this explicitly for yourself if unclear.) Also, after $^{(*)}$ we used that $\mathrm{tr}[S_{AB}] = \mathrm{tr}[\mathrm{tr}_B[S_{AB}]]$.

The maximum can be achieved when V_A^T produces the polar decomposition of $\sqrt{\rho'}$ – which in this case is trivially $V_A = \mathbb{1}_A$. We obtain $F(\rho', \sigma') = (\sqrt{p} + \sqrt{1-p})/\sqrt{2}$.

(b) The general case follows immediately from the original definition of the fidelity:

$$
\begin{aligned}
F(\rho, \sigma) &= \mathrm{tr}\left[\sqrt{\sqrt{\sigma'}\rho'\sqrt{\sigma'}}\right] \\
&= \frac{1}{\sqrt{n}} \mathrm{tr}\left[\sqrt{\rho'}\right] \\
&= \frac{1}{\sqrt{n}}\,.
\end{aligned} \tag{S.10}
$$

The last equality follows from the fact that $\sqrt{\rho'} = \rho'$ for pure states.

Exercise 3. *An interpretation of the quantum trace distance*

In Exercise Sheet 3 we have seen an interpretation of the classical trace distance. We have shown that two probability distributions that are ε-close in trace distance allow for a joint distribution s.t. the corresponding reduced random variables differ with probability at most ε. In the quantum case, where probability distributions are replaced by quantum states, say ρ and $\sigma \in \mathcal{S}(\mathcal{H})$, this statement does not have a direct translation. Instead, as we will see in this exercise, there is a similar but different way of interpreting the trace distance.

(a) Thinking of the classical version from Exercise Sheet 3 again, what goes wrong when trying to 'quantize' this interpretation directly? Why does this not work?

Suppose that $\delta(\rho, \sigma) = \varepsilon$. We will show that there is a quantum state $\omega \in \mathcal{S}(\mathcal{H})$ that can be written in two ways,

$$
\begin{aligned}
\omega &= (1-\varepsilon)\rho + \varepsilon\hat{\rho} \\
&= (1-\varepsilon)\sigma + \varepsilon\hat{\sigma},
\end{aligned} \tag{6}
$$

where $\hat{\rho}, \hat{\sigma} \in \mathcal{S}(\mathcal{H})$ are some quantum states.

(b) *Use the fact that the operator $\rho - \sigma$ can be decomposed into $\rho - \sigma = R - S$, where both R and S are positive operators with mutually orthogonal support, and $\mathrm{tr}[R] = \varepsilon$, to construct ω.*

The above statement has two interpretations: (i) there exists a state ω that behaves as if it was ρ with probability $1-\varepsilon$; (ii) the same state ω behaves exactly like σ with probability $1-\varepsilon$.

(c) *Can you construct a classical example of this interpretation in the language of density operators to illustrate the connection to the classical version?*

Solution.

(a) For completeness we repeat the classical statement once more. Let P_X and $P_{X'}$ be two probability distributions on the same alphabet with $\delta(P_X, P_{X'}) = \varepsilon$. Then there is a joint distribution $\bar{P}_{XX'}$ s.t. $\bar{P}_X = P_X$, $\bar{P}_{X'} = P_{X'}$, and $\bar{P}[X \neq X'] \leq \varepsilon$.

A quantum version of this would now work with density operators instead of probability distributions and could state that if for $\rho, \sigma \in \mathcal{S}(\mathcal{H})$: $\delta(\rho, \sigma) = \varepsilon$, then there is a joint state $\omega \in \mathcal{S}(\mathcal{H} \otimes \mathcal{H})$ s.t. $\mathrm{tr}_2[\omega] = \rho$, $\mathrm{tr}_1[\omega] = \sigma$, and .... what? This is where it goes wrong. It is unclear how to translate the event $\{X \neq X'\}$ to the quantum setting. To do so, one would have to talk about measurements (projective or POVMs) and outcomes of those, only then we can speak of 'events'.

(b) As in Exercise 1 we use that $\rho - \sigma = R - S$ for positive R and S with orthogonal support, and $\mathrm{tr}[R] = \mathrm{tr}[S] = \varepsilon$. Obviously the statement is trivial for $\varepsilon = 1$. Also, because δ is a metric, the case $\varepsilon = 0$ implies that $\rho = \sigma$, and again the statement is straight forward. Hence, for the rest of this exercise we assume that $\varepsilon \in (0, 1)$.

One possibility to construct ω is to choose $\omega := (1-\varepsilon)\rho + S$, i.e. $\hat{\rho} = S/\varepsilon$We claim that this can be written as $\omega = (1-\varepsilon)\sigma + \varepsilon\hat{\sigma}$ for some state $\hat{\sigma}$. Using $\rho - \sigma = R - S$ we can rewrite ω as

$$\begin{aligned} \omega &= (1-\varepsilon)(\sigma + R - S) + S \\ &= (1-\varepsilon)\sigma + (1-\varepsilon)R + \varepsilon S \\ &= (1-\varepsilon)\sigma + \varepsilon\hat{\sigma}\,, \end{aligned} \tag{S.11}$$

where $\hat{\sigma} = \frac{1-\varepsilon}{\varepsilon}R + S$. So we are left with showing that $\hat{\sigma}$ is indeed a state. This is easy, because

$$\mathrm{tr}[\hat{\sigma}] = \frac{1-\varepsilon}{\varepsilon}\mathrm{tr}[R] + \mathrm{tr}[S] = 1 - \varepsilon + \varepsilon = 1\,, \tag{S.12}$$

and R and S are positive, as well as $\frac{1-\varepsilon}{\varepsilon} \geq 0$, hence $\hat{\sigma} \geq 0$.

(c) Consider the two states of a qutrit system

$$\rho = \begin{pmatrix} 1-\varepsilon & 0 & 0 \\ 0 & \varepsilon & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \text{and} \qquad \sigma = \begin{pmatrix} 1-\varepsilon & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \varepsilon \end{pmatrix}. \tag{S.13}$$

By construction they commute, $[\rho, \sigma] = 0$, which we chose to have a classical example. Taking the above construction for ω we obtain

$$\omega = \begin{pmatrix} (1-\varepsilon)^2 & 0 & 0 \\ 0 & (1-\varepsilon)\varepsilon & 0 \\ 0 & 0 & \varepsilon \end{pmatrix}. \tag{S.14}$$

We have observed in (a) that the problem occurring when trying to translate the classical interpretation to the quantum setting lies in the definition of the event $\{X \neq X'\}$. In this special case it is possible to give this event a natural meaning, which allows us to establish a correspondence between the two interpretation of trace distance.

Think of a measurement of in the standard basis[1] $\{e_0, e_1, e_2\}$ giving an outcome in $\{0, 1, 2\}$. Not knowing which state was actually measured, the event $\{2\}$ then excludes that ρ is the state at hand, while the event $\{1\}$ excludes σ. Taking the coarse-grained event $E = \{1 \textit{ or } 2\}$ then defines an event which allows us to exclude one of the two possible states. But there is another way of interpreting this event: one could say that E characterizes the cases when ρ and σ yield different outcomes. In this sense, E can be seen as the equivalent of $\{X \neq X'\}$ in the classical case.

If one carried out this measurement on ω the probabilities of the respective outcomes would be

$$P_\omega[\{0\}] = (1-\varepsilon)^2 \approx 1 - 2\varepsilon\,, \quad P_\omega[\{1\}] = (1-\varepsilon)\varepsilon \approx \varepsilon\,, \quad P_\omega[\{2\}] = \varepsilon\,, \tag{S.15}$$

where we assumed that ε is small (which is when this interpretation becomes interesting). Hence, the probability to notice a difference between ρ and σ (i.e. the probability to witness event E) is

$$P_\omega[E] \approx 2\varepsilon\,. \tag{S.16}$$

Summing up, we can say that the event E together with P_ω yield a connection of the classical interpretation to the quantum one encountered here by means of $E \leftrightarrow \{X \neq X'\}$ and $P_\omega \leftrightarrow \bar{P}$. The factor of 2 in (S.16) is not relevant when ε is sufficiently small.

[1] Notice that the standard basis is also the common eigenbasis of ρ and σ. If such an eigenbasis did not exist, i.e. if $[\rho, \sigma] \neq 0$, the argument below would not go through.