ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Quantum Information Theory
Tips 5

HS 11
Lídia and Norm

Exercise 5.1 Purification

Purification is explained in detail in the script (pages 32–33). In a nutshell, for every mixed state $\rho_A \in \mathcal{H}_A$, it is possible to find a pure state $|\psi\rangle$ in a larger system $\mathcal{H}_A \otimes \mathcal{H}_B$ such that when we trace out the purification system $\mathcal{H}_B$ we recover the original state: $\rho_A = \mathrm{Tr}_B|\psi\rangle\langle\psi|$. Note the existence of a purifying space $\mathcal{H}_B$ does not tell us where to find it, or even if it represents any physical system!
In part *a*) we are given a formula for purification and have to check that it actually works. First step: any density operator may be expanded in its eigenbasis (spectral decomposition, pages 26–27 of the script) as

$$\rho_A = \sum_x \lambda_x |a_x\rangle\langle a_x|.$$

We expand the operator like that and then build a pure state in $\mathcal{H}_A \otimes \mathcal{H}_B$ as

$$|\psi\rangle = \sum_x \sqrt{\lambda_x}|a_x\rangle_A \otimes |b_x\rangle_B,$$

where $\{b_x\}_x$ forms an orthonormal basis of $\mathcal{H}_B$. Note that this implies that the dimension of the purification space $\mathcal{H}_B$ is the same as the dimension of the original space $\mathcal{H}_A$. Here you only have to check that $|\psi\rangle$ is indeed a purification of ρ_A, ie. that $\rho_A = \mathrm{Tr}_B|\psi\rangle\langle\psi|$.
In part *b*) you have to prove that any two purifications are equivalent up to an isometry on the purifying system. For instance, both

$$|\psi\rangle_{AB} = \frac{1}{\sqrt{2}}\Big(|0\rangle_A|+\rangle_B + |1\rangle_A|-\rangle_B\Big) \quad \text{and} \quad |\phi\rangle_{AC} = \frac{1}{\sqrt{2}}\Big(|0\rangle_A|000\rangle_C + |1\rangle_A|110\rangle_C\Big)$$

purify the fully mixed state of A, $\rho_A = \mathbb{1}_A/2$, even if B is a 1-qubit system while C is a 3-qubit system, and even if the state is purified in different bases. But alas, these two states are related by an isometry on the purifying system: $|\phi\rangle_{AC} = (\mathbb{1}_A \otimes U_{B\mapsto C})|\psi\rangle_{AB}$, where $U_{B\mapsto C}$ maps $|+\rangle_B$ to $|000\rangle_C$ and $|-\rangle_B$ to $|110\rangle_C$. To prove that this is always the case, write a Schmidt decomposition of the purified state, and see which elements can be changed so that the state still purifies ρ_A.
Let us now approach part *c*) of the exercise. Suppose you want to create a certain mixed state ρ'. It is relatively easy to create pure states because you know exactly what the state should be like—things like a spin up or a bunch of photons with a certain polarisation—but mixed states are more tricky as they are states about which we do not have full information, i.e. we are not sure about their exact states. One way to do it is to diagonalise $\rho' = \sum_z \alpha_z |z\rangle\langle z|$ and then get a machine that produces the pure state $|z\rangle$ with probability α_z. Of course you need to be sure that the machine is genuinely random and that you have no access to the information "which state has been created". The first condition in particular is hard to achieve classically, and one may think that there has to be a neater way to do it, maybe using quantum mechanics. Well, there is, and it involves purification.
The idea is that you prepare a special pure state in a bipartite system (easy) and then perform a measurement in a part of that system (also easy). What is left is a mixed state in the other subsystem, and you know which mixed state it is according to the result of your measurement. What we will now see is what pure state and measurements we should prepare for a desired set of mixed states $\{\rho_A^x\}_x$.
The first step is to write down a mixed state that is a convex combination of $\{\rho_A^x\}_x$, $\rho_A = \sum_x \lambda_x \rho_A^x$ (ie. the state ρ_A is decomposed in the possibly mixed states $\{\rho_A^x\}_x$). We will purify ρ_A, but instead of first mixing the components $\{\rho_A^x\}_x$ and then purify the resulting state ρ_A, we will do it the other way around, purifying the components, then mixing the resulting pure states and finally purifying the state we get from there. We will see that the two processes are equivalent.
We can decompose every $\rho_A^x \in \mathcal{H}_A$ in its eigenbasis $\big\{|a_y^x\rangle\big\}_y$ as $\rho_A^x = \sum_y \alpha_y^x |a_y^x\rangle\langle a_y^x|$ and then purify it as usual using the extra Hilbert space $\mathcal{H}_C$,

$$|\phi^x\rangle = \sum_y \sqrt{\alpha_y^x}|a_y^x\rangle_A \otimes |c_y^x\rangle_C.$$

We have now a set of pure states $\{|\phi^x\rangle\}_x$, and we will combine them to make a mixed state using the coefficients $\{\lambda_x\}_x$ used to make ρ_A. The state we will obtain belongs to the composed space $\mathcal{H}_A \otimes \mathcal{H}_C$,

$$\rho_{AC} = \sum_x \lambda_x |\phi^x\rangle\langle\phi^x|.$$

Now we purify this state using a purifying system $\mathcal{H}_D$,

$$|\phi\rangle = \sum_x \sqrt{\lambda_x} |\phi^x\rangle_{AC} \otimes |d_x\rangle_D.$$

This state lives in $\mathcal{H}_A \otimes \mathcal{H}_C \otimes \mathcal{H}_D$ and you will see that $|\phi\rangle$ is a purification of ρ_A in system $\mathcal{H}_B = \mathcal{H}_C \otimes \mathcal{H}_D$. For this you have to check that $\mathrm{Tr}_B(|\phi\rangle\langle\phi|) = \rho_A$.
So now we have a pure state that is a purification of ρ_A. The next step is to choose a measurement in the purifying system $\mathcal{H}_B$ such that after measuring and tracing $\mathcal{H}_B$ out we obtain a mixed state from the set $\{\rho_A^x\}_x$, as we wanted. Remember what we saw last week about how measurements can be represented by a set of operators out of which one is chosen at random? You may want to test the following measurements:

$$M_B = \{M_B^x = \mathbb{1}_C \otimes |d_x\rangle\langle d_x|_D\}_x .$$

In particular, check that $\rho_A^x = \frac{\mathrm{Tr}_B[|\Phi\rangle\langle\Phi|(\mathbb{1}_A \otimes M_B^x)]}{\lambda_x}$ and $\lambda_x = \mathrm{Tr}\left[|\Phi\rangle\langle\Phi|(\mathbb{1}_A \otimes M_B^x)\right]$.
Summary of the mixed state recipe: you have a set of mixed states $\{\rho_A^x\}_x$ you would like to create (or alternatively you have one mixed state and you would like to obtain a certain decomposition). Mix the states in a convex combination ρ_A with coefficients $\{\lambda_x\}_x$. Go back to your initial set of states. Purify them. Now mix the purified states just like you did with the mixed ones before. Purify that global mixed state. Now apply a measurement in the purifying systems that measures the element of basis of this last purifying system and acts as the identity in the first one. Trace out both the purifying systems (the state had collapsed there anyway). Now with probability λ_x you measured $|d_x\rangle$ in the last purifying system and are left with ρ_A^x. You always know which state you have because you know the outcome of your measurement. Simple, right?

Exercise 5.2 Distinguishing two quantum states

This exercise is similar to the one where we knew the behaviour of two dice and after throwing one of them we had to guess which one that was. The difference now is that we are dealing with quantum states with density matrices instead of classical probability distributions. Back then our strategy was to choose the die that was more likely to outcome the event we measured. Here we will do the same. However, we have to decide in which basis to perform the measurement. In the case of two density matrices that share the same eigenstates, this is easy – we can simply perform a measurement in that basis. Take the example of two states given with equal probability,

$$\rho = 0.7\ |0\rangle\langle 0| + 0.3\ |1\rangle\langle 1|, \qquad \sigma = 0.1\ |0\rangle\langle 0| + 0.9\ |1\rangle\langle 1|.$$

The probability of measuring eg. $|0\rangle$ when looking at state ρ is given by $\mathrm{Tr}(|0\rangle\langle 0|\ \rho)$. We could perform a measurement in their common eigenbasis $\{|0\rangle, |1\rangle\}$. If we obtained 0 we should guess we had measured state ρ and vice-versa. Things get more complicated when the two states are not diagonalised in the same basis. Consider for instance

$$\rho = 0.7\ |0\rangle\langle 0| + 0.3\ |1\rangle\langle 1|, \qquad \sigma = \frac{\mathbb{1}}{2}, \quad \tau = 0.7\ |+\rangle\langle +| + 0.3.\ |-\rangle\langle -|$$

Let us see what happens if we perform a measurement in three different bases. The probabilities of obtaining the different outcomes for these states are
Even though $\{|+\rangle, |-\rangle\}$ is an eigenbasis of σ, a measurement in this basis would not help distinguish σ from ρ. A measurement in basis $\{|\circlearrowleft\rangle, |\circlearrowright\rangle\}$ would be even worse: all three states have the same measurement statistics in this basis. And although the statistics for σ and ρ are different in the computational basis, one may wonder if there is a better choice of basis where they're even more apart.

outcome:	$\|0\rangle$	$\|1\rangle$	$\|+\rangle$	$\|-\rangle$	$\|\circlearrowright\rangle$	$\|\circlearrowleft\rangle$
ρ	0.7	0.3	0.5	0.5	0.5	0.5
σ	0.5	0.5	0.5	0.5	0.5	0.5
τ	0.5	0.5	0.7	0.3	0.5	0.5

In general, we are looking for the measurement basis $\{|y\rangle\}_y$ that maximises our probability of sucessfully distinguishing the two states. Depending on the outcome of the measurement, we will guess that we had either ρ or σ. So we can define a set of all the measurement outcomes that will make us choose ρ: call this set G, so that $\{y\}_y = G \cup \bar{G}$.
Recall that the probability of obtaining outcome y when we measure ρ is $\mathrm{Tr}(|y\rangle\langle y|\rho)$. Remember also that, in this exercise, ρ was given to us with probability p, and σ with probability $1-p$. Our probability of guessing correctly when we measure in basis $\{|y\rangle\}_y$ is

$$\begin{aligned}
\Pr^p_{\checkmark}(\rho,\sigma) &= \Pr(\text{state was } \rho \text{ and we guess } \rho) + \Pr(\text{state was } \sigma \text{ and we guess } \sigma) \\
&= p\ \Pr(\text{we get one of the outcomes in } G \mid \text{we measure } \rho) \\
&\quad + (1-p)\ \Pr(\text{we get one of the outcomes in } \bar{G} \mid \text{we measure } \sigma) \\
&= p\ \Big(\sum_{y\in G} \Pr(\text{we get outcome } y \mid \text{we measure } \rho)\Big) \\
&\quad + (1-p)\ \Big(\sum_{y\in \bar{G}} \Pr(\text{we get outcome } y \mid \text{we measure } \rho)\Big) \\
&= p\ \Big(\sum_{y\in G} \mathrm{Tr}(|y\rangle\langle y|\rho)\Big) + (1-p)\ \Big(\sum_{y\in \bar{G}} \mathrm{Tr}(|y\rangle\langle y|\sigma)\Big).
\end{aligned}$$

Show, using properties of the trace, that

$$\Pr^p_{\checkmark}(\rho,\sigma) = 1 - p + \mathrm{Tr}\Big(P_G[p\ \rho - (1-p)\ \sigma]\Big), \qquad P_G = \sum_{y\in G} |y\rangle\langle y|.$$

So now we want to choose the projector P_G that maximises this quantity. For simplicity, call $A = p\ \rho - (1-p)\ \sigma$. This is an operator (matrix) with positive and negative eigenvalues; for instance

$$A = \begin{pmatrix} 0.8 & 0 & 0 & 0 \\ 0 & -0.7 & 0 & 0 \\ 0 & 0 & 0.2 & 0 \\ 0 & 0 & 0 & -0.3 \end{pmatrix}$$

when diagonalised. The projector P_G is a matrix that only has eigenvalues 0 and 1, e.g.,

$$P_G = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Show that, in order to maximise $\mathrm{Tr}(P_G A)$, we should choose P_G to be the projector onto the subspace that corresponds to *positive eigenvalues* of A (like in the example above).
This solution corresponds to the following strategy. We measure our state (ρ or σ) in the eigenbasis of $A = p\ \rho - (1-p)\ \sigma$. If we the outcome corresponds to a positive eigenvalue of A (i.e. $y \in G$), then it is more likely that we have measured ρ. If it corresponds to a negative eigenvalue of A (i.e. $y \in \bar{G}$), we should say that the state was σ.
In the particular case where the two density operators share the same eigenbasis this corresponds to following the classical strategy after measuring the state in their common eigenbasis.
Also, when ρ and σ are given with equal probability, we recover the result $\Pr^{0.5}_{\checkmark}(\rho,\sigma) = \frac{1}{2}[1+\delta(\rho,\sigma)]$, where $\delta(\rho,\sigma)$ is the trace distance between the two operators (defined in the script).

Exercise 5.3 Data hiding

Nah, there's enough information in the exercise sheet. This result means that to encode one bit of information with security ϵ (i.e., such that the agents' probability of successfully guessing the bit is just $1/2 + \epsilon$), you need a global system of approximately $4/\epsilon$ qubits (and give half the qubits to each agent). For instance, if you want $\epsilon = 1\%$, you need 400 qubits!