Exercise Sheet 5: Entropies and Examples of Quantum Channels

Entropic Quantities

Properties of the Shannon entropy

The Shannon entropy and related entropic quantities are of fundamental importance in classical information theory. We now invest some time and effort to better understand these notions.

8 P. **Exercise 1.** Recall the definition of the Shannon entropies for random variables X, Y which take values in (for simplicity finite) sets $\mathcal{X}, \mathcal{Y}$, and are distributed according to probability distributions p, q over $\mathcal{X}$ and $\mathcal{Y}$, respectively:

$$\text{Shannon entropy:} \qquad H(X) = -\sum_{x\in\mathcal{X}} p(x)\log p(x) \tag{1}$$

$$\text{Conditional entropy:} \qquad H(X|Y) = H(X,Y) - H(Y) = \sum_{y\in\mathcal{Y}} p(y) H(X|Y=y) = -\sum_{x\in\mathcal{X}, y\in\mathcal{Y}} p(x,y)\log\frac{p(x,y)}{p(y)} \tag{2}$$

Note that here, as is common in information theory, log denotes the logarithm with base 2. For these definitions, recall that we take the convention $0\log(0) \equiv 0$.

2 P. (a) Show that $0 \leq H(X) \leq \log|\mathcal{X}|$. Moreover, show that the first equality holds if there is an $x \in \mathcal{X}$ for which $p(x) = 1$, and that the second equality holds if $p(x) = 1/|\mathcal{X}|$ for all x.

Hint: To prove the second inequality, use Jensen's inequality for a suitably chosen concave function.

Solution

First inequality. Since $p(x) \leq 1$ for all x, we have $\log(p(x)) \leq 0$ for all x, which together with $p(x) \geq 0$ implies that $0 \leq H(X)$. If there exist $x \in \mathcal{X}$ with $p(x) = 1$, we have equality, since $\log(1) = 0$ and $0 \cdot \log(0) = 0$. (This can be justified via $\lim_{x \searrow 0} x \log(x) = 0$.) Bonus note: The reverse direction also holds. This follows from the facts that $H(X)$ is concave and that the set of probability distributions is convex, so the entropy attains its minimum value exactly at the extreme points. So, the entropy attains its minimum value at the extreme points, which are exactly the "delta distributions" that have $p(x) = 1$ for exactly one $x \in \mathcal{X}$.

Second inequality. Using Jensen's inequality for the convex function $x \mapsto \log(x)$, we get

$$H(X) = \mathbb{E}\left[\log_2 \frac{1}{p(x)}\right] \leq \log_2 \mathbb{E}\left[\frac{1}{p(x)}\right] = \log_2 \sum_{x \in \mathcal{X}} \frac{p(x)}{p(x)} = \log_2 |\mathcal{X}|.$$

This proves the upper bound. And it is an easy computation to see that this upper bound is attained for the uniform distribution.

Second inequality – Alternative solution. The second inequality can also be proven using Lagrange multipliers. In particular, if all $p_x := p(x) > 0$, we can compute the gradient $(\mathrm{grad} H(X))_{p_x} = -\log(p_x) - 1$. Together with the restriction $\sum_x p_x = 1$, we obtain the equations $-\log(p_x) - 1 + \lambda = 0$. As log is an injective function, this can only be the case if all p_x are equal. Then they all have to be equal to $1/|\mathcal{X}|$, so that evaluating $H(X)$ at $p_x = 1/|\mathcal{X}|$ yields the second inequality. In fact, the uniform distribution over $\mathcal{X}$ is the unique maximizer of the entropy. This can be seen as follows: It can easily be checked that the case with $p_x = 0$ for some multiple $x \in X$ does not yield a larger value simply by repeating the above argument on the non-vanishing p_x.

3 P. (b) Show that the Shannon entropy is *subadditive*, i.e., that $H(X,Y) \leq H(X) + H(Y)$ with equality if X and Y are independent.

Hint: Show that $H(X,Y) - H(X) - H(Y) \leq 0$ *using that* $\ln(2) \cdot \log_2(x) = \ln(x) \leq x - 1$.

Solution

Using that $\sum_x p(x,y) = p(y)$ and $\sum_y p(x,y) = p(x)$, we have that

$$\begin{aligned} H(X,Y) - H(X) - H(Y) &= -\sum_{x,y} p(x,y)(\log p(x,y) - \log p(x) - \log p(y)) \\ &= \sum_{x,y} p(x,y) \log \frac{p(x)p(y)}{p(x,y)} \leq \frac{1}{\ln 2} \sum_{x,y} p(x,y) \left(\frac{p(x)p(y)}{p(x,y)} - 1\right) \\ &= \frac{1}{\ln 2} \sum_{x,y} (p(x)p(y) - p(x,y)) = \frac{1}{\ln 2}(1-1) = 0 \end{aligned}$$

We also immediately see that equality holds if $p(x,y) = p(x)p(y)$ for all x, y, i.e. for independent random variables X and Y, because then $p(x,y)(\log p(x,y) - \log p(x) - \log p(y)) = 0$ for all x, y.

2 P. (c) Show that $H(Y|X) \geq 0$, with equality if Y is a (deterministic) function of X.

Hint: Use Bayes' rule: $p(x,y) = p(y|x)p(x)$.

Solution

Non-negativity of $H(Y|X)$ can be seen as follows:

$$\begin{aligned} H(Y|X) &= -\sum_{x\in\mathcal{X},y\in\mathcal{Y}} p(x,y)\log_2\frac{p(x,y)}{p(x)} \\ &= -\sum_{x\in\mathcal{X},y\in\mathcal{Y}} p(x,y)\log_2\frac{p(y|x)p(x)}{p(x)} = -\sum_{x\in\mathcal{X},y\in\mathcal{Y}} p(x,y)\log_2 p(y|x)\,. \end{aligned}$$

Now, we have $p(x,y) \geq 0$ and $\log_2 p(y|x) \leq 0$ (because $0 \leq p(y|x) \leq 1$), so the above rewriting of $H(Y|X)$ is ≥ 0.

Next, suppose $Y = f(X)$ for some deterministic function $f : \mathcal{X} \to \mathcal{Y}$. Then, $p(y|x) = \delta_{f(x),y}$. Rewriting

$$H(Y|X) = -\sum_{x\in\mathcal{X}} p(x)\sum_{y\in\mathcal{Y}} p(y|x)\log_2 p(y|x)\,,$$

we can now again use the convention $0\log(0) \equiv 0$. This gives us

$$\begin{aligned} H(Y|X) &= -\sum_{x\in\mathcal{X}} p(x)\sum_{y\in\mathcal{Y}} \delta_{f(x),y}\log_2\delta_{f(x),y} \\ &= -\sum_{x\in\mathcal{X}} p(x)\left((|\mathcal{Y}|-1)0\log(0) + 1\log(1)\right) \\ &= 0\,. \end{aligned}$$

1 P. (d) Show that $H(Y|X) \leq H(Y)$ with equality if X and Y are independent random variables.

Solution

Using subadditivity, we have

$$H(Y|X) = H(X,Y) - H(X) \leq H(X) + H(Y) - H(X) = H(Y).$$

Equality holds if $H(X,Y) = H(X) + H(Y)$, which is the case if $p(x,y) = p(x)p(y)$ as we have seen when proving subadditivity.

Properties of the von Neumann entropy

Above, we have investigated some useful properties of the Shannon entropy. Now, we do the same for the von Neumann entropy. For any state $\rho \in \mathcal{D}(\mathcal{H})$ with $\dim\mathcal{H} = d$, the von Neumann entropy is defined as $S(\rho) = -\operatorname{Tr}[\rho\log\rho]$. Recall from previous exercise sheets that $\log\rho$ is obtained by taking the eigendecomposition $\rho = UDU^\dagger$, with $D = \operatorname{diag}((\lambda_i)_i)$ and applying the log function to the (non-zero) eigenvalues, leaving everything else untouched $\log\rho = U\operatorname{diag}((\log\lambda_i)_i)V^\dagger$.

6 P. Exercise 2.

1 P. (a) Show that $0 \leq S(\rho) \leq \log d$ with equality if and only if ρ is pure.

Hint: You've already proved something quite similar for the Shannon entropy in Exercise 1. So, don't start from scratch but instead try to use that already established result.

Solution

Diagonalize ρ to obtain $S(\rho) = -\sum_x \lambda_x \log \lambda_x$, where the λ_x are the eigenvalues of ρ. The λ_x form a probability distribution and $S(\rho) = H((\lambda_x)_x)$. The claim now follows from the lower and upper bounds on the Shannon entropy (proven in Exercise 1).

2 P. (b) Prove that for any two positive definite matrices A and B we have $\log(A \otimes B) = \log(A) \otimes \mathbb{I} + \mathbb{I} \otimes \log(B)$.

Note: We only ask you to prove for $A, B > 0$, but this result can be extended to $A, B \geq 0$ given a proper definition of the matrix logarithm.

Solution

We first address the case $A, B > 0$. Let A and B have spectral decompositions

$$A = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i|, \quad B = \sum_j \mu_j |\phi_j\rangle\langle\phi_j|\,,$$

where the eigenvalues by the positive definiteness assumption satisfy $\lambda_i, \mu_j > 0$ for all i, j. In particular, $\log(\lambda_i \mu_j)$, $\log(\lambda_i)$, and $\log(\mu_j)$ are well-defined for all i and j. Then, the spectral decomposition of $A \otimes B$ is

$$A \otimes B = \sum_{ij} \lambda_i \mu_j |\psi_i\rangle\langle\psi_i| \otimes |\phi_j\rangle\langle\phi_j|\,,$$

so we can compute

$$\begin{aligned}
&\log(A \otimes B) \\
&= \sum_{ij} \log(\lambda_i \mu_j) |\psi_i\rangle\langle\psi_i| \otimes |\phi_j\rangle\langle\phi_j| \\
&= \sum_{ij} (\log(\lambda_i) + \log(\mu_j))\, |\psi_i\rangle\langle\psi_i| \otimes |\phi_j\rangle\langle\phi_j| \\
&= \left(\sum_i \log(\lambda_i) |\psi_i\rangle\langle\psi_i|\right) \otimes \left(\sum_j |\phi_j\rangle\langle\phi_j|\right) + \left(\sum_i |\psi_i\rangle\langle\psi_i|\right) \otimes \left(\sum_j \log(\mu_j) |\phi_j\rangle\langle\phi_j|\right) \\
&= \log(A) \otimes \mathbb{I} + \mathbb{I} \otimes \log(B)\,,
\end{aligned}$$

where the last step used that $\sum_i |\psi_i\rangle\langle\psi_i| = \mathbb{I}$ and $\sum_j |\phi_j\rangle\langle\phi_j| = \mathbb{I}$.

Note on positive semidefinite matrices. Now for the case $A, B \geq 0$. If A and B have 0 as an eigenvalue, then so does $A \otimes B$. However, by definition of the matrix logarithm, we apply the logarithm only to the non-zero eigenvalues occurring in the spectral decompositions and leave the 0-eigenvalue untouched (since the corresponding summands in the spectral decomposition anyways vanish). Hence, we don't run into any issues about the logarithm not being well-defined. (This can be made more formal when written out in terms of eigenprojections.)

3 P. (c) Show that the von Neumann entropy is *subadditive* in the following sense: If two distinct systems A and B have a joint quantum state ρ_{AB} then $S(A, B) \leq S(A) + S(B)$, with equality if $\rho_{AB} = \rho_A \otimes \rho_B$.

Hint: You may use the inequality $S(\rho) \leq -\operatorname{Tr}[\rho \log \sigma]$ for an arbitrary quantum state σ (without proving it).

Solution

We want to show

$$S(A,B) = S(\rho_{AB}) \leq S(\mathrm{Tr}_B[\rho_{AB}]) + S(\mathrm{Tr}_A[\rho_{AB}]) = S(A) + S(B).$$

Let $\rho = \rho_{AB}$, $\sigma = \rho_A \otimes \rho_B$. Then:

$$\begin{aligned} S(\rho) &\leq -\mathrm{Tr}[\rho \log(\rho_A \otimes \rho_B)] \\ &= -\mathrm{Tr}[\rho(\log(\rho_A) \otimes \mathbb{I} + \mathbb{I} \otimes \log(\rho_B))] \\ &= -\mathrm{Tr}[\rho(\log(\rho_A) \otimes \mathbb{I})] - \mathrm{Tr}[\rho(\mathbb{I} \otimes \log(\rho_B))] \\ &= S(\rho_A) + S(\rho_B). \end{aligned}$$

Here, the first step used the inequality in the hint for $\sigma = \rho_A \otimes \rho_B$, the second step used (b), and the third step used that the partial trace satisfies $\mathrm{Tr}[\rho(X_A \otimes \mathbb{I})] = \mathrm{Tr}[\rho_A X_A]$ for every operator X_A on the A-system (and similarly for the B-system). This proves the desired inequality.

Now, for the claimed equality: Suppose $\rho_{AB} = \rho_A \otimes \rho_B$. Then, the first step in our derivation of the inequality above is clearly an equality. As the other steps are equalities anyways, this shows that we get the overall equality $S(\rho) = S(\rho_A) + S(\rho_B)$ as claimed.

4 P. Bonus Exercise 1.

3 P. (a) Suppose that $p = (p_i)_i$ is a discrete probability distribution and that the states ρ_i are mutually orthogonal. Show that

$$S\left(\sum_i p_i \rho_i\right) = H(p) + \sum_i p_i S(\rho_i).$$

and use this result to infer that

$$S\left(\sum_i p_i \rho_i \otimes |\psi_i\rangle\langle\psi_i|\right) = H(p) + \sum_i p_i S(\rho_i),$$

where $\langle\psi_i|\psi_j\rangle = \delta_{ij}$ and the ρ_i are arbitrary quantum states.

Solution

Take spectral decompositions

$$\rho_i = \sum_j \lambda_i^j \Pi_i^j \,,$$

where the Π_i^j are eigenprojectors. By assumption, the states ρ_i are mutually orthogonal, hence we have $\mathrm{Tr}[\Pi_i^j \Pi_k^l] = 0$ for $i \neq k$ whenever $\max\{\lambda_i^j, \lambda_k^l\} > 0$. Also, we have $\mathrm{Tr}[\Pi_i^j \Pi_i^l] = 0$ whenever $j \neq l$. Thus, we can compute:

$$\begin{aligned}
S\left(\sum_i p_i \rho_i\right) &= -\,\mathrm{Tr}\left[\sum_{i,j} p_i \lambda_i^j \Pi_i^j \log\left(\sum_{k,l} p_k \lambda_k^l \Pi_k^l\right)\right] \\
&= -\,\mathrm{Tr}\left[\sum_{i,j} p_i \lambda_i^j \Pi_i^j \left(\sum_{k,l:\lambda_k^l \neq 0} (\log p_k + \log \lambda_k^l) \Pi_k^l\right)\right] \\
&= -\left(\sum_{i,j} \sum_{k,l:\lambda_k^l \neq 0} p_i \lambda_i^j (\log p_k + \log \lambda_k^l)\, \mathrm{Tr}\left[\Pi_i^j \Pi_k^l\right]\right) \\
&= -\left(\sum_{i,j} p_i \lambda_i^j (\log p_i + \log \lambda_i^j)\right) \\
&= H(p) + \sum_i p_i S(\rho_i)\,.
\end{aligned}$$

Here, the second-to-last step used the orthogonality relations among the eigenprojections discussed above and the last step used that $\sum_j \lambda_i^j = 1$ for every i.
To get the second of the two claimed equalities, first observe that the states $\rho_i \otimes |\psi_i\rangle\langle\psi_i|$ are mutually orthogonal, simply because the $|\psi_i\rangle\langle\psi_i|$ are: $\mathrm{Tr}[(\rho_i \otimes |\psi_i\rangle\langle\psi_i|)^\dagger(\rho_j \otimes |\psi_j\rangle\langle\psi_j|)] = \mathrm{Tr}[\rho_i^\dagger \rho_j] \cdot |\langle\psi_i|\psi_j\rangle|^2 = \mathrm{Tr}[\rho_i^\dagger \rho_j]\delta_{ij}$. So, we can apply the equality that we have just established. If we additionally use the equality in the subadditivity condition for $S(\rho \otimes \sigma) = S(\rho) + S(\sigma)$, this gives $S(\rho_i \otimes |\psi_i\rangle\langle\psi_i|) = S(\rho_i) + S(|\psi_i\rangle\langle\psi_i|) = S(\rho_i)$ and thus the claimed equality.

1 P. (b) Use the subadditivity and the result from (a) to infer that the von Neumann entropy S is concave, that is, $S(\sum_i p_i \rho_i) \geq \sum_i p_i S(\rho_i)$ for any probability distribution $\{p_i\}$ and states ρ_i.

Solution

To show: $S(\sum_i p_i \rho_i) \geq \sum_i p_i S(\rho_i)$.
We start from the rhs:

$$\begin{aligned}
\sum_i p_i S(\rho_i) &= S\left(\sum_i p_i \rho_i \otimes |i\rangle\langle i|\right) - H(p) \\
&\leq S\left(\sum_i p_i \rho_i\right) + S\left(\sum_i p_i |i\rangle\langle i|\right) - H(p) = S\left(\sum_i p_i \rho_i\right)
\end{aligned}$$

Here, the first step used (a), the second step used that $S(A, B) \leq S(A) + S(B)$, and the final step used $S(\sum_i p_i |i\rangle\langle i|) = H(p)$.

5 P. **Exercise 3.** Now we want to set our sight back on bipartite systems and entanglement entropies. Throughout this problem, if the global state being referred to is clear, we will denote entropies of the reduced states using the corresponding Hilbert space as an argument, e.g. the entropy of a state ρ_{AB} reduced on subsystem A, that is $S\left(\mathrm{Tr}_B\left[\rho_{AB}\right]\right)$, is denoted $S(A)$.

1 P. (a) Let $|\psi\rangle_{AB} = |\xi\rangle_A \otimes |\eta\rangle_B$ be some product pure state. Compute the entanglement entropy of the state, as well as the conditional von Neumann entropy $S(A|B) \equiv S(A,B) - S(B)$.

Solution

$$E(\rho) = S(B) = -\operatorname{Tr}\rho_B \log_2 \rho_B$$

For a product state, the reduced state is also a pure state $\rho_B = |\eta\rangle\langle\eta|$, thus its von Neumann entropy is 0. Consequently, the entanglement entropy of a pure product state also is 0.

$$S(A|B) = S(A,B) - S(B) = -\operatorname{Tr}[\rho_{AB}\log_2(\rho_{AB})] + \operatorname{Tr}[\rho_B \log_2(\rho_B)] = 0 + 0$$

since both the bipartite state $\rho_{AB} = |\psi\rangle\langle\psi|$ and the reduced state are pure states, their von Neumann entropy is 0.

1 P. (b) Let Ω_{AB} be the maximally entangled state on two Hilbert spaces of equal dimension d, i.e. $\Omega = |\Omega\rangle\langle\Omega|$ with

$$|\Omega\rangle = \frac{1}{\sqrt{d}}\sum_{i=1}^{d}|ii\rangle.$$

Compute the conditional von Neumann entropy $S(A|B) \equiv S(A,B) - S(B)$. What do you conclude when comparing your result to that of Exercise 1(c)?

Solution

$$S(A|B) = S(AB) - S(B) = -S(B) = -\log(d) = -E(\rho).$$

The conditional von Neumann entropy can be negative for entangled states. This is in contrast to the classical conditional Shannon entropy, which is always non-negative.

2 P. (c) Let ρ_{AB} be a bipartite (potentially mixed) state. Show that if ρ_A is separable, i.e. $\rho_{AB} = \sum_i p_i \sigma_A^i \otimes \tau_B^i$ where σ^i and τ^i are states and $\{p_i\}$ is a probability distribution, then $S(A|B) \geq 0$.

Hint: The concavity of the von Neumann entropy established in (e) can be helpful here.

Solution

The reduced density matrices of ρ_{AB} are

$$\rho_A = \sum_i p_i \sigma_A^i, \quad \rho_B = \sum_i p_i \tau_B^i \,.$$

Thus, $S(B) = S(\sum_i p_i \tau_B^i)$. Let's write the spectral decomposition $\sigma_A^i = \sum_j \lambda_i^j |\psi_i^j\rangle\langle\psi_i^j|$. Now, we have

$$\begin{aligned}
S(AB) &= S(\sum_i p_i \sigma_A^i \otimes \tau_B^i) \\
&= S(\sum_{ij} p_i \lambda_i^j |\psi_i^j\rangle\langle\psi_i^j| \otimes \tau_B^i) \\
&\geq \sum_j \lambda_i^j S(\sum_i |\psi_i^j\rangle\langle\psi_i^j| \otimes \tau_B^i) \\
&= \sum_j \lambda_i^j S(\sum_i \tau_B^i) \\
&= S(B),
\end{aligned}$$

where the inequality is an application of concavity of the von Neumann entropy (applicable since each $(\lambda_i^j)_j$ forms a probability distribution) and the second-to-last step uses the equality case of subadditivity and the fact that pure states have vanishing entropy (just as we did in the solution to (d) above). Thus $S(A|B) \equiv S(AB) - S(B) \geq 0$.

1 P. (d) In the lecture, you have encountered the entropy of entanglement, defined via the entropy of a reduced density matrix, as a measure for pure state entanglement. Show via an example that the entropy of a reduced density matrix is not suitable for quantifying entanglement in mixed states.

Solution

Take a bipartite maximally mixed state. This is a tensor product state, thus not at all entangled, but it achieves the maximum possible value for the entropy of a reduced density matrix. Hence, while large entropy of a reduced density matrix is indicates the presence of entanglement if we started from a pure bipartite state, this is no longer true when starting from a bipartite mixed state.

Examples of Quantum Channels

On the last sheet, we looked at quantum channels on a general, abstract level. Here, we investigate some examples of quantum channels acting on qubits, i.e., our Hilbert space is $\mathcal{H} = \mathbb{C}^2$. The following maps are important so-called noise channels

$$\begin{aligned}
F_\epsilon(\rho) &:= \epsilon X \rho X + (1-\epsilon)\rho) \\
D_\epsilon(\rho) &:= \epsilon \operatorname{Tr}[\rho] \frac{\mathbb{I}}{2} + (1-\epsilon)\rho \\
A_\epsilon(\rho) &:= \epsilon \operatorname{Tr}[\rho] |0\rangle\langle 0| + (1-\epsilon)\rho,
\end{aligned}$$

where $\epsilon \in [0,1]$.

10 P. **Exercise 4.** In this exercise, we will familiarize ourselves with the three channels above. In the process, we will also learn to use the three important representations of channels that we have encountered so far – Choi-Jamiołkowski, Kraus, and Stinespring – in concrete examples rather than just abstractly.

4 P. (a) Show that F_ϵ, D_ϵ, and A_ϵ indeed are linear CPTP maps and thus valid quantum channels.

Solution

Trace preservation and linearity are obvious. Therefore we just show complete positivity. If the Choi state of a linear superoperator has non-negative eigenvalues the map is completely positive, so we show that the Choi states of these linear maps are PSD.
Because of linearity we can write

$$J(F_\epsilon) = \epsilon J(F_1) + (1-\epsilon)J(\mathrm{id}) = \epsilon J(F_1) + (1-\epsilon)|\Omega\rangle\langle\Omega| \,,$$

and similarly for $J(D_\epsilon)$ and $J(A_\epsilon)$. As $|\Omega\rangle\langle\Omega|$ is PSD, we now focus on $J(F_1)$, $J(D_1)$, and $J(A_1)$.
The Choi state for F_1 is given by

$$J(F_1) = \frac{1}{2}\sum_{i,j} |i \text{ xor } 1, i\rangle\langle j \text{ xor } 1, j| = \frac{1}{2}\sum_{i,j} |i \oplus 1, i\rangle\langle j \oplus 1, j| \,,$$

where 1 xor 1 = 0 and 0 xor 1 = 1 and $\oplus$ is just another way of writing this flip. Written in matrix from, this is

$$J(F_1) = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} ,$$

so $J(F_1)$ has eigenvalues 0 (of multiplicity 3) and 1 (of multiplicity 1) and is thus PSD. Hence, $J(F_\epsilon)$ is a convex combination of two PSD matrices and thus itself PSD.
The depolarising channel has the Choi state

$$\begin{aligned} J(D_1) &= (D_1 \otimes \mathbb{I})|\Omega\rangle\langle\Omega| \\ &= \frac{1}{d^2}\mathbb{I} \otimes \mathrm{Tr}_1\left[\sum_{ij} |ii\rangle\langle jj|\right] \\ &= \frac{1}{d^2}\mathbb{I} \otimes \sum_{ij} \delta_{ij}|i\rangle\langle j| \\ &= \frac{1}{d^2}\mathbb{I} \otimes \sum_{i} |i\rangle\langle i| \\ &= \frac{1}{d^2}\mathbb{I} \otimes \mathbb{I} \geq 0 \,, \end{aligned}$$

So, $J(D_\epsilon)$ is PSD as a convex combination of PSD matrices.
Finally, we have

$$J(A_1) = |0\rangle\langle 0| \otimes \frac{\mathbb{I}}{2} = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \geq 0 \,.$$

Again, $J(A_\epsilon)$ is PSD as a convex combination of PSD matrices.

3 P. (b) Describe the action of each of the channels F_ϵ, D_ϵ, and A_ϵ in words.

Solution

F_ϵ does the following: With probability $1-\epsilon$, it leaves the input state unchanged, and with probability ϵ, it "flips" the input state by applying a Pauli X.

D_ϵ does the following: With probability $1-\epsilon$, it leaves the input state unchanged, and with probability ϵ, it completely depolarizes the input state, i.e., it throws the input state away and replaces it by a maximally mixed state.

A_ϵ does the following: With probability $1-\epsilon$, it leaves the input state unchanged, and with probability ϵ, it throws the input state away and replaces it by $|0\rangle\langle 0|$.

3 P. (c) Compute the action of each of the three channels on the input states $|0\rangle\langle 0|$ and $\rho = \mathbb{I}/2$.

Solution

For the input $|0\rangle\langle 0|$:

$$
\begin{aligned}
F_\epsilon(|0\rangle\langle 0|) &= \epsilon|1\rangle\langle 1| + (1-\epsilon)|0\rangle\langle 0|\,, \\
D_\epsilon(|0\rangle\langle 0|) &= \epsilon\frac{\mathbb{I}}{d} + (1-\epsilon)|0\rangle\langle 0|\,, \\
A_\epsilon(|0\rangle\langle 0|) &= |0\rangle\langle 0|\,.
\end{aligned}
$$

For the input $\rho = \mathbb{I}/2$:

$$
\begin{aligned}
F_\epsilon(\mathbb{I}/2) &= \mathbb{I}/2\,, \\
D_\epsilon(\mathbb{I}/2) &= \mathbb{I}/2\,, \\
A_\epsilon(\mathbb{I}/2) &= \epsilon|0\rangle\langle 0| + (1-\epsilon)\mathbb{I}/2\,.
\end{aligned}
$$

Recap

As we keep gaining in speed and confidence when talking about mixed and entangled states, we take a moment to revisit some core concepts via direct, small examples.

5 P. Bonus Exercise 2.

3 P. (a) Come up with three mixed quantum states as density matrices ρ_1, ρ_2, and ρ_3. Let the states be such that the density matrix is not diagonal in the computational basis, and let each of them have the corresponding purity: $\mathrm{Tr}(\rho_k^2) = k/(k+1)$.

Solution

Let's make it nice and easy. Let σ_1, σ_2 be any two orthogonal states. Our ansatz for the ρ_i will be mixtures of σ_1 and σ_2. That is, we consider states

$$\rho(p) = p\sigma_1 + (1-p)\sigma_2\,,$$

where $p \in [0,1]$. It follows that $\mathrm{Tr}((\rho(p))^2) = p^2 + (1-p)^2 = 2p^2 - 2p + 1$. Now let's equate this to the required purities, and find the value p_k that we need to set p to, for each k:

$$\begin{aligned}
2p_k^2 - 2p_k + 1 &= \frac{k}{k+1} \\
2p_k^2 - 2p_k + 1 - \frac{k}{k+1} &= 0 \\
\implies p_k &= \frac{2 \pm \sqrt{4 - 8(1 - \frac{k}{k+1})}}{4} \\
&= \frac{1 \pm \sqrt{1 - \frac{2}{k+1}}}{2} \\
&= \frac{1 \pm \sqrt{\frac{k-1}{k+1}}}{2} \\
&= \frac{1 \pm \frac{\sqrt{k^2-1}}{k+1}}{2} \\
&= \frac{k+1 \pm \sqrt{k^2-1}}{2(k+1)}.
\end{aligned}$$

Okay, this is as far as we'll go, now you just plug in any $k \geq 1$ and that gives you a corresponding probability p_i.
Since we ask for mixed states that are not diagonal in the computational basis, it's left to fix the arbitrary orthogonal states σ_1 and σ_2 such that the sum is not diagonal on the computational basis. It should be enough to take $\sigma_1 = |+\rangle\langle+|$ and $\sigma_2 = |-\rangle\langle-|$.

2 P. (b) Compute the purity of the product states $\rho_{i,j} = \rho_i \otimes \rho_j$ for every combination of $i, j \in \{1, 2, 3\}$.

Hint: There is of course a brute force solution, but spend a few minutes thinking about whether you can save yourself computational effort.

Solution

We can answer this very quickly, without even considering the form of the individual quantum states. In particular, it is true that the purity of a product state is the product of the purities of each tensor factors. To see this, we only need two ingredients:

(a) The trace of a tensor product is the product of traces: $\operatorname{Tr}(A \otimes B) = \operatorname{Tr}(A)\operatorname{Tr}(B)$.

(b) The square of a tensor product is the tensor product of the squares: $(A \otimes B)^2 = A^2 \otimes B^2$.

The purity is the trace of the square of a matrix, and for this exercise we consider tensor product matrices. Now we use the two ingredients: First, we know the square of the tensor product is the tensor product of squares: $\rho_{i,j}^2 = \rho_i^2 \otimes \rho_j^2$. Second, we know the trace of a tensor product is the product of traces: $\operatorname{Tr}(\rho_i^2 \otimes \rho_j^2) = \operatorname{Tr}\rho_i^2 \operatorname{Tr}\rho_j^2$, completing the necessary bit of math.
Now we just plug in the formula for the individual purity $\operatorname{Tr}\rho_i^2 = i/(i+1)$ to obtain

$$\operatorname{Tr}\rho_{i,j}^2 = \frac{ij}{(i+1)(j+1)}.$$

One way of interpreting this result is that putting together states that are not pure results in a larger state with lower purity.

3 P. **Bonus Exercise 3.** Consider a mixed state as a classical mixture of pure states.

1 P. (a) Can you obtain an entangled mixed state by mixing only pure tensor product states? (Either give an illustrative example or prove that it's not possible.)

Solution

No, this is not possible. By definition, a mixture (aka convex combination) of (possibly mixed) tensor product states is separable and thus not entangled.

1 P. (b) Can you obtain a tensor product mixed state by mixing only pure entangled states? (Either give an illustrative example or prove that it's not possible.)

Solution

Yes, illustrative example: take the Bell basis, which is an ONB of maximally entangled 2-qubit states. Consider the classical mixture of preparing each of the Bell states with uniform probability. Then the resulting state is the maximally mixed state (which would be the case for any ONB of states, irrespective of their entanglement). The maximally mixed state of a d-dimensional system is $\frac{1}{d}\mathbb{I}$. In the case of a 4-dimensional system (2 qubits), we can write the 4-dimensional identity matrix as the tensor product of two 2-dimensional identity matrices, with the correct normalization:

$$\frac{1}{4}\mathbb{I}_4 = \left(\frac{1}{2}\mathbb{I}_2\right) \otimes \left(\frac{1}{2}\mathbb{I}_2\right),$$

so indeed we are left with a product state as a classical mixture of entangled states.

1 P. (c) Given the results of (a) and (b), what do you conclude about how entanglement behaves under probabilistic mixtures?

Solution

Probabilistic mixing cannot create entanglement but it can destroy entanglement.

Total Points: 29 (+12)