Theory of Quantum Information Assignment 3

Name: Nate Stemen (20906566) **Due:** Nov 26, 2021
Email: `nate@stemen.email` **Course:** QIC 820

Problem 1

Let X, Y and Z be registers. Prove that, for any state of these registers, the following two inequalities are true:

(a) $\mathrm{I}(\mathsf{X}, \mathsf{Y} : \mathsf{Z}) + \mathrm{I}(\mathsf{Y} : \mathsf{Z}) \geq \mathrm{I}(\mathsf{X} : \mathsf{Z})$.

(b) $\mathrm{I}(\mathsf{X}, \mathsf{Y} : \mathsf{Z}) \leq \mathrm{I}(\mathsf{Y} : \mathsf{X}, \mathsf{Z}) + 2\,\mathrm{H}(\mathsf{X})$.

Solution. (a) First note that the mutual information is symmetric with respect to it's arguments. This property is inherited from the symmetry of the von Neumann entropy for a compound register: $\mathrm{H}(\mathsf{X}, \mathsf{Y}) = \mathrm{H}(\mathsf{Y}, \mathsf{X})$.

$$\mathrm{I}(\mathsf{X} : \mathsf{Y}) = \mathrm{H}(\mathsf{X}) + \mathrm{H}(\mathsf{Y}) - \mathrm{H}(\mathsf{X}, \mathsf{Y}) = \mathrm{H}(\mathsf{Y}) + \mathrm{H}(\mathsf{X}) - \mathrm{H}(\mathsf{Y}, \mathsf{X}) = \mathrm{I}(\mathsf{Y} : \mathsf{X})$$

We also have the ability to permute registers inside either argument to the mutual information.

$$\mathrm{I}(\mathsf{X} : \mathsf{Y}, \mathsf{Z}) = \mathrm{H}(\mathsf{X}) + \mathrm{H}(\mathsf{Y}, \mathsf{Z}) - \mathrm{H}(\mathsf{X}, \mathsf{Y}, \mathsf{Z}) = \mathrm{H}(\mathsf{X}) + \mathrm{H}(\mathsf{Z}, \mathsf{Y}) - \mathrm{H}(\mathsf{X}, \mathsf{Z}, \mathsf{Y}) = \mathrm{I}(\mathsf{X} : \mathsf{Z}, \mathsf{Y})$$

Using these properties along with corollary 5.37 which says $\mathrm{I}(\mathsf{A} : \mathsf{B}, \mathsf{C}) \geq \mathrm{I}(\mathsf{A} : \mathsf{C})$ we have the following manipulation.

$$\mathrm{I}(\mathsf{X}, \mathsf{Y} : \mathsf{Z}) = \mathrm{I}(\mathsf{Z} : \mathsf{X}, \mathsf{Y}) = \mathrm{I}(\mathsf{Z} : \mathsf{Y}, \mathsf{X}) \geq \mathrm{I}(\mathsf{Z} : \mathsf{X}) = \mathrm{I}(\mathsf{X} : \mathsf{Z})$$

To complete the proof of the desired inequality, we now need to show $\mathrm{I}(\mathsf{Y} : \mathsf{Z}) \geq 0$ which follows from the subadditivity of the von Neumann entropy.

$$\begin{aligned} \mathrm{H}(\mathsf{Y}) + \mathrm{H}(\mathsf{Z}) &\geq \mathrm{H}(\mathsf{Y}, \mathsf{Z}) \\ \mathrm{H}(\mathsf{Y}) + \mathrm{H}(\mathsf{Z}) - \mathrm{H}(\mathsf{Y}, \mathsf{Z}) &\geq 0 \\ \mathrm{I}(\mathsf{Y} : \mathsf{Z}) &\geq 0. \end{aligned}$$

This completes the proof.

(b) In this part we use both the subadditivity of the von Neumann entropy, together with Theorem 5.25 which states $\mathrm{H}(\mathsf{X}) \leq \mathrm{H}(\mathsf{Y}) + \mathrm{H}(\mathsf{X}, \mathsf{Y})$.

$$\begin{aligned} \mathrm{I}(\mathsf{X}, \mathsf{Y} : \mathsf{Z}) &:= \mathrm{H}(\mathsf{X}, \mathsf{Y}) + \mathrm{H}(\mathsf{Z}) - \mathrm{H}(\mathsf{X}, \mathsf{Y}, \mathsf{Z}) \\ &\leq \mathrm{H}(\mathsf{X}) + \mathrm{H}(\mathsf{Y}) + \mathrm{H}(\mathsf{Z}) - \mathrm{H}(\mathsf{X}, \mathsf{Y}, \mathsf{Z}) && \text{(subadditivity)} \\ &\leq \mathrm{H}(\mathsf{X}) + \mathrm{H}(\mathsf{Y}) + \Big[\mathrm{H}(\mathsf{X}) + \mathrm{H}(\mathsf{X}, \mathsf{Z})\Big] - \mathrm{H}(\mathsf{X}, \mathsf{Y}, \mathsf{Z}) && \text{(5.25)} \\ &= \mathrm{H}(\mathsf{Y}) + \mathrm{H}(\mathsf{X}, \mathsf{Z}) - \mathrm{H}(\mathsf{Y}, \mathsf{X}, \mathsf{Z}) + 2\,\mathrm{H}(\mathsf{X}) && \text{(symmetry of H)} \\ &= \mathrm{I}(\mathsf{Y} : \mathsf{X}, \mathsf{Z}) + 2\,\mathrm{H}(\mathsf{X}) \end{aligned}$$

So there we go.

Problem 2

Let $\mathcal{X}$ be a complex Euclidean space, let Σ be an alphabet, let $p \in \mathcal{P}(\Sigma)$ be a probability vector, and let $\{\rho_a : a \in \Sigma\} \subseteq \mathrm{D}(\mathcal{X})$ be a collection of states. Prove that, for every state $\sigma \in \mathrm{D}(\mathcal{X})$, we have

$$\sum_{a\in\Sigma} p(a)\mathrm{D}(\rho_a\|\sigma) = \sum_{a\in\Sigma} p(a)\mathrm{D}(\rho_a\|\rho) + \mathrm{D}(\rho\|\sigma),$$

where

$$\rho = \sum_{a\in\Sigma} p(a)\rho_a.$$

Solution. By the linearity of the trace we have

$$\mathrm{D}(\rho\|\sigma) = \mathrm{tr}(\rho\log\rho) - \mathrm{tr}(\rho\log\sigma) = \sum_{a\in\Sigma} p(a)\big[\mathrm{tr}(\rho_a\log\rho) - \mathrm{tr}(\rho_a\log\sigma)\big].$$

Thus the left hand side of the equation to be proven is finite when $\mathrm{im}(\rho_a) \subseteq \mathrm{im}(\rho)$ which is always true given the form of ρ, and $\mathrm{im}(\rho_a) \subseteq \mathrm{im}(\sigma)$. Given that holds for each $a \in \Sigma$, then it must also hold for a convex combination of them, and hence $\mathrm{im}(\rho) \subseteq \mathrm{im}(\sigma)$. This implies that when the left hand side is finite, so is the right hand side, and the reasoning can be run in reverse to show both sides are infinite when $\mathrm{D}(\rho_a\|\sigma)$ is infinite.

Now we can address the equality when both sides are finite.

$$\begin{aligned}
&\sum_{a\in\Sigma} p(a)\mathrm{D}(\rho_a\|\rho) + \mathrm{D}(\rho\|\sigma) \\
&\qquad := \sum_{a\in\Sigma} p(a)\Big[\mathrm{tr}(\rho_a\log\rho_a) - \mathrm{tr}(\rho_a\log\rho)\Big] + \mathrm{tr}(\rho\log\rho) - \mathrm{tr}(\rho\log\sigma) \\
&\qquad = \sum_{a\in\Sigma} p(a)\,\mathrm{tr}(\rho_a\log\rho_a) - \mathrm{tr}(\rho\log\rho) + \mathrm{tr}(\rho\log\rho) - \mathrm{tr}(\rho\log\sigma) \\
&\qquad = \sum_{a\in\Sigma} p(a)\,\mathrm{tr}(\rho_a\log\rho_a) - \mathrm{tr}(\rho\log\sigma) \\
&\qquad = \sum_{a\in\Sigma} p(a)\Big[\mathrm{tr}(\rho_a\log\rho_a) - \mathrm{tr}(\rho_a\log\sigma)\Big] \\
&\qquad =: \sum_{a\in\Sigma} p(a)\mathrm{D}(\rho_a\|\sigma)
\end{aligned}$$

Problem 3

Let X and Y be registers, let Σ be an alphabet, let $p \in \mathcal{P}(\Sigma)$ be a probability vector, let $\{\sigma_a : a \in \Sigma\} \subseteq \mathrm{D}(\mathcal{X})$ and $\{\xi_a : a \in \Sigma\} \subseteq \mathrm{D}(\mathcal{Y})$ be collections of density operators, and define $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$ as

$$\rho = \sum_{a \in \Sigma} p(a) \sigma_a \otimes \xi_a.$$

States that can be expressed in this way are called *separable states*, and will be discussed in the fourth part of the course—but nothing from that part of the course is needed to solve this problem.

(a) Prove that $\mathrm{I}(\mathsf{X} : \mathsf{Y})_\rho \leq \mathrm{H}(p)$.

(b) Prove that

$$\mathrm{H}(\rho) \geq \sum_{a \in \Sigma} p(a)\, \mathrm{H}(\sigma_a) + \mathrm{H}\left(\sum_{a \in \Sigma} p(a) \xi_a \right).$$

Solution completed in collaboration with Mohammad Ayyash,[1] and Nicholas Zutt.[2]

Solution. (a) By the definition of mutual information we have $\mathrm{I}(\mathsf{X} : \mathsf{Y}) := \mathrm{H}(\mathsf{X}) + \mathrm{H}(\mathsf{Y}) - \mathrm{H}(\mathsf{X}, \mathsf{Y})$ and the fact we are making this calculation with respect to ρ we have $\mathrm{H}(\mathsf{X}, \mathsf{Y}) = \mathrm{H}(\rho)$. We can then write the mutual information as $\mathrm{I}(\mathsf{X} : \mathsf{Y}) = \mathrm{H}(\rho[\mathsf{X}]) = \mathrm{H}(\rho[\mathsf{Y}]) - \mathrm{H}(\rho)$.

Let's now spectral decompose each σ_a as $\sigma_a = \sum_{b \in \Gamma} \lambda_{ab} |\psi_{ab}\rangle\langle\psi_{ab}|$ where we've taken $\mathcal{X} = \mathbb{C}^\Gamma$. We can now calculate the associated terms of $\mathrm{I}(\mathsf{X} : \mathsf{Y})$. Here we use a fact that $\mathrm{H}(\sum_{a \in \Sigma} p(a) \rho_a) \leq \mathrm{H}(p) + \sum_{a \in \Sigma} p(a)\, \mathrm{H}(\rho_a)$ for $p \in \mathcal{P}(\Sigma)$ and ρ_a density operators. This fact is formalized in the following theorem from Nielsen & Chuang's *Quantum Computation and Quantum Information* in Theorem 11.10. I've taken the liberty to formulate the theorem from their book into the language and notation of this course, as well as omit some unnecessary information.

Fact (Theorem 11.10 Neilsen & Chuang)**.** *Take Σ to be an alphabet, $p \in \mathcal{P}(\Sigma)$ a probability vector, and $\{\rho_a : a \in \Sigma\} \subseteq \mathrm{D}(\mathcal{X})$ a collection of density operators for some complex Euclidean space $\mathcal{X}$. If $\rho := \sum_{a \in \Sigma} p(a) \rho_a$ is a convex combination of density operators, then*

$$\mathrm{H}(\rho) \leq \mathrm{H}(p) + \sum_{a \in \Sigma} p(a)\, \mathrm{H}(\rho_a).$$

Using this we then have

$$\mathrm{H}(\rho[\mathsf{X}]) = \mathrm{H}\left(\sum_{a \in \Sigma} p(a) \sigma_a \right) \leq \mathrm{H}(p) + \sum_{a \in \Sigma} p(a)\, \mathrm{H}(\sigma_a),$$

$$\mathrm{H}(\rho[\mathsf{Y}]) = \mathrm{H}\left(\sum_{a \in \Sigma} p(a) \xi_a \right) \leq \mathrm{H}(p) + \sum_{a \in \Sigma} p(a)\, \mathrm{H}(\xi_a).$$

[1] mmayyash@uwaterloo.ca
[2] nzutt@uwaterloo.ca

Finally for the last term we have

$$\begin{aligned}
\mathrm{H}(\rho) &= \mathrm{H}\left(\sum_{\substack{a\in\Sigma\\ b\in\Gamma}} p(a)\lambda_{ab}\,|\psi_{ab}\rangle\langle\psi_{ab}|\otimes\xi_a\right) \\
&\geq \sum_{a\in\Sigma} p(a)\,\mathrm{H}\left(\sum_{b\in\Gamma}\lambda_{ab}\,|\psi_{ab}\rangle\langle\psi_{ab}|\otimes\xi_a\right) && \text{(Concavity)} \\
&= \sum_{a,b} p(a)\,\mathrm{H}(\lambda_{ab}\xi_a).
\end{aligned}$$

We can now use Equation 5.98 in the book which states $\mathrm{H}(\alpha P) = \alpha\,\mathrm{H}(P) - \alpha\log(\alpha)\,\mathrm{tr}(P)$ to pull the constants out.

$$\begin{aligned}
\sum_{a,b} p(a)\,\mathrm{H}(\lambda_{ab}\xi_a) &= \sum_{a,b} p(a)\lambda_{ab}\,\mathrm{H}(\xi_a) - p(a)\lambda_{ab}\log(p(a)\lambda_{ab})\,\mathrm{tr}(\xi_a) \\
&= \sum_{a,\,b} p(a)\,\lambda_{ab}\,\mathrm{H}(\xi_a) - p(a)\,\lambda_{ab}\,\log(p(a)) - p(a)\lambda_{ab}\log(\lambda_{ab})
\end{aligned}$$

The highlighted terms vanish (with the sum on b) because $\mathrm{tr}(\sigma_a) = \sum_b \lambda_{ab} = 1$.

$$\begin{aligned}
\mathrm{H}(\rho) &\geq \sum_{a\in\Sigma} p(a)\,\mathrm{H}(\xi_a) - p(a)\log(p(a)) - p(a)\sum_b \lambda_{ab}\log(\lambda_{ab}) \\
&= \mathrm{H}(p) + \sum_{a\in\Sigma} p(a)\,\mathrm{H}(\xi_a) + p(a)\,\mathrm{H}(\sigma_a)
\end{aligned}$$

We can now put together all of the terms as follows.

$$\begin{aligned}
\mathrm{I}(\mathsf{X}:\mathsf{Y})_\rho &:= \mathrm{H}(\rho[X]) + \mathrm{H}(\rho[Y]) - \mathrm{H}(\rho) \\
&\leq 2\,\mathrm{H}(p) + \left[\sum_{a\in\Sigma} p(a)\Big(\mathrm{H}(\sigma_a) + \mathrm{H}(\xi_a)\Big)\right] - \mathrm{H}(p) - \sum_{a\in\Sigma} p(a)\Big(\mathrm{H}(\xi_a) + \mathrm{H}(\sigma_a)\Big) \\
&= \mathrm{H}(p)
\end{aligned}$$

(b) In the previous part we showed

$$\mathrm{H}(\rho) \geq \mathrm{H}(p) + \sum_{a\in\Sigma} p(a)\Big(\mathrm{H}(\xi_a) + \mathrm{H}(\sigma_a)\Big)$$

and hence to show $\mathrm{H}(\rho) \geq \sum_{a\in\Sigma} p(a)\,\mathrm{H}(\sigma_a) + \mathrm{H}\big(\sum_{a\in\Sigma} p(a)\xi_a\big)$ we can prove

$$\mathrm{H}(p) + \sum_{a\in\Sigma} p(a)\,\mathrm{H}(\xi_a) \geq \mathrm{H}\left(\sum_{a\in\Sigma} p(a)\xi_a\right)$$

which is exactly the same fact we used above with $\rho_a = \xi_a$.

Problem 4

Let X be a register having alphabet Σ, and also let Y and Z be registers (having arbitrary alphabets we need not name). Let $\{\sigma_a : a \in \Sigma\} \subseteq \mathrm{D}(\mathcal{Y} \otimes \mathcal{Z})$ be a collection of density operators, let $p \in \mathcal{P}(\Sigma)$ be a probability vector, and define $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Y} \otimes \mathcal{Z})$ as

$$\rho = \sum_{a \in \Sigma} p(a) \, |a\rangle\langle a| \otimes \sigma_a.$$

Prove that, with respect to the state ρ, one has

$$\mathrm{I}(\mathsf{X}, \mathsf{Y} : \mathsf{Z}) \leq \mathrm{I}(\mathsf{Y} : \mathsf{X}, \mathsf{Z}) + \mathrm{H}(\mathsf{X}).$$

Solution completed in collaboration with Margie Christ.[3]

Solution. To prove this inequality we will first calculate the needed entropies.

$$\mathrm{H}(\mathsf{X}) = \mathrm{H}(\rho[\mathsf{X}]) = \mathrm{H}\left(\sum_{a \in \Sigma} p(a) \, |a\rangle\langle a|\right) = \mathrm{H}(p)$$

$$\mathrm{H}(\mathsf{Y}) = \mathrm{H}(\rho[\mathsf{Y}]) = \mathrm{H}\left(\sum_{a \in \Sigma} p(a) \, \mathrm{tr}(|a\rangle\langle a|) \otimes \mathrm{tr}_{\mathcal{Z}} \, \sigma_a\right) = \mathrm{H}\left(\sum_{a \in \Sigma} p(a) \, \mathrm{tr}_{\mathcal{Z}} \, \sigma_a\right)$$

$$\mathrm{H}(\mathsf{Z}) = \mathrm{H}(\rho[\mathsf{Z}]) = \mathrm{H}\left(\sum_{a \in \Sigma} p(a) \, \mathrm{tr}(|a\rangle\langle a|) \otimes \mathrm{tr}_{\mathcal{Y}} \, \sigma_a\right) = \mathrm{H}\left(\sum_{a \in \Sigma} p(a) \, \mathrm{tr}_{\mathcal{Y}} \, \sigma_a\right)$$

$$\begin{aligned}
\mathrm{H}(\mathsf{X}, \mathsf{Y}) &= \mathrm{H}(\rho[\mathsf{X}, \mathsf{Y}]) = \mathrm{H}\left(\sum_{a \in \Sigma} p(a) \, |a\rangle\langle a| \otimes \mathrm{tr}_{\mathcal{Z}} \, \sigma_a\right) \\
&= \sum_{a \in \Sigma} \mathrm{H}(p(a) \, \mathrm{tr}_{\mathcal{Z}} \, \sigma_a) && \text{(Block diagonal form of } |a\rangle\langle a| \otimes \mathrm{tr}_{\mathcal{Z}} \, \sigma_a) \\
&= \mathrm{H}(p) + \sum_{a \in \Sigma} p(a) \, \mathrm{H}(\mathrm{tr}_{\mathcal{Z}} \, \sigma_a) && (\mathrm{H}(\alpha P) = \alpha \, \mathrm{H}(P) - \alpha \log \alpha \, \mathrm{tr}(P))
\end{aligned}$$

$$\begin{aligned}
\mathrm{H}(\mathsf{X}, \mathsf{Z}) &= \mathrm{H}(\rho[\mathsf{X}, \mathsf{Z}]) = \mathrm{H}\left(\sum_{a \in \Sigma} p(a) \, |a\rangle\langle a| \otimes \mathrm{tr}_{\mathcal{Y}} \, \sigma_a\right) \\
&= \sum_{a \in \Sigma} \mathrm{H}(p(a) \, \mathrm{tr}_{\mathcal{Y}} \, \sigma_a) \\
&= \mathrm{H}(p) + \sum_{a \in \Sigma} p(a) \, \mathrm{H}(\mathrm{tr}_{\mathcal{Y}} \, \sigma_a)
\end{aligned}$$

$$\mathrm{H}(\mathsf{Y}, \mathsf{Z}) = \mathrm{H}(\rho[\mathsf{Y}, \mathsf{Z}]) = \mathrm{H}\left(\sum_{a \in \Sigma} p(a) \sigma_a\right)$$

We can cancel out the $\mathrm{H}(\mathsf{X}, \mathsf{Y}, \mathsf{Z})$ term from both sides of our desired inequality to get

$$\mathrm{H}(\mathsf{X}, \mathsf{Y}) + \mathrm{H}(\mathsf{Z}) \leq \mathrm{H}(\mathsf{Y}) + \mathrm{H}(\mathsf{X}, \mathsf{Z}) + \mathrm{H}(\mathsf{X}).$$

[3] mchrist@uwaterloo.ca

Starting on the right hand side and using the above calculated entropies we have

$$\begin{aligned}
\mathrm{H}(\mathsf{Y}) + \mathrm{H}(\mathsf{X},\mathsf{Z}) + \mathrm{H}(\mathsf{X}) &= \mathrm{H}\left(\sum_{a\in\Sigma} p(a)\,\mathrm{tr}_{\mathcal{Z}}\,\sigma_a\right) + \mathrm{H}(p) + \sum_{a\in\Sigma} p(a)\,\mathrm{H}(\mathrm{tr}_{\mathcal{Y}}\,\sigma_a) + \mathrm{H}(p) \\
&= \mathrm{H}(\mathsf{Y}) + 2\,\mathrm{H}(\mathsf{X}) + \sum_{a\in\Sigma} p(a)\,\mathrm{H}(\mathrm{tr}_{\mathcal{Y}}\,\sigma_a) \\
&\geq \mathrm{H}(\mathsf{X}) + \mathrm{H}(\mathsf{X},\mathsf{Y}) + \sum_{a\in\Sigma} p(a)\,\mathrm{H}(\mathrm{tr}_{\mathcal{Y}}\,\sigma_a) \qquad \text{(subadditivity)}
\end{aligned}$$

Since $\mathrm{H}(\mathsf{X},\mathsf{Y})$ appears on both sides, we can cancel the term and it will suffice to show that $\mathrm{H}(\mathsf{X}) + \sum_{a\in\Sigma} p(a)\,\mathrm{H}(\mathrm{tr}_{\mathcal{Y}}\,\sigma_a) \geq \mathrm{H}(\mathsf{Z})$. Expressed slightly differently we have

$$\mathrm{H}(p) + \sum_{a\in\Sigma} p(a)\,\mathrm{H}(\mathrm{tr}_{\mathcal{Y}}\,\sigma_a) \geq \mathrm{H}\left(\sum_{a\in\Sigma} p(a)\,\mathrm{tr}_{\mathcal{Y}}\,\sigma_a\right).$$

Which is an immediate consequence of the following theorem from Nielsen & Chuang with $\rho_a = \mathrm{tr}_{\mathcal{Y}}\,\sigma_a$. I've taken the liberty to formulate the theorem from their book into the language and notation of this course, as well as omit some unnecessary information.

Fact (Theorem 11.10 Neilsen & Chuang)**.** *Take Σ to be an alphabet, $p \in \mathcal{P}(\Sigma)$ a probability vector, and $\{\rho_a : a \in \Sigma\} \subseteq \mathrm{D}(\mathcal{X})$ a collection of density operators for some complex Euclidean space $\mathcal{X}$. If $\rho := \sum_{a\in\Sigma} p(a)\rho_a$ is a convex combination of density operators, then*

$$\mathrm{H}(\rho) \leq \mathrm{H}(p) + \sum_{a\in\Sigma} p(a)\,\mathrm{H}(\rho_a).$$