Part III Quantum Information Theory: Example Sheet 3 for the class on 27th Nov, 2pm in MR14

16th November

If anything is unclear or you think you've found a mistake, please email wm266@cam.ac.uk.

1. Show that for any random variable (RV) X taking values in a finite set $\mathcal{A}_X$:
   (a) $0 \leq H(X)$ with equality iff there is some $k \in \mathcal{A}_X$ such that $\Pr(X = k) = 1$.
   (b) $H(X) \leq \log |\mathrm{supp}(P_X)| \leq \log |\mathcal{A}_X|$, with equality in the first inequality iff $\Pr(X = k) = 1/|\mathrm{supp}(P_X)|$ for all $k \in \mathrm{supp}(P_X)$.
2. When is $S(\rho) = 0$?
3. For a memoryless source $Z_i \overset{iid}{\sim} P_Z$ where $\mathcal{A}_Z = \{0, 1\}$, and $P_Z(1) = 1/4$ give an explicit blocklength n_0 such that, for all $n \geq n_0$ a compression rate of $1.05 \times H(Z)$ can be achieved with error probability no more than $1/1000$.
4. Show that for random variables X and Y:
$$I(X : Y) = D(P_{XY} \| P_X P_Y) = \sum_x P_X(x) D(P_{Y|X=x} \| P_Y) = \sum_y P_Y(y) D(P_{X|Y=y} \| P_X)$$
5. Prove the chain rule for conditional mutual information:
$$I(X : Y, Z|W) = I(X : Z|W) + I(X : Y|Z, W).$$
6. Give examples of RVs X, Y, Z for which
   (a) $I(X : Y) = 1$ and $I(X : Y|Z) = 0$,
   (b) $I(X : Y) = 0$ and $I(X : Y|Z) = 1$.
7. Show that, if X, Y, Z form a Markov chain $X - Y - Z$, then the *data processing inequality* $I(X : Z) \leq I(X : Y)$ is satisfied.

8. Suppose that random variables X_1, X_2, X_3, X_4 form a Markov chain $X_1 - X_2 - X_3 - X_4$. Use the properties of conditional mutual information / conditional entropy to show that (a) $X_1 - X_2 - X_3$, (b) $X_2 - X_3 - X_4$, (c) $X_1 - X_2 - X_4$, (d) $X_1 - X_3 - X_4$.

9. Prove that for a bipartite system AB in any state ρ_{AB}, $H(\mathsf{AB})_\rho \geq |H(\mathsf{A})_\rho - H(\mathsf{B})_\rho|$.

10. Prove that $I(\mathsf{A}:\mathsf{B})_\rho := D(\rho_{\mathsf{AB}} \| \rho_{\mathsf{A}} \otimes \rho_{\mathsf{B}})$.

11. Given states $\rho^{(j)}$, $\sigma^{(j)}$ for $j \in \mathcal{A} = \{0, 1, \ldots, d_{\mathsf{R}} - 1\}$ and probability distributions p and q on $\mathcal{A}$, by considering the quantity

$$D\left(\sum_{j\in\mathcal{A}} p(j)|j\rangle\langle j|_{\mathsf{R}} \otimes \rho^{(j)}_{\mathsf{Q}} \middle\| \sum_{j\in\mathcal{A}} q(j)|j\rangle\langle j|_{\mathsf{R}} \otimes \sigma^{(j)}_{\mathsf{Q}}\right)$$

prove the joint convexity of the quantum relative entropy:

$$D\left(\sum_{j\in\mathcal{A}} p(j)\rho^{(j)}_{\mathsf{Q}} \middle\| \sum_{j\in\mathcal{A}} p(j)\sigma^{(j)}_{\mathsf{Q}}\right) \leq \sum_{j\in\mathcal{A}} p(j) D(\rho^{(j)}_{\mathsf{Q}} \| \sigma^{(j)}_{\mathsf{Q}}).$$

12. (a) Use the chain rule for conditional entropies and the positivity of mutual information to prove that the entropy $H(X)$ is a concave function of P_X (here we are regarding P_X as belonging to the convex subset of probability mass functions in $\mathbb{R}^{\mathcal{A}_X}$). (Hint: introduce a binary RV W, which "selects" the distribution for X.)

    (b) Given $P_{YX} = N_{Y|X} P_X$, (i.e. $P_{YX}(y,x) = N_{Y|X}(y|x)P_X(x)$, for all x, y), consider the function $f : (N_{Y|X}, P_X) \mapsto I(X:Y)$. Show that f is a convex function of the conditional probability distribution $N_{Y|X}$ (for any fixed P_X) and a concave function of the input probability distribution P_X (for any fixed $N_{Y|X}$).

13. Use Shannon's noisy channel coding theorem to find an expression for the capacity of a binary symmetric channel in terms of its 'bit-flip' probability f.

14. Suppose that for a particular channel there is a block error correcting code of block length n, message set $\{0,1\}^k$, and error probability $\Pr(M \neq \hat{M}) \leq \epsilon$. Here $\hat{M}$ is the estimate made by the decoder of the message M, and M has the uniform distribution on $\{0,1\}^k$. Show that there exists another code for the same channel, with block length n, message set $\{0,1\}^{k-1}$ and *maximum* error probability $\max_{m\in\{0,1\}^{k-1}} \Pr(M \neq \hat{M} | M = m) \leq 2\epsilon$.

15. Show that the Holevo information of an operation is always attained by an input ensemble of *pure* states: That is

$$\chi(\mathcal{N}^{\mathsf{B}\leftarrow\mathsf{A}}) = I(\hat{\mathsf{X}}:\mathsf{B})_\sigma$$

where $\sigma_{\hat{\mathsf{X}}\mathsf{B}} = \mathcal{N}^{\mathsf{B}\leftarrow\mathsf{A}} \rho_{\hat{\mathsf{X}}\mathsf{A}}$ for some $\rho_{\hat{\mathsf{X}}\mathsf{A}} = \sum_x p(x)|x\rangle\langle x|_{\hat{\mathsf{X}}} \otimes |\psi^{(x)}\rangle\langle\psi^{(x)}|_{\mathsf{A}}$.