Q. Inf. Science 3 (8.372) — Fall 2024

Assignment 4

Due: **Tuesday**, *Oct 8, 2024 at* ***9pm***

Turning in your solutions: Upload a single pdf file to gradescope.

1. **Mutual information and relative entropy** Given a state ρ_{AB} prove that

$$\min_{\sigma_A,\omega_B} D(\rho_{AB}\|\sigma_A \otimes \omega_B) = I(A;B)_\rho \tag{1}$$

2. **Compression with side information.**

   (a) *Conditionally typical set.* For a probability distribution p_{XY} define $J^n_{p,\delta}$ to be the jointly typical set: formally $J^n_{p,\delta} := T^n_{p,\delta} \cap (T^n_{p_X,\delta} \times T^n_{p_Y,\delta})$. Given y^n, define the conditionally typical set $J(y^n) := J^n_{p,\delta}(y^n)$ by

$$J(y^n) = \left\{x^n \in X^n : (x^n, y^n) \in J^n_{p,\delta}\right\}. \tag{2}$$

   Observe that if $y^n \notin T^n_{p_Y,\delta}$ then $J(y^n)$ is empty. If $y^n \in T^n_{p_Y,\delta}$ then what bounds can you place on $p^n(x^n|y^n)$ for $x^n \in J(y^n)$? Prove that

$$|J(y^n)| \leq \exp(n(H(X|Y) + 2\delta)). \tag{3}$$

   (b) Let $(X^n, Y^n) \sim p^n_{XY}$, i.e. each (X_i, Y_i) is drawn independently from p_{XY}. Suppose that Alice knows X^n and Y^n, Bob holds Y^n and Alice wishes to transmit X^n to Bob. Shannon's noiseless coding theorem tells her how to do this using $\approx nH(X)$ bits, but this would not take advantage of the correlations between X^n and Y^n. Show that she can transmit X^n to Bob using $n(H(X|Y) + \delta)$ bits and error ϵ, with $\epsilon, \delta \to 0$ as $n \to \infty$. (Note: the δ in (a) might not be the same δ as the one here.)

   (c) Now suppose that Alice knows only X^n and Bob knows Y^n. This is significantly more challenging than the situation in (b). Suppose that Alice uses a random codebook, as we will also see in Shannon's noisy coding theorem. To compress to rate R, Alice uses a random function $E : X^n \to [2^{nR}] := \{1, 2, \ldots, 2^{nR}\}$, meaning that each $E(x^n)$ is chosen independently and uniformly from $[2^{nR}]$. As in the channel coding theorem, E is chosen randomly and then fixed and can be assumed to be known by both parties.

   Given message m, Bob decodes by choosing the unique x^n such that $E(x^n) = m$ and $(x^n, Y^n) \in J$, i.e. in the set $E^{-1}(m) \cap J(Y^n)$. If this x^n either doesn't exist or isn't unique, then he declares failure. Let WRONG be the event where

$$E^{-1}(m) \cap J(Y^n) \tag{4}$$

contains a string x^n that is not equal to the correct string X^n. Prove that $p^n(\text{WRONG}) \to 0$ if $R > H(X|Y) + 3\delta$ as $n \to \infty$.

(d) What other errors are possible? By bounding their probabilities show that the coding strategy in (c) can work with error approaching 0 as $n \to \infty$ for any $R > H(X|Y)$.

3. **Feedback-assisted capacity** The proof in lecture of Shannon's noisy coding theorem did not allow Bob to send messages back to Alice, an ability called "feedback." Suppose that after receiving each channel output Y_i, Bob can noiselessly send Alice an arbitrary message. Modify the proof in lecture to show that the same converse still holds. You do not need to repeat the parts of the proof that are unchanged. As a hint, try to show that $I(M; Y^n) \leq H(Y^n) - \sum_{i=1}^{n} H(Y_i|X_i)$.