1. *Multiple access channel* Consider a multiple access channel with inputs $\mathcal{X}_1 = \{0, 1, 2, 3\}$ and $\mathcal{X}_2 = \{0, 1\}$ The channel is given by:

$$\mathsf{Y} = \mathsf{X}_1 + \mathsf{X}_2 \bmod 4 \tag{1}$$

Find the capacity region for this channel.

2. *Slepian-Wolf* Let X_i and Z_i be independent random variables with $p_{\mathsf{X}}(x) = [1-p, p]$ and $p_{\mathsf{Z}}(z) = [1-r, r]$ for $0 \leq p, r \leq 1$. Let $\mathsf{Y}_i = \mathsf{X}_i \oplus \mathsf{Z}_i$ where $\oplus$ denotes addition modulo 2. Let the source vector $\mathbf{X} = (\mathsf{X}_1, \dots, X_n)$ be encoded at rate R_1 and let $\mathbf{Y} = (\mathsf{Y}_1, \dots, Y_n)$ be encoded at rate R_2.

   (a) What is $p_{\mathsf{Y}}(y)$?

   (b) What region of rates allows recovering of $\mathbf{X}, \mathbf{Y}$ with probability of error tending to zero as $n \to \infty$? Draw a pentagon-shpaed region and label the key points.

   (c) On the same figure, draw the region of achievable rates, assuming the correlation between X and Y is ignored.

3. Suppose that $(\mathsf{X}, \mathsf{Y}, \mathsf{Z})$ are jointly Gaussian and that $\mathsf{X} \to \mathsf{Y} \to \mathsf{Z}$ forms a Markov chain. Let X and Y have correlation coefficient ρ_1 and let Y and Z have correlation coefficient ρ_2.

   1. Find $E[\mathsf{XZ}]$. (Hint: $E[\mathsf{X}|\mathsf{Y} = y] = \rho_1 \frac{\sigma_X^2}{\sigma_Y^2} y$.)
   2. Find $I(\mathsf{X}; \mathsf{Z})$.

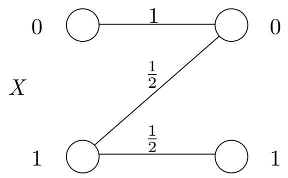
4. Let $\Pr(\mathsf{X}=1)=p$, $\Pr(\mathsf{X}=0)=1-p$, and let $\mathsf{Y}=\mathsf{X}+\mathsf{Z}$, where Z is uniform over the interval $[0,a]$, $a>1$, and Z is independent of X.

    1. Calculate $I(\mathsf{X};\mathsf{Y})=H(\mathsf{X})-H(\mathsf{X}|\mathsf{Y})$.
    2. Now calculate $I(\mathsf{X};\mathsf{Y})$ the other way by $I(\mathsf{X};\mathsf{Y})=H(\mathsf{Y})-H(\mathsf{Y}|\mathsf{X})$.
    3. Calculate the capacity of this channel by maximizing over p.

5. *Joint AEP for the binary Z channel* The binary Z channel is a DMC with binary inputs $\mathcal{X} = \{0, 1\}$ and binary outputs $\mathcal{Y} = \{0, 1\}$ and conditional probability distribution $p_{\mathsf{Y}|\mathsf{X}}(y|x)$ given by matrix:

$$\begin{bmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}. \tag{32}$$

The channel diagram looks like a "Z":



The channel is memoryless:

$$p_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p_{\mathsf{Y}|\mathsf{X}}(y|x)(y_i|x_i)$$

Assume the input distribution is uniform $p_{\mathsf{X}}(x) = [\frac{1}{2}, \frac{1}{2}]$.

(a) Let m be the number of one's in $\mathbf{y}$. Find $p_{\mathbf{Y}}(\mathbf{y})$.

(b) For what values of m and n does the following hold (equivalent to $\mathcal{T}_\epsilon^{(n)}$ with $\epsilon = 0$):

$$-\frac{1}{n} \log p_{\mathbf{Y}}(\mathbf{y}) = H(\mathsf{Y}).$$

(c) Using the memoryless property, compute the following quantities:

- $p_{Y|X}(000|001) =$
- $p_{Y|X}(001|001) =$
- $p_{Y|X}(010|001) =$
- $p_{Y|X}(011|001) =$

Continued on next page

Let $\mathbf{x} = (x_1, \dots, x_n)$ be a sequence with k ones, with $k \leq n$. Consider any sequence $\mathbf{x}$ and $\mathbf{y}$. If $x_i = 0$ and $y_i = 1$, then $p_{\mathsf{XY}}(\mathbf{x}, \mathbf{y}) = 0$ (for any $i = 1, 2, \dots, n$). Let $\mathcal{V}$ be the valid sequences:

$$\mathcal{V} \;=\; \big\{(\mathbf{x}, \mathbf{y}) : x_i \neq 0 \;\text{ or }\; y_i \neq 1, \text{all } i = 1, 2, \dots, n\big\}$$

Note that $p(\mathbf{x}, \mathbf{y}|\mathcal{V}^c) = 0$, where $\mathcal{V}^c$ is the complement of $\mathcal{V}$. For $n = 2$, here is a list of valid sequences V and invalid sequences V^c:

Valid $\mathcal{V}$, $(x_1 x_2, y_1 y_2)$	Not Valid $\mathcal{V}^c$, (x_1, x_2, y_1, y_2)
(00,00)	
(00,01)	(01,00)
(01,01)	(10,00)
(00,10)	(11,00)
(10,10)	(10,01)
(00,11)	(11,01)
(01,11)	(01,10)
(10,11)	(11,10)
(11,11)	

(d) Find $\Pr(\mathbf{y}|\mathbf{x}\mathcal{V})$ (That is, find $\Pr(\mathbf{y}|\mathbf{x})$, given a valid input/output sequence). Express using k, the number of ones in $\mathbf{x}$.

(e) Using your answer to part (d), find $p_{\mathsf{XY}}(\mathbf{x}, \mathbf{y})$. Then, find $-\frac{1}{n} \log p_{\mathsf{XY}}(\mathbf{x}, \mathbf{y})$. Express using n and k.

(f) Find $H(\mathsf{X}, \mathsf{Y})$.

(g) Let $\mathcal{T}'_\epsilon$ be:

$$\mathcal{T}'_\epsilon = \left\{ (x^n, y^n) : \left| -\frac{1}{n} \log p(x^n, y^n) - H(X,Y) \right| < \epsilon \right\}.$$

For $n = 20$ and $\epsilon = 0.06$, describe the sequences that are in the set $\mathcal{T}'_\epsilon$.

7. *General information theory* For each question (a)-(d), write an expression

(a) For the Markov chain $X \to Y \to Z$, write the data processing inequality.

(b) For variables X and Y, write an inequality expressing "independence bound on entropy"

(c) For variables X and Y, write the entropy chain rule:

(d) For variables X and Y write an inequality expressing "conditioning reduces entropy":

(e) **True** or **False** ? A Markov chain $X_1, X_2, X_3 \ldots,$ has entropy rate $H(\mathcal{X})$:

$$H(\mathcal{X}) \quad \leq \quad H(X_2).$$

(f) **True** or **False**? For a continuous random variable X, the uniform distribution $f(x) = \frac{1}{a}$ for $0 \leq x \leq a$ maximizes the differential entropy $H(X) = \int f(x) \ln f(x) dx$.

(g) Consider the random variable X:

$$\Pr(X = i) \quad = \quad \begin{cases} \frac{1}{4} & \text{if } \ i = -1 \\ \frac{1}{2} & \text{if } \ i = 0 \\ \frac{1}{4} & \text{if } \ i = 1 \end{cases}$$

Find $H(X)$. If $g(x) = x^2$, find $H(g(X))$ and $H(g(X)|X)$.

9. *Proof of Fano's Inequality* Write a justification (for example, "data processing inequality") for each step in the proof of Fano's inequality

   Fano's Inequality For any estimator $\widehat{X}$ such that $X \to Y \to \widehat{X}$, with event $E = \{X \neq \widehat{X}\}$ and with $P_e = \Pr(E)$, we have:

$$h(P_e) + P_e \log|\mathcal{X}| \quad \geq \quad H(X|\widehat{X}). \tag{34}$$

   Proof:

Justification

$$
\begin{aligned}
H(E, X|\widehat{X}) &= H(X|\widehat{X}) + H(E|X, \widehat{X}) && \text{(a) \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_} \\
&= H(X|\widehat{X}) && \text{(b) \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_} \\
H(E, X|\widehat{X}) &= H(E|\widehat{X}) + H(X|E, \widehat{X}) && \text{(c) \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_} \\
H(X|\widehat{X}) &= H(E|\widehat{X}) + H(X|E, \widehat{X}) && \text{equality of (b) and (c)} \\
&\leq H(E) + H(X|E, \widehat{X}) && \text{(d) \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_} \\
&= h(P_e) + H(X|E, \widehat{X}) && H(E) = h(P_e) \\
&= h(P_e) + H(X|\widehat{X}, E = 1)P_e && \\
&\quad + H(X|\widehat{X}, E = 0)(1 - P_e) && \text{(e) \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_} \\
&= h(P_e) + H(X|\widehat{X}, E = 1)P_e && \text{(f) \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_} \\
&\leq h(P_e) + P_e \log|\mathcal{X}| && \text{(g) \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_}
\end{aligned}
$$