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.

Solution: Expand the mutual information terms as:

$$\begin{aligned} I(\mathsf{X}_1;\mathsf{Y}|\mathsf{X}_2) &= H(\mathsf{Y}|\mathsf{X}_2) - H(\mathsf{Y}|\mathsf{X}_1\mathsf{X}_2) = H(\mathsf{Y}|\mathsf{X}_2) && (2)\\ I(\mathsf{X}_2;\mathsf{Y}|\mathsf{X}_1) &= H(\mathsf{Y}|\mathsf{X}_1) - H(\mathsf{Y}|\mathsf{X}_1\mathsf{X}_2) = H(\mathsf{Y}|\mathsf{X}_1) && (3)\\ I(\mathsf{X}_1,\mathsf{X}_2;\mathsf{Y}) &= H(\mathsf{Y}) - H(\mathsf{Y}|\mathsf{X}_1\mathsf{X}_2) = H(\mathsf{Y}) && (4) \end{aligned}$$

where the last equality follows because the channel is noiseless.

Now, observe that for this channel:

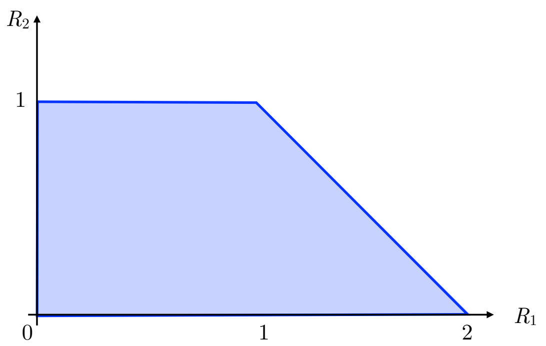
$$\begin{aligned} H(\mathsf{Y}|\mathsf{X}_1) &= H(\mathsf{X}_2) && (5)\\ H(\mathsf{Y}|\mathsf{X}_2) &= H(\mathsf{X}_1) && (6) \end{aligned}$$

so that:

$$\begin{aligned} R_1 &< H(\mathsf{X}_1) \le \log|\mathcal{X}_1| && (7)\\ R_2 &< H(\mathsf{X}_2) \le \log|\mathcal{X}_2| && (8)\\ R_1 + R_2 &< H(\mathsf{Y}) \le \log|\mathcal{Y}| && (9) \end{aligned}$$

where entropy upper bounds are shown. If we choose $p_{\mathsf{X}_1}(x_1) = [\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}]$ and $p_{\mathsf{X}_2}(x_2) = [\frac{1}{2}, \frac{1}{2}]$, then $p_{\mathsf{Y}}(y) = [\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}]$, and all three bounds are achieved with equality. So the rate region is:

$$\begin{aligned} R_1 &< 2 && (10)\\ R_2 &< 1 && (11)\\ R_1 + R_2 &< 2 && (12) \end{aligned}$$



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 $0 \le p, r \le 1$. Let $\mathsf{Y}_i = \mathsf{X}_i \oplus \mathsf{Z}_i$ where $\oplus$ denotes addition modulo 2. Let the source vector $\mathbf{X} = (\mathsf{X}_1, \ldots, \mathsf{X}_n)$ be encoded at rate R_1 and let $\mathbf{Y} = (\mathsf{Y}_1, \ldots, \mathsf{Y}_n)$ 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.

   Solution: No solution provided at this time.

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: $E[\mathsf{X}|\mathsf{Y}=y] = \rho_1 \frac{\sigma_\mathsf{X}^2}{\sigma_\mathsf{Y}^2} y$.)

   Solution:

$$
\begin{aligned}
E[\mathsf{X}|\mathsf{Y}=y] &= \rho_1 \frac{\sigma_\mathsf{X}}{\sigma_\mathsf{Y}} y && E[\mathsf{X}|\mathsf{Y}=y] \text{ is a function of } y \\
E[\mathsf{XZ}] &= E\big[E[\mathsf{XZ}|\mathsf{Y}]\big] && \text{by law of total expectation} \\
&= E\big[E[\mathsf{X}|\mathsf{Y}]E[\mathsf{Z}|\mathsf{Y}]\big] && \text{by conditional independence of } \mathsf{X} \text{ and } \mathsf{Z} \\
&= \int \left(\rho_1 \frac{\sigma_\mathsf{X}}{\sigma_\mathsf{Y}} y\right)\left(\rho_2 \frac{\sigma_\mathsf{Z}}{\sigma_\mathsf{Y}} y\right) p_\mathsf{Y}(y) dy \\
&= \rho_1 \frac{\sigma_\mathsf{X}}{\sigma_\mathsf{Y}} \rho_2 \frac{\sigma_\mathsf{Z}}{\sigma_\mathsf{Y}} \underbrace{\int y^2 p_\mathsf{Y}(y) dy}_{=\sigma_\mathsf{Y}^2} \\
&= \rho_1 \rho_2 \sigma_\mathsf{X} \sigma_\mathsf{Z}
\end{aligned}
$$

   2. Find $I(\mathsf{X};\mathsf{Z})$.

   Solution: Covariance matrix:

$$
\begin{aligned}
K &= \begin{bmatrix} \sigma_\mathsf{X}^2 & \rho_1\rho_2\sigma_\mathsf{X}\sigma_\mathsf{Z} \\ \rho_1\rho_2\sigma_\mathsf{X}\sigma_\mathsf{Z} & \sigma_\mathsf{Z}^2 \end{bmatrix} && (13) \\
\det K &= \sigma_\mathsf{X}^2\sigma_\mathsf{Z}^2 - \rho_1\rho_2\sigma_\mathsf{X}^2\sigma_\mathsf{Z}^2 = \sigma_\mathsf{X}^2\sigma_\mathsf{Z}^2(1-\rho_1^2\rho_2^2) && (14) \\
I(\mathsf{X};\mathsf{Z}) &= H(\mathsf{X}) + H(\mathsf{Y}) - H(\mathsf{X},\mathsf{Y}) && (15) \\
&= \frac{1}{2}\ln 2\pi e\sigma_\mathsf{X}^2 + \frac{1}{2}\ln 2\pi e\sigma_\mathsf{Z}^2 - \frac{1}{2}\ln(2\pi e)^2\sigma_\mathsf{X}^2\sigma_\mathsf{Z}^2(1-\rho_1^2\rho_2^2) && (16) \\
&= -\frac{1}{2}\ln(1-\rho_1^2\rho_2^2) && (17)
\end{aligned}
$$

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 $I(\mathsf{X};\mathsf{Y})=H(\mathsf{X})-H(\mathsf{X}|\mathsf{Y})$.

Solution:

$$f_{\mathsf{Y}}(y|\mathsf{X}=0)=\begin{cases}\frac{1}{a} & if\ 0\le y\le a\\ 0 & otherwise\end{cases} \tag{18}$$

and

$$f_{\mathsf{Y}}(y|\mathsf{X}=1)=\begin{cases}\frac{1}{a} & if\ 1\le y\le a+1\\ 0 & otherwise\end{cases} \tag{19}$$

Therefore

$$f_{\mathsf{Y}}(y)=\begin{cases}(1-p)\frac{1}{a} & if\ 0\le y<1\\ \frac{1}{a} & if\ 1\le y\le a\\ p\frac{1}{a} & if\ a<y<1+a\end{cases} \tag{20}$$

$H(\mathsf{X})=h(p)$ and $H(\mathsf{X}|\mathsf{Y})$ is not 0 for $1\le y\le a$

$$H(\mathsf{X}|\mathsf{Y})=P(1\le y\le a)h(p)=\frac{a-1}{a}h(p) \tag{21}$$

Mutual information is:

$$I(\mathsf{X};\mathsf{Y})=H(\mathsf{X})-H(\mathsf{X}|\mathsf{Y})=\frac{1}{a}h(p) \tag{22}$$

   2. Now calculate $I(\mathsf{X};\mathsf{Y})$ the other way by $I(\mathsf{X};\mathsf{Y})=H(\mathsf{Y})-H(\mathsf{Y}|\mathsf{X})$.

Solution:

$$H(\mathsf{Y}|\mathsf{X}=0)=-\int_0^a\frac{1}{a}\log\frac{1}{a}dy=\log a \tag{23}$$

$$H(\mathsf{Y}|\mathsf{X}=1)=-\int_1^{a+1}\frac{1}{a}\log\frac{1}{a}dy=\log a \tag{24}$$

$$H(\mathsf{Y}|\mathsf{X})=P(\mathsf{X}=0)H(\mathsf{Y}|\mathsf{X}=0)+P(\mathsf{X}=1)H(\mathsf{Y}|\mathsf{X}=1)=\log a \tag{25}$$

$$\begin{aligned}H(\mathsf{Y})&=-\int_0^1(1-p)\frac{1}{a}\log(1-p)\frac{1}{a}dy-\int_1^a\frac{1}{a}\log\frac{1}{a}dy-\int_a^{a+1}p\frac{1}{a}\log p\frac{1}{a}dy && (26)\\ &=\frac{1}{a}h(p)+\log a && (27)\end{aligned}$$

Mutual information is:

$$I(\mathsf{X};\mathsf{Y})=H(\mathsf{Y})-H(\mathsf{Y}|\mathsf{X})=\frac{1}{a}h(p) \tag{28}$$

   3. Calculate the capacity of this channel by maximizing over p.

Solution: mutual information is $\frac{1}{a}h(p)$ and is maximum when p has uniform distribution:

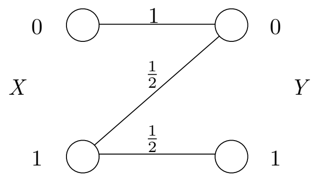
$$p=\frac{1}{2} \tag{29}$$

$$C=\max I(X;Y)=\frac{1}{a}h(p)=\frac{1}{a} \tag{30}$$

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:

$$\begin{bmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}. \tag{31}$$

The channel diagram looks like a "Z":



The channel is memoryless:

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

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 $p_{\mathbf{Y}}(\mathbf{y})$.

Solution: $\Pr(Y = 1) = \frac{p}{2} = \frac{1}{4}$. So:

$$p_{\mathbf{Y}}(\mathbf{y}) \quad = \quad \left(\frac{1}{4}\right)^m \left(\frac{3}{4}\right)^{n-m}$$

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

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

Solution:

$$\begin{aligned}
-\frac{1}{n} \log p(y^n) &= H(Y) \\
-\frac{1}{n}\Big(\log(\frac{1}{4})^m(\frac{3}{4})^{n-m}\Big) &= -(\frac{1}{4}\log\frac{1}{4} + \frac{3}{4}\log\frac{3}{4}) \\
\frac{1}{n}\Big(m\log(\frac{1}{4}) + (n-m)\log(\frac{3}{4})\Big) &= \log\frac{1}{4} + \frac{3}{4}\log 3 \\
\frac{1}{n}\Big(n\log(\frac{1}{4}) + (n-m)\log 3\Big) &= \log\frac{1}{4} + \frac{3}{4}\log 3 \\
\log(\frac{1}{4}) + \frac{n-m}{n}\log 3 &= \log\frac{1}{4} + \frac{3}{4}\log 3 \\
\frac{n-m}{n} &= \frac{3}{4} \\
m &= \frac{1}{4}n
\end{aligned}$$

That is, the "most typical" sequence has 1/4 ones and 3/4 zeros.

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

- $p_{\mathsf{Y}|\mathsf{X}}(000|001) =$ $\underline{p(0|0)p(0|0)p(0|1) = 1 \cdot 1 \cdot \frac{1}{2} = \frac{1}{2}}$

- $p_{\mathsf{Y|X}}(001|001) =$ $\underline{p(0|0)p(0|0)p(1|1) = 1 \cdot 1 \cdot \frac{1}{2} = \frac{1}{2}}$
- $p_{\mathsf{Y|X}}(010|001) =$ $\underline{p(0|0)p(1|0)p(0|1) = 1 \cdot 0 \cdot \frac{1}{2} = 0}$
- $p_{\mathsf{Y|X}}(011|001) =$ $\underline{p(0|0)p(1|0)p(1|1) = 1 \cdot 0 \cdot \frac{1}{2} = 0}$

Continued on next page

Let $\mathbf{x} = (x_1, \ldots, x_n)$ 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 $i = 1, 2, \ldots, n$). Let $\mathcal{V}$ be the valid sequences:

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

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 $\mathbf{x}$.

Solution:

$$p(y^n|x^n, V) \quad = \quad \prod_{i=1}^{n} p(y_i|x_i, \{x_i = 1 \ \text{ or } \ y_i = 0\})$$

Any sequence in V can be written as:

$$\begin{aligned} p(y^n|x^n, V) &= \prod_{(x_i,y_i)=(0,0)} p(y_i|x_i) \cdot \prod_{(x_i,y_i)=(1,0)\text{or}(1,1)} p(y_i|x_i) \\ &= \prod_{(x_i,y_i)=(0,0)} 1 \cdot \prod_{(x_i,y_i)=(1,0)\text{or}(1,1)} (\frac{1}{2}) \end{aligned}$$

Since there are k ones in x^n:

$$= \quad 1 \cdot (\frac{1}{2})^k = \frac{1}{2^k}$$

(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.

Solution:

$$\begin{aligned} p(x^n, y^n) &= \prod p(y_i|x_i)p(x_i) \\ &= \begin{cases} \frac{1}{2^k}\frac{1}{2^n} & \text{if } (x^n, y^n) \in V \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

Then,

$$\begin{aligned} -\frac{1}{n}\log p(x^n, y^n) &= -\frac{1}{n}\log 2^{-n-k} \\ &= \frac{n+k}{n} \end{aligned}$$

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

Solution:

$$\begin{aligned} H(X,Y) &= -\sum_x \sum_y p(x,y)\log p(x,y) \\ &= -(\frac{1}{2}\log\frac{1}{2} + \frac{1}{4}\log\frac{1}{4} + \frac{1}{4}\log\frac{1}{4}) \\ &= 1.5 \text{ bits} \end{aligned}$$

(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 $\mathcal{T}'_\epsilon$.

Solution:

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.

> **Solution:** $I(X;Y) \geq I(X;Z)$

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

> **Solution:** $H(X;Y) \leq H(X) + H(Y)$, with equality iff X and Y are independent.

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

> **Solution:** $H(X;Y) = H(X) + H(Y|X)$

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

> **Solution:** $H(X|Y) \leq H(X)$

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

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

> **Solution:** True. $H(\mathcal{X}) \leq H(X_2|X_1)$ and $H(X_2|X_1) \leq H(X_2)$.

(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 $H(X) = \int f(x) \ln f(x) dx$.

> **Solution:** False. A Gaussian (normal) maximizes the differential entropy, a uniform distribution maximizes the (discrete) entropy.

(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 $H(g(X)|X)$.

> **Solution:**
>
> $$H(X) \quad = \quad -\Big(\frac{1}{4}\log\frac{1}{4} + \frac{1}{2}\log\frac{1}{2} + \frac{1}{4}\log\frac{1}{4}\Big) = \Big(\frac{1}{4}\cdot 2 + \frac{1}{2}\cdot 1 + \frac{1}{4}\cdot 2\Big) = 1.5 \text{ bits}$$
>
> $Y = g(X), \Pr(Y=0) = \frac{1}{2}, \Pr(Y=1) = \frac{1}{4} + \frac{1}{4}$. So, $H(Y) = H(g(X)) = h(\frac{1}{2}) = 1$.
> $H(g(X)|X) = 0$, since X tells us $g(X)$ exactly.

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:

$$h(P_e) + P_e \log |\mathcal{X}| \quad \geq \quad H(X|\widehat{X}). \tag{33}$$

   Proof:

Justification

$$\begin{aligned}
H(E, X|\widehat{X}) &= H(X|\widehat{X}) + H(E|X, \widehat{X}) && \text{(a) \underline{ Entropy chain rule }} \\
&= H(X|\widehat{X}) && \text{(b) \underline{ $E = \{X \neq wX\} \to$ Conditionally, E is known exactly }} \\
H(E, X|\widehat{X}) &= H(E|\widehat{X}) + H(X|E, \widehat{X}) && \text{(c) \underline{ Entropy chain rule }} \\
H(X|\widehat{X}) &= H(E|\widehat{X}) + H(X|E, \widehat{X}) && \text{equality of (b) and (c)} \\
&\leq H(E) + H(X|E, \widehat{X}) && \text{(d) \underline{ Conditioning reduces entropy }} \\
&= h(P_e) + H(X|E, \widehat{X}) && H(E) = h(P_e) \\
&= h(P_e) + H(X|\widehat{X}, E = 1)P_e && \\
&\quad +H(X|\widehat{X}, E = 0)(1 - P_e) && \text{(e) \underline{ Definition of conditional entropy }} \\
&= h(P_e) + H(X|\widehat{X}, E = 1)P_e && \text{(f) \underline{ entropy of X is 0 , given no errors and $\widehat{X}$ known }} \\
&\leq h(P_e) + P_e \log |\mathcal{X}| && \text{(g) \underline{ uniform distribution upper bound on entropy }}
\end{aligned}$$