Homework Set #6
Source channel separation, Max entropy principle, and channel coding with side information

1. Fading channel.
   Consider an additive noise fading channel
   \[ \begin{align*}
   &V \\
   &Z \\
   &\sum \\
   &Y = XV + Z,
   \end{align*} \]
   where \( Z \) is additive noise, \( V \) is a random variable representing fading, and \( Z \) and \( V \) are independent of each other and of \( X \).
   (a) Argue that knowledge of the fading factor \( V \) improves capacity by showing
   \[ I(X;Y|V) \geq I(X;Y). \]
   (b) Incidentally, conditioning does not always increase mutual information. Give an example of \( p(u,r,s) \) such that \( I(U;R|S) < I(U;R) \).

2. Diversity System
   For the following system, a message \( W \in \{1,2,\ldots,2^{nR}\} \) is encoded into two symbol blocks \( X^n_1 = (X_{1,1},X_{1,2},\ldots,X_{1,n}) \) and \( X^n_2 = (X_{2,1},X_{2,2},\ldots,X_{2,n}) \) that are being transmitted over a channel. The average power constrain on the inputs are \( \frac{1}{n} E[\sum_{i=1}^{n} X^2_{1,i}] \leq P_1 \) and \( \frac{1}{n} E[\sum_{i=1}^{n} X^2_{2,i}] \leq P_2 \). The channel has a multiplying effect on \( X_1, X_2 \) by factor \( h_1, h_2 \), respectively, i.e., \( Y = h_1X_1 + h_2X_2 + Z \), where \( Z \) is a white Gaussian noise \( Z \sim N(0,\sigma^2) \).
   (a) Find the joint distribution of \( X_1 \) and \( X_2 \) that bring the mutual information \( I(Y;X_1,X_2) \) to a maximum? (You need to find \( \arg\max P_{X_1,X_2} I(X_1,X_2;Y) \)).
   (b) What is the capacity of the system?
(c) Express the capacity for the following cases:

i. $h_1 = 1, h_2 = 1$?

ii. $h_1 = 1, h_2 = 0$?

iii. $h_1 = 0, h_2 = 0$?

3. **AWGN with two noises** (15 points)

Figure 2 depicts a communication system with an AWGN (Additive white noise Gaussian) channel with two i.i.d. noises $Z_1 \sim N(0, \sigma_1^2)$, $Z_2 \sim N(0, \sigma_2^2)$ that are independent of each other and are added to the signal $X$, i.e., $Y = X + Z_1 + Z_2$. The average power constrain on the input is $P$, i.e., $\frac{1}{n}E[\sum_{i=1}^{n} X_i^2] \leq P$. In the sub-questions below we consider the cases where the noise $Z_2$ may or may not be known to the encoder and decoder.
(a) Find the channel capacity for the case in which the noise in not known to either sides (lines 1 and 2 are disconnected from the encoder and the decoder).

(b) Find the capacity for the case that the noise $Z_2$ is known to the encoder and decoder (lines 1 and 2 are connected to both the encoder and decoder). This means that the codeword $X^n$ may depend on the message $W$ and the noise $Z_2^n$ and the decoder decision $\hat{W}$ may depend on the output $Y^n$ and the noise $Z_2^n$. (Hint: Could the capacity be larger than $\frac{1}{2} \log(1 + \frac{P}{\sigma^2})$?)

(c) Find the capacity for the case that the noise $Z_2$ is known only to the decoder. (line 1 is disconnected from the encoder and line 2 is connected to the decoder). This means that the codewords $X^n$ may depend only on the message $W$ and the decoder decision $\hat{W}$ may depend on the output $Y^n$ and the noise $Z_2^n$.

4. Source and channel. (Please read the relevant lecture on source channel separation)
We wish to encode a Bernoulli($\alpha$) process $V_1, V_2, \ldots$ for transmission over a binary symmetric channel with error probability $p$.

\[
\begin{array}{ccccccc}
V^n & \rightarrow & X^n(V^n) & \rightarrow & 1 & \xrightarrow{p} & 1 & \rightarrow & Y^n & \rightarrow & \hat{V}^n \\
011011101 & & & & 0 & \xrightarrow{p} & 0 & & 011011101 \\
\end{array}
\]

Find conditions on $\alpha$ and $p$ so that the probability of error $P(\hat{V}^n \neq V^n)$ can be made to go to zero as $n \rightarrow \infty$.

5. Maximum entropy.
Find the maximum entropy density $f$ satisfying $EX = \alpha_1, E \ln X = \alpha_2$. That is,

\[
\text{maximize } h(f)
\]

subject to $\int x f(x) \, dx = \alpha_1, \int (\ln x) f(x) \, dx = \alpha_2$. What family of densities is this?
6. **Minimum relative entropy** $D(P \parallel Q)$ **under constraints on** $P$.

We wish to find the (parametric form) of the probability mass function $P(x), x \in \{1, 2, \ldots\}$ that minimizes the relative entropy $D(P \parallel Q)$ over all $P$ such that $\sum P(x)g_i(x) = \alpha_i, i = 1, 2, \ldots$.

(a) Use Lagrange multipliers to guess that $P^*(x) = Q(x)e^{\sum_{i=1}^{\infty} \lambda_i g_i(x) + \lambda_0}$ achieves this minimum if there exist $\lambda_i$’s satisfying the $\alpha_i$ constraints. This generalizes the theorem on maximum entropy distributions subject to constraints.

(b) Verify that $P^*$ minimizes $D(P \parallel Q)$.

7. **Maximum entropy with marginals.**

What is the maximum entropy probability mass function $p(x, y)$ with the following marginals? You may wish to guess and verify a more general result.

<table>
<thead>
<tr>
<th></th>
<th>$y_1$</th>
<th>$y_2$</th>
<th>$y_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$</td>
<td>$p_{11}$</td>
<td>$p_{12}$</td>
<td>$p_{13}$</td>
</tr>
<tr>
<td>$x_2$</td>
<td>$p_{21}$</td>
<td>$p_{22}$</td>
<td>$p_{23}$</td>
</tr>
<tr>
<td>$x_3$</td>
<td>$p_{31}$</td>
<td>$p_{32}$</td>
<td>$p_{33}$</td>
</tr>
</tbody>
</table>

$2/3 \quad 1/6 \quad 1/6$