Problem Solving Session 6

28) Choice of channels.

Find the capacity $C$ of the union of 2 channels $(X_1, P_1(y_1|x_1), Y_1)$ and $(X_2, P_2(y_2|x_2), Y_2)$ where, at each time, one can send a symbol over channel 1 or over channel 2 but not both. Assume the output alphabets are distinct and do not intersect.

a) Show $2^C = 2^{C_1} + 2^{C_2}$. Thus $2^C$ is the effective alphabet size of a channel with capacity $C$.

28) Choice of Channels

(a) This is solved by using the very same trick that was used to solve problem 2.10.

Consider the following communication scheme:

$$X = \begin{cases} 
X_1 & \text{Probability } \alpha \\
X_2 & \text{Probability } (1 - \alpha)
\end{cases}$$

Let

$$\theta(X) = \begin{cases} 
1 & X = X_1 \\
2 & X = X_2
\end{cases}$$

Since the output alphabets $Y_1$ and $Y_2$ are disjoint, $\theta$ is a function of $Y$ as well, i.e. $X \rightarrow Y \rightarrow \theta$.

$$I(X;Y,\theta) = I(X;\theta) + I(X;Y|\theta)$$
$$= I(X;Y) + I(X;\theta|Y)$$

Since $X \rightarrow Y \rightarrow \theta$, $I(X;\theta|Y) = 0$. Therefore,

$$I(X;Y) = I(X;\theta) + I(X;Y|\theta)$$
$$= H(\theta) - H(\theta|X) + \alpha I(X_1;Y_1) + (1 - \alpha)I(X_2;Y_2)$$
$$= H(\alpha) + \alpha I(X_1;Y_1) + (1 - \alpha)I(X_2;Y_2)$$

Thus, it follows that

$$C = \sup_\alpha \{ H(\alpha) + \alpha C_1 + (1 - \alpha)C_2 \}.$$  

Maximizing over $\alpha$ one gets the desired result. The maximum occurs for $H'(\alpha) + C_1 - C_2 = 0$, or $\alpha = 2^C/(2^{C_1} + 2^{C_2})$.

7.33

33) BSC with feedback. Suppose that feedback is used on a binary symmetric channel with parameter $p$. Each time a $Y$ is received, it becomes the next transmission. Thus $X_1$ is Bern(1/2), $X_2 = Y_1$, $X_3 = Y_2$, ..., $X_n = Y_{n-1}$.

a) Find $\lim_{n \to \infty} \frac{1}{n} I(X^n;Y^n)$.

b) Show that for some values of $p$, this can be higher than capacity.

33) BSC with feedback solution.

a)

$$I(X^n;Y^n) = H(Y^n) - H(Y^n|X^n).$$

$$H(Y^n|X^n) = \sum_i H(Y_i|Y^{i-1}, X^n) = H(Y_1|X_1) + \sum_i H(Y_i|Y^n) = H(p) + 0.$$  

$$H(Y^n) = \sum_i H(Y_i|Y^{i-1}) = H(Y_1) + \sum_i H(Y_i|X_i) = 1 + (n - 1)H(p)$$

So,
\[ I(X^n;Y^n) = 1 + (n-1)H(p) - H(p) = 1 + (n-2)H(p) \]

and,

\[ \lim_{n \to \infty} \frac{1}{n} I(X^n;Y^n) = \lim_{n \to \infty} \frac{1 + (n-2)H(p)}{n} = H(p) \]

b) For the BSC \( C = 1 - H(p) \). For \( p = 0.5 \), \( C = 0 \), while \( \lim_{n \to \infty} \frac{1}{n} I(X^n;Y^n) = H(0.5) = 1 \).

### 7.35

**Capacity.**

Suppose channel \( \mathcal{P} \) has capacity \( C \), where \( \mathcal{P} \) is an \( m \times n \) channel matrix.

a) What is the capacity of

\[ \bar{\mathcal{P}} = \begin{bmatrix} \mathcal{P} & 0 \\ 0 & 1 \end{bmatrix} \]

b) What about the capacity of

\[ \hat{\mathcal{P}} = \begin{bmatrix} \mathcal{P} & 0 \\ 0 & I_k \end{bmatrix} \]

where \( I_k \) is the \( k \times k \) identity matrix.

### 35) Solution: Capacity.

a) By adding the extra column and row to the transition matrix, we have two channels in parallel. You can transmit on either channel. From problem 7.28, it follows that

\[ \hat{C} = \log(2^9 + 2^C) \]

\[ \hat{C} = \log(1 + 2^C) \]

b) This part is also an application of the conclusion problem 7.28. Here the capacity of the added channel is \( \log k \).

\[ \hat{C} = \log(2^{\log k} + 2^C) \]

\[ \hat{C} = \log(k + 2^C) \]

### 8.10

**The Shape of the Typical Set**

Let \( X_i \) be i.i.d. \( \sim f(x) \), where

\[ f(x) = ce^{-x^4} \]

Let \( h = -\int f \ln f \). Describe the shape (or form) or the typical set \( A_x^{(n)} = \{ x^n \in \mathcal{R}^n : f(x^n) \in 2^{-n(h+\epsilon)} \} \).
10) \textbf{The Shape of the Typical Set}

We are interested in the set \( \{x^n \in \mathbb{R} : f(x^n) \in 2^{-n(h+\epsilon)} \} \). This is:

\[
2^{-n(h-\epsilon)} \leq f(x^n) \leq 2^{-n(h+\epsilon)}
\]

Since \( X_i \) are i.i.d.,

\[
f(x^n) = \prod_{i=1}^{n} f(x) = \prod_{i=1}^{n} c e^{-x_i^4} = e^{n \ln(c) - \sum_{i=1}^{n} x_i^4} \]

(734)
(735)
(736)
(737)

Plugging this in for \( f(x^n) \) in the above inequality and using algebraic manipulation gives:

\[
n(\ln(c) + (h-\epsilon) \ln(2)) \geq \sum_{i=1}^{n} x_i^4 \geq n(\ln(c) + (h+\epsilon) \ln(2))
\]

So the shape of the typical set is the shell of a \( 4 \)-norm ball \( \{x^n : ||x^n||_4 \in (n(\ln(c) + (h \pm \epsilon) \ln(2)))^{1/4} \} \).

8.11

11) \textbf{Non ergodic Gaussian process.}

Consider a constant signal \( V \) in the presence of iid observational noise \( \{Z_i\} \). Thus \( X_i = V + Z_i \), where \( V \sim N(0, S) \), and \( Z_i \) are iid \( \sim N(0, N) \). Assume \( V \) and \( \{Z_i\} \) are independent.

a) Is \( \{X_i\} \) stationary?

b) Find \( \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} X_i \). Is the limit random?

11) \textbf{Nonergodic Gaussian process}

a) Yes. \( EX_i = EV + Z_i = 0 \) for all \( i \), and

\[
EX_iX_j = E(V + Z_i)(V + Z_j) = \begin{cases} S, & i = j \\ S + N, & i \neq j \end{cases}
\]

(738)

Since \( X_i \) is Gaussian distributed it is completely characterized by its first and second moments. Since the moments are stationary, \( X_i \) is wide sense stationary, which for a Gaussian distribution implies that \( X_i \) is stationary.

b)

\[
\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} X_i = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} (Z_i + V)
\]

(739)
\[
= V + \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} Z_i
\]

(740)
\[
= V + EZ_i (\text{by the strong law of large numbers})
\]

(741)
\[
= V
\]

(742)

The limit is a random variable \( \mathcal{N}(0, S) \).