Problem Solving Session 5

7.2

2) An additive noise channel. Find the channel capacity of the following discrete memoryless channel:

\[ Z \]

\[ X \rightarrow Y \]

where \( \Pr\{Z = 0\} = \Pr\{Z = a\} = \frac{1}{2} \). The alphabet for \( x \) is \( X = \{0, 1\} \). Assume that \( Z \) is independent of \( X \). Observe that the channel capacity depends on the value of \( a \).

2) A sum channel.

\[ Y = X + Z \quad X \in \{0, 1\}, \quad Z \in \{0, a\} \]

We have to distinguish various cases depending on the values of \( a \).

- \( a = 0 \) In this case, \( Y = X \), and \( \max I(X; Y) = \max H(X) = 1 \). Hence the capacity is 1 bit per transmission.

- \( a \neq 0 \) In this case, \( Y \) has four possible values 0, 1, \( a \) and \( 1 + a \). Knowing \( Y \), we know the \( X \) which was sent, and hence \( H(X|Y) = 0 \). Hence \( \max I(X; Y) = \max H(X) = 1 \), achieved for a uniform distribution on the input \( X \).

- \( a = 1 \) In this case \( Y \) has three possible output values, 0, 1 and 2, and the channel is identical to the binary erasure channel discussed in class, with \( a = 1/2 \). As derived in class, the capacity of this channel is \( 1 - a = 1/2 \) bit per transmission.

- \( a = -1 \) This is similar to the case when \( a = 1 \) and the capacity here is also 1/2 bit per transmission.

7.3

3) Channels with memory have higher capacity. Consider a binary symmetric channel with \( Y_i = X_i \oplus Z_i \), where \( \oplus \) is mod 2 addition, and \( X_i, Y_i \in \{0, 1\} \).

Suppose that \( \{Z_i\} \) has constant marginal probabilities \( \Pr\{Z_i = 1\} = p = 1 - \Pr\{Z_i = 0\} \), but that \( Z_1, Z_2, \ldots, Z_n \) are not necessarily independent. Assume that \( Z^n \) is independent of the input \( X^n \). Let \( C = 1 - H(p, 1-p) \). Show that \( \max_{p(x_1, x_2, \ldots, x_n)} I(X_1, X_2, \ldots, X_n; Y_1, Y_2, \ldots, Y_n) \geq nC \).
3) **Channels with memory have a higher capacity.**

\[ Y_i = X_i \oplus Z_i, \]

where

\[ Z_i = \begin{cases} 
1 & \text{with probability } p \\
0 & \text{with probability } 1 - p 
\end{cases} \]

and \( Z_i \) are not independent.

\[
I(X_1, X_2, \ldots, X_n; Y_1, Y_2, \ldots, Y_n) \\
= H(X_1, X_2, \ldots, X_n) - H(X_1, X_2, \ldots, X_n|Y_1, Y_2, \ldots, Y_n) \\
= H(X_1, X_2, \ldots, X_n) - H(Z_1, Z_2, \ldots, Z_n|Y_1, Y_2, \ldots, Y_n) \\
\geq H(X_1, X_2, \ldots, X_n) - H(Z_1, Z_2, \ldots, Z_n) \\
\geq H(X_1, X_2, \ldots, X_n) - \sum H(Z_i) \\
= H(X_1, X_2, \ldots, X_n) - nH(p) \\
= n - nH(p),
\]

if \( X_1, X_2, \ldots, X_n \) are chosen i.i.d. \( \sim \) Bern(\( \frac{1}{2} \)). The capacity of the channel with memory over \( n \) uses of the channel is

\[
nC^{(n)} = \max_{p(x_1, x_2, \ldots, x_n)} I(X_1, X_2, \ldots, X_n; Y_1, Y_2, \ldots, Y_n) \\
\geq I(X_1, X_2, \ldots, X_n; Y_1, Y_2, \ldots, Y_n)_{p(x_1, x_2, \ldots, x_n) = \text{Bern}(\frac{1}{2})} \\
\geq n(1 - H(p)) \\
= nC.
\]

Hence channels with memory have higher capacity. The intuitive explanation for this result is that the correlation between the noise decreases the effective noise; one could use the information from the past samples of the noise to combat the present noise.

7.4

4) **Channel capacity.** Consider the discrete memoryless channel \( Y = X + Z \) (mod 11), where

\[ Z = \begin{pmatrix} 
1, & 2, & 3 \\
1/3, & 1/3, & 1/3 
\end{pmatrix} \]

and \( X \in \{0, 1, \ldots, 10\} \). Assume that \( Z \) is independent of \( X \).

a) Find the capacity.

b) What is the maximizing \( p^*(x) \)?

4) **Channel capacity.**

\[ Y = X + Z \text{ (mod 11)} \]

where

\[ Z = \begin{cases} 
1 & \text{with probability } 1/3 \\
2 & \text{with probability } 1/3 \\
3 & \text{with probability } 1/3 
\end{cases} \]

In this case,

\[ H(Y|X) = H(Z|X) = H(Z) = \log 3, \]
independent of the distribution of $X$, and hence the capacity of the channel is

$$
C = \max_{p(x)} I(X;Y) \\
= \max_{p(x)} H(Y) - H(Y|X) \\
= \max_{p(x)} H(Y) - \log 3 \\
= \log 11 - \log 3,
$$

which is attained when $Y$ has a uniform distribution, which occurs (by symmetry) when $X$ has a uniform distribution.

a) The capacity of the channel is $\log \frac{11}{3}$ bits/transmission.

b) The capacity is achieved by an uniform distribution on the inputs. $p(X = i) = \frac{1}{11}$ for $i = 0, 1, \ldots, 10$. 