Problem 1. 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 independent. Assume that \((Z_1, \ldots, Z_n)\) is independent of the input \((X_1, \ldots, X_n)\). Let \( C = 1 - H(p, 1-p) \). Show that

\[
\max_{p_{X_1, X_2, \ldots, X_n}} I(X_1, X_2, \ldots, X_n; Y_1, Y_2, \ldots, Y_n) > nC.
\]

Problem 2. Consider two discrete memoryless channels. The input alphabet, output alphabet, transition probabilities and capacity of the \( k \)'th channel is given by \( X_k, Y_k, p_k \) and \( C_k \) respectively \((k = 1, 2)\). The channels operate independently. A communication system has access to both channels, that is, the effective channel between the transmitter and receiver has input alphabet \( X_1 \times X_2 \), output alphabet \( Y_1 \times Y_2 \) and transition probabilities \( p_1(y_1|x_1)p_2(y_2|x_2) \). Find the capacity of this channel.

Problem 3. Show that a cascade of \( n \) identical binary symmetric channels,

\[
X_0 \rightarrow \text{BSC} \#1 \rightarrow X_1 \rightarrow \cdots \rightarrow X_{n-1} \rightarrow \text{BSC} \#n \rightarrow X_n
\]

each with raw error probability \( p \), is equivalent to a single BSC with error probability \( \frac{1}{2}(1 - (1 - 2p)^n) \) and hence that \( \lim_{n \to \infty} I(X_0; X_n) = 0 \) if \( p \neq 0, 1 \). Thus, if no processing is allowed at the intermediate terminals, the capacity of the cascade tends to zero.

Problem 4. Consider a memoryless channel with transition probability matrix \( P_{Y|X}(y|x) \), with \( x \in \mathcal{X} \) and \( y \in \mathcal{Y} \). For a distribution \( Q \) over \( \mathcal{X} \), let \( I(Q) \) denote the mutual information between the input and the output of the channel when the input distribution is \( Q \). Show that for any two distributions \( Q \) and \( Q' \) over \( \mathcal{X} \),

(a)

\[
I(Q') \leq \sum_{x \in \mathcal{X}} Q'(x) \sum_{y \in \mathcal{Y}} P_{Y|X}(y|x) \log \left( \frac{P_{Y|X}(y|x)}{\sum_{x' \in \mathcal{X}} P_{Y|X}(y|x') Q(x')} \right)
\]

(b)

\[
C \leq \max_x \sum_{y \in \mathcal{Y}} P_{Y|X}(y|x) \log \left( \frac{P_{Y|X}(y|x)}{\sum_{x' \in \mathcal{X}} P_{Y|X}(y|x') Q(x')} \right)
\]

where \( C \) is the capacity of the channel. Notice that this upper bound to the capacity is independent of the maximizing distribution.