1. 20 points Consider the following $k$ input and $k+1$ output discrete erasure channel:

(a) For a given input distribution $p(x)$, what is the mutual information $I(X;Y)$? (Express your answer in terms of $H(X)$)

(b) Define the random variable $E$ as

$$E = \begin{cases} 1 & Y = e \\ 0 & \text{otherwise} \end{cases}$$

What are $H(Y,E)$ and $H(Y,E|X)$?

(c) Suppose an arbitrary $j$ input, $k$ output channel from $W$ to $X$ is followed in cascade by the $X$, $Y$ erasure channel from part (a) as follows:

What is $I(W;Y)$? Your answer should be expressed in terms of $I(W;X)$. Hint: consider the auxiliary random variable $E$.

2. 20 points Consider a channel consisting of two parallel AWGN channels with inputs $X_1, X_2$ and outputs

$$Y_1 = X_1 + Z_1$$
$$Y_2 = X_2 + Z_2$$

The noises $Z_1$ and $Z_2$ are independent and have variances $N_1$ and $N_2$ with $N_1 < N_2$. However, we are constrained to use the same symbol on both channels, i.e. $X_1 = X_2 = X$, where $X$ is constrained to have power $E[X^2] = P$. 
(a) Suppose at the receiver, we combine the outputs to produce \( Y = Y_1 + Y_2 \)? What is the capacity \( C_1 \) of channel with input \( X \) and output \( Y \)? What type of signaling achieves this capacity?

(b) Suppose the receiver can view both outputs \( Y_1 \) and \( Y_2 \). What is the capacity \( C_2 \) of this system? Does the optimal signaling change from part (a)?

(c) Suppose the receiver must combine the two received signals to produce \( Y' = \alpha Y_1 + (1 - \alpha) Y_2 \) where \( 0 \leq \alpha \leq 1 \). However, as the receiver designer, you can choose the best \( \alpha \) for combining. What is the capacity \( C' \) of this system with input \( X \) and output \( Y' \)? Is there a loss in capacity relative to \( C_2 \)?

(d) Suppose the transmitter, is still constrained to transmit the same signal on both channels, but can choose how much power to use on each channel. That is, for constants \( a \) and \( b \), \( X_1 = aX \) and \( X_2 = bX \). Subject to a constraint that the total transmitted power is bounded by \( 2P \), what are the optimal \( a \) and \( b \) and corresponding capacity \( C'' \) of the system with outputs \( (Y_1, Y_2) \)?

3. Consider the binary symmetric channel and the binary erasure channel shown below:

\[
\begin{array}{ccc}
X=0 & \rightarrow & Y=0 \\
X=1 & \rightarrow & Y=1 \\
\end{array}
\]

BSC

\[
\begin{array}{ccc}
X=0 & \rightarrow & Y=0 \\
X=1 & \rightarrow & Y=e \\
Y=1 & \rightarrow & Y=1 \\
\end{array}
\]

BEC

(a) Find the capacity \( C_{BSC}(\epsilon) \) of the BSC and \( C_{E}(\delta) \) of the erasure channel.

(b) When \( \delta = \epsilon \), use the data processing theorem to prove that the BEC has higher capacity than the BSC.

(c) Consider the Z-channel:

\[
\begin{array}{ccc}
X=0 & \rightarrow & Y=0 \\
X=1 & \rightarrow & Y=1 \\
\end{array}
\]

Use the data processing theorem to find an upper bound and a lower bound to the capacity \( C_Z(\alpha) \) of the \( z \) channel. Express these bounds in terms of \( C_{BSC}(\cdot) \) and \( C_{E}(\cdot) \).

4. Suppose we have a wireless network with \( n \) hops. For \( i = 0, 1, \ldots, n - 1 \), node \( i \) transmits coded messages at rate \( R_i \) to node \( i + 1 \) over an AWGN channel with noise variance \( N_i \):

Assume each node transmits in an orthogonal channel. Node \( i \) decodes messages the transmission of node \( i - 1 \) and forwards to node \( i + 1 \). Note that the nodes may or may not use different coding strategies. Node \( i \) transmits at power \( P_i \) and the multihop system is subject to the constraint \( \sum_{i=0}^{n-1} P_i = P \).

(a) In terms of \( P_i \) and \( N_i \), what is the capacity \( C_i \) of the channel \( i \) from node \( i \) to node \( i + 1 \)? (Yes, this is a gift.)

(b) For a given set of powers \( P_0, \ldots, P_n \), what is the capacity \( C \) of the multihop communication system from node 0 to node \( n \)? Express your answer in terms of \( C_i \). Explain your answer in terms of the end-to-end data rate \( R \).

(c) What is the optimal power allocation \( P_0, \ldots, P_{n-1} \)? What is the corresponding channel capacity \( C \)?