(1) True or False?

(a) True

\[ H(x, y, z) - I(x; g(x)) = H(x, y, z) - H(g(x)) + H(g(x) | x) \]

\[ = H(x, y, z, g(x)) - H(g(x)) = H(x, y, z | g(x)) \]

(b) (i) True

\[ h(x | y) = h(x - f(y) | y) \quad \forall \text{ function } f(\cdot) \]

so \[ h(x | y) = h(x - \frac{x - \text{lin} 1y}{f(\cdot)}) \]

(ii) True

For a jointly Gaussian RVs the optimal linear estimator is also the optimal estimator for which orthogonality principle holds, i.e.

\[ E = x - \frac{x - \text{lin} g(y)}{g(y)} \quad \text{for every } g(\cdot) \]

Thus:

\[ h(x - \frac{x - \text{lin} 1y}{f(\cdot)}) = h(x - \frac{x - \text{lin}}{x - \text{lin} 1y}) \]

C) (i) True

(ii) True

(ii') False

(d) False

If \( f(x) \) is convex so by the perspective transform \( g(t, x) = t \cdot f(x/t) \) is convex for \( t > 0 \) and not convex for \( t < 0 \).

Therefore a linear transformation will also not be convex for \( t < 0 \).
(2) Channel with State:

(a) Denote the capacity of the $S$-Channel by

$$C_S = \max_{p(x|s=0)} I(X; Y|S=1)$$

$$= \max_{p(x|s=1)} \left[ H(Y|S=1) - H(Y|X,S=1) \right]$$

Assume that the input $X$ is distributed according to

$X \sim$ Bernoulli ($d$) ($p(X=0) = d$, $p(X=1) = 1-d$) for $s=1$. and let us calculate the entropy terms:

$$H(Y|X, S=1) = \sum_{x \in \mathcal{X}} p(x) H(Y|X=x, S=1)$$

$$= d H(Y|X=0, S=1) + (1-d) H(Y|X=1, S=1)$$

$$= -d H_b(\varepsilon) + (1-d) \cdot \frac{-d}{\varepsilon} = d H_b(\varepsilon)$$

$$H(Y|S=1)$$

$$p(y=0|S=0) = p(y=0|x=0|S=1) + p(y=0|x=1|S=1)$$

$$= p(x=0|S=1) p(y=0|x=0, S=1) + p(x=1|S=1) p(y=0 |x=1, S=1)$$

$$= d \cdot (1-\varepsilon) + (1-d) \cdot \varepsilon = d(1-\varepsilon)$$

$$p(y=1|S=1) = 1 - d(1-\varepsilon)$$

$$H(Y|S=1) = H_b(d(1-\varepsilon))$$

In order to find the capacity we differentiate

$$I(X; Y|S=1) = H_b(d(1-\varepsilon)) -d H_b(\varepsilon)$$

with respect to $d$ and find the roots of the derivative.
\[
\frac{\partial}{\partial \alpha} I(x;y|s=1) = \frac{\partial}{\partial \alpha} \left[ H_b(0.8\alpha) - \alpha H_b(0.3\alpha) \right] = 0
\]

\Rightarrow \quad \alpha^* = 0.435664

\[
C_s = I(x;y|s=1) = 0.618931 \quad \text{[bits/channel]}
\]

(b) For the binary symmetric channel (BSC), the capacity is given by:

\[
C_{BSC} = I(x;y|s=2) = 1 - H_b(\alpha)
\]

Taking \( \alpha = 0.8 \):

\[
C_{BSC} = 1 - H_b(0.4) = 1 - 0.468996 = 0.531 \quad \text{[bits/channel]}
\]

(c) The capacity of the Z-Channel and the S-channel are equal since the channels are equivalent up to switching 0 with 1 and vice versa. We have:

\[
C_Z = I(x;y|s=3) = C_s = I(x;y|s=1) = 0.618934 \quad \text{[bits/channel]}
\]

(d) The capacity of a channel with states:

\[
C = \max_{p(x|s)} I(X;y|S)
\]

\[
= \max_{p(x|s)} \left[ p(s=1) I(X;Y|S=1) + p(s=2) I(X;Y|S=2) + p(s=3) I(X;Y|S=3) \right]
\]

\[
= \frac{1}{2} C_s + \frac{1}{3} C_{BSC} + \frac{1}{6} C_Z
\]

\[
= 0.583154
\]
(e) The rows of the matrix $P_{X|S}$ are the probability function which achieve capacity for each of the sub channels, i.e.,

$$P_{X|S} = \begin{bmatrix}
    p(x=0|s=1) & p(x=1|s=1) \\
    p(x=0|s=2) & p(x=1|s=2) \\
    p(x=0|s=3) & p(x=1|s=3)
\end{bmatrix} = \begin{bmatrix}
    0.435 & 0.565 \\
    0.5 & 0.5 \\
    0.565 & 0.435
\end{bmatrix}$$

where we have used the fact that the input probability that achieves capacity for the BSC is $[0.5, 0.5]$ (i.e. Bernouli ($\frac{1}{2}$)).
(3) a) We need to find a row vector \( \mu \) such that
\[ \mu \cdot P = \mu \]
holds. Denote \( \mu = \begin{bmatrix} q & 1-q \end{bmatrix} \)

(First equation)
\[ \mu \cdot P = \mu \Rightarrow q(1-a) + (1-1) = q \Rightarrow q = \frac{1}{1+a} \]

Thus:
\[ \mu = \begin{bmatrix} \frac{1}{1+a} & \frac{a}{1+a} \end{bmatrix} \]

b) We need to calculate \( \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} H(X_i|X_{i-1}) \)

\[ \lim_{n \to \infty} \frac{1}{n} H(X_i^n) = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} H(X_i|X_{i-1}) \]

Markov property
\[ X_i = X_{i-1} \quad X_i \sim \quad \text{for all } i \geq 1 \]

\[ = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \left[ p(x_{i-1}=0) H(X_i|X_{i-1}=0) + p(x_{i-1}=1) H(X_i|X_{i-1}=1) \right] \]

\[ = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} qH(a) + (1-q) \cdot 0 \]

\[ = \frac{H(a)}{1+a} \]

\[ \Rightarrow \text{ The entropy rule } \]

\[ H(X) = \frac{H(a)}{1+a} \]
4. (a) Since we need to reconstruct the data given any two hardisks and the source is distributed according to Bernoulli \(\frac{1}{2}\), i.e., no further compression is possible, each of the three hardisks must be at least of size 6 Gbit.

(b) These vertices represent the memory units (disks).

These vertices represent the choice of any two disks from the three disks used.

Since \(\text{mincut} = 2\) in this scheme, it is possible to transfer two bits using this scheme (although each memory unit (disk) receives only one bit).

The achievable scheme is: \(\text{in red}\).

(c) To solve this network coding problem, we need to find 4 vectors which are pairwise independent. This cannot be done using a binary alphabet \(1+1=2\) but it is possible using a ternary alphabet \(1+1+1=3\). The achievable scheme: \(\text{in red}\).