Solutions to Homework Set #3
Channel and Source coding

1. **Rates**

   (a) **Channels coding Rate:** Assuming you are sending 1024 different messages using 20 usages of a channel. What is the rate (in bits per channel use) that you send.

   (b) **Source coding Rate:** Assuming you have a file with 10^6 Ascii characters, where the alphabet of Ascii characters is 256. Assume that each Ascii character is represented by bits (binary alphabet before compression). After compressing it we get $4 * 10^6$ bits. What is the compression rate?

   Solution: Rates.

   (a)
$$R = \frac{1}{20} \log_2 1024 = \frac{1}{2}.$$

   (b)
$$R = \frac{4 * 10^6}{10^6 * \log_2(256)} = \frac{1}{2}.$$

2. **Preprocessing the output.**
   One is given a communication channel with transition probabilities $p(y \mid x)$ and channel capacity $C = \max_{p(x)} I(X;Y)$. A helpful statistician preprocesses the output by forming $\tilde{Y} = g(Y)$, yielding a channel $p(\tilde{y}|x)$. He claims that this will strictly improve the capacity.

   (a) Show that he is wrong.

   (b) Under what conditions does he not strictly decrease the capacity?

   Solution: Preprocessing the output.

(a) The statistician calculates $\tilde{Y} = g(Y)$. Since $X \to Y \to \tilde{Y}$ forms a Markov chain, we can apply the data processing inequality. Hence for every distribution on x,

$$I(X;Y) \geq I(X;\tilde{Y}).$$

Let $\tilde{p}(x)$ be the distribution on x that maximizes $I(X;\tilde{Y})$. Then

$$C = \max_{p(x)} I(X;Y) \geq I(X;Y)_{p(x)=\tilde{p}(x)} \geq I(X;\tilde{Y})_{p(x)=\tilde{p}(x)} = \max_{p(x)} I(X;\tilde{Y}) = \tilde{C}.$$

Thus, the helpful suggestion is wrong and processing the output does not increase capacity.

(b) We have equality (no decrease in capacity) in the above sequence of inequalities only if we have equality in data processing inequality, i.e., for the distribution that maximizes $I(X;\tilde{Y})$, we have $X \to \tilde{Y} \to Y$ forming a Markov chain. Thus, $\tilde{Y}$ should be a sufficient statistic.

3. **The Z channel.**
The Z-channel has binary input and output alphabets and transition probabilities $p(y|x)$ given by the following matrix:

$$Q = \begin{bmatrix} 1 & 0 \\ 1/2 & 1/2 \end{bmatrix} \qquad x, y \in \{0, 1\}$$

Find the capacity of the Z-channel and the maximizing input probability distribution.

Solution: The Z channel.
First we express $I(X;Y)$, the mutual information between the input and output of the Z-channel, as a function of $\alpha = \Pr(X = 1)$:

$$\begin{aligned} H(Y|X) &= \Pr(X=0) \cdot 0 + \Pr(X=1) \cdot 1 = \alpha \\ H(Y) &= \mathbf{H}(\Pr(Y=1)) = \mathbf{H}(\alpha/2) \\ I(X;Y) &= H(Y) - H(Y|X) = \mathbf{H}(\alpha/2) - \alpha \end{aligned}$$

Since $I(X;Y)$ is strictly concave on α (why?) and $I(X;Y) = 0$ when $\alpha = 0$ and $\alpha = 1$, the maximum mutual information is obtained for some value of α such that $0 < \alpha < 1$.

Using elementary calculus, we determine that

$$\frac{d}{d\alpha} I(X;Y) = \frac{1}{2} \log_2 \frac{1-\alpha/2}{\alpha/2} - 1\,,$$

which is equal to zero for $\alpha = 2/5$. (It is reasonable that $\Pr(X=1) < 1/2$ since $X = 1$ is the noisy input to the channel.) So the capacity of the Z-channel in bits is $H(1/5) - 2/5 = 0.722 - 0.4 = 0.322$.

4. **Using two channels at once.**
Consider two discrete memoryless channels $(\mathcal{X}_1, p(y_1 \mid x_1), \mathcal{Y}_1)$ and $(\mathcal{X}_2, p(y_2 \mid x_2), \mathcal{Y}_2)$ with capacities C_1 and C_2 respectively. A new channel $(\mathcal{X}_1 \times \mathcal{X}_2, p(y_1 \mid x_1) \times p(y_2 \mid x_2), \mathcal{Y}_1 \times \mathcal{Y}_2)$ is formed in which $x_1 \in \mathcal{X}_1$ and $x_2 \in \mathcal{X}_2$, are *simultaneously* sent, resulting in y_1, y_2. Find the capacity of this channel.

Solution: Using two channels at once.
To find the capacity of the product channel $(\mathcal{X}_1 \times \mathcal{X}_2, p(y_1, y_2|x_1, x_2), \mathcal{Y}_1 \times \mathcal{Y}_2)$, we have to find the distribution $p(x_1, x_2)$ on the input alphabet $\mathcal{X}_1 \times \mathcal{X}_2$ that maximizes $I(X_1, X_2; Y_1, Y_2)$. Since the transition probabilities are given as $p(y_1, y_2|x_1, x_2) = p(y_1|x_1)p(y_2|x_2)$,

$$\begin{aligned} p(x_1, x_2, y_1, y_2) &= p(x_1, x_2)p(y_1, y_2|x_1, x_2) \\ &= p(x_1, x_2)p(y_1|x_1)p(y_2|x_2), \end{aligned}$$

Therefore, $Y_1 \to X_1 \to X_2 \to Y_2$ forms a Markov chain and

$$\begin{aligned} I(X_1, X_2; Y_1, Y_2) &= H(Y_1, Y_2) - H(Y_1, Y_2|X_1, X_2) \\ &= H(Y_1, Y_2) - H(Y_1|X_1, X_2) - H(Y_2|X_1, X_2) \quad (1) \\ &= H(Y_1, Y_2) - H(Y_1|X_1) - H(Y_2|X_2) \quad (2) \\ &\le H(Y_1) + H(Y_2) - H(Y_1|X_1) - H(Y_2|X_2) \quad (3) \\ &= I(X_1; Y_1) + I(X_2; Y_2), \end{aligned}$$

where Eqs. (1) and (2) follow from Markovity, and Eq. (3) is met with equality if X_1 and X_2 are independent and hence Y_1 and Y_2 are independent. Therefore

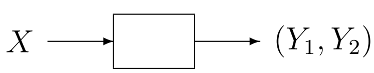
$$\begin{aligned} C &= \max_{p(x_1,x_2)} I(X_1, X_2; Y_1, Y_2) \\ &\le \max_{p(x_1,x_2)} I(X_1; Y_1) + \max_{p(x_1,x_2)} I(X_2; Y_2) \\ &= \max_{p(x_1)} I(X_1; Y_1) + \max_{p(x_2)} I(X_2; Y_2) \\ &= C_1 + C_2. \end{aligned}$$

with equality iff $p(x_1, x_2) = p^*(x_1)p^*(x_2)$ and $p^*(x_1)$ and $p^*(x_2)$ are the distributions that maximize C_1 and C_2 respectively.

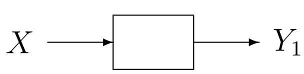
5. **A channel with two independent looks at Y.**
Let Y_1 and Y_2 be conditionally independent and conditionally identically distributed given X. Thus $p(y_1, y_2|x) = p(y_1|x)p(y_2|x)$.

   (a) Show $I(X; Y_1, Y_2) = 2I(X; Y_1) - I(Y_1; Y_2)$.

   (b) Conclude that the capacity of the channel



is less than twice the capacity of the channel



Solution: A channel with two independent looks at Y.

(a)

$$\begin{aligned} I(X; Y_1, Y_2) &= H(Y_1, Y_2) - H(Y_1, Y_2|X) && (4) \\ &= H(Y_1) + H(Y_2) - I(Y_1; Y_2) - H(Y_1|X) - H(Y_2|X) && (5) \\ &\quad \text{(since } Y_1 \text{ and } Y_2 \text{ are conditionally independent given } X\text{)} && (6) \\ &= I(X; Y_1) + I(X; Y_2) - I(Y_1; Y_2) && (7) \\ &= 2I(X; Y_1) - I(Y_1; Y_2) \quad \text{(since } Y_1 \text{ and } Y_2 \text{ are conditionally identically distributed)} && (8) \end{aligned}$$

(b) The capacity of the single look channel $X \to Y_1$ is

$$C_1 = \max_{p(x)} I(X; Y_1). \tag{9}$$

The capacity of the channel $X \to (Y_1, Y_2)$ is

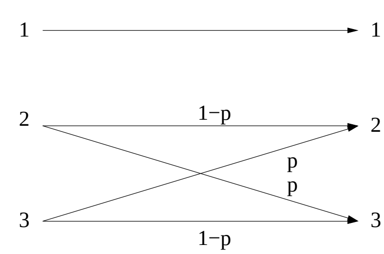
$$\begin{aligned} C_2 &= \max_{p(x)} I(X; Y_1, Y_2) & (10) \\ &= \max_{p(x)} 2I(X; Y_1) - I(Y_1; Y_2) & (11) \\ &\leq \max_{p(x)} 2I(X; Y_1) & (12) \\ &= 2C_1. & (13) \end{aligned}$$

Hence, two independent looks cannot be more than twice as good as one look.

6. **Choice of channels.**
Find the capacity C of the union of 2 channels $(\mathcal{X}_1, p_1(y_1|x_1), \mathcal{Y}_1)$ and $(\mathcal{X}_2, p_2(y_2|x_2), \mathcal{Y}_2)$ 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 $2^C = 2^{C_1} + 2^{C_2}$.

(b) What is the capacity of this Channel?



Solution: Choice of channels.

(a) Let

$$\theta = \begin{cases} 1, & \text{if the signal is sent over the channel 1} \\ 2, & \text{if the signal is sent over the channel 2} \end{cases}.$$

Consider the following communication scheme: The sender chooses between two channels according to Bern(α) coin flip. Then the channel input is $X = (\theta, X_\theta)$.

Since the output alphabets $\mathcal{Y}_1$ and $\mathcal{Y}_2$ are disjoint, θ is a function of Y, i.e. $X \to Y \to \theta$.

Therefore,

$$\begin{aligned} I(X;Y) &= I(X;Y,\theta) \\ &= I(X_\theta,\theta;Y,\theta) \\ &= I(\theta;Y,\theta) + I(X_\theta;Y,\theta|\theta) \\ &= I(\theta;Y,\theta) + I(X_\theta;Y|\theta) \\ &= H(\theta) + \alpha I(X_\theta;Y|\theta=1) + (1-\alpha)I(X_\theta;Y|\theta=2) \\ &= H(\alpha) + \alpha I(X_1;Y_1) + (1-\alpha)I(X_2;Y_2). \end{aligned}$$

Thus, it follows that

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

which is a strictly concave function on α. Hence, the maximum exists and by elementary calculus, one can easily show $C = \log_2(2^{C_1} + 2^{C_2})$, which is attained with $\alpha = 2^{C_1}/(2^{C_1}+2^{C_2})$.

If one interprets $M = 2^C$ as the effective number of noise free symbols, then the above result follows in a rather intuitive manner: we have $M_1 = 2^{C_1}$ noise free symbols from channel 1, and $M_2 = 2^{C_2}$ noise free symbols from channel 2. Since at each step we get to choose which channel to use, we essentially have $M_1 + M_2 = 2^{C_1} + 2^{C_2}$ noise free symbols for the new channel. Therefore, the capacity of this channel is $C = \log_2(2^{C_1} + 2^{C_2})$.

This argument is very similiar to the effective alphabet argument given in Problem 19, Chapter 2 of the text.

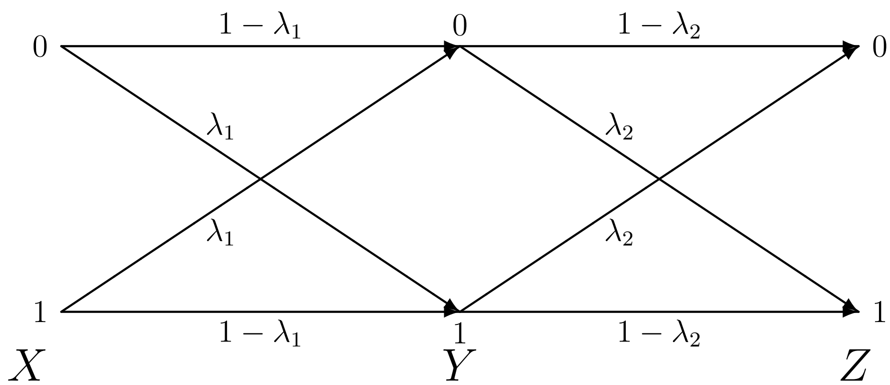
(b) From part (b) we get capacity is

$$\log(2^{1-H(p)} + 2^0).$$

7. **Cascaded BSCs.**

Consider the two discrete memoryless channels $(\mathcal{X}, p_1(y|x), \mathcal{Y})$ and $(\mathcal{Y}, p_2(z|y), \mathcal{Z})$.

Let $p_1(y|x)$ and $p_2(z|y)$ be binary symmetric channels with crossover probabilities λ_1 and λ_2 respectively.



(a) What is the capacity C_1 of $p_1(y|x)$?

(b) What is the capacity C_2 of $p_2(z|y)$?

(c) We now cascade these channels. Thus $p_3(z|x) = \sum_y p_1(y|x)p_2(z|y)$. What is the capacity C_3 of $p_3(z|x)$? Show $C_3 \leq \min\{C_1, C_2\}$.

(d) Now let us actively intervene between channels 1 and 2, rather than passively transmitting y^n. What is the capacity of channel 1 followed by channel 2 if you are allowed to decode the output y^n of channel 1 and then reencode it as $\tilde{y}^n$ for transmission over channel 2? (Think $W \longrightarrow x^n(W) \longrightarrow y^n \longrightarrow \tilde{y}^n(y^n) \longrightarrow z^n \longrightarrow \hat{W}$.)

(e) What is the capacity of the cascade in part c) if the receiver can view *both* Y and Z?

Solution: Cascaded BSCs.

(a) This is a simple BSC with capacity $C_1 = 1 - H_b(\lambda_1)$.

(b) Similarly, $C_2 = 1 - H_b(\lambda_2)$.

(c) This is also a BSC channel with transition probability $\lambda_1 * \lambda_2 = \lambda_1(1-\lambda_2) + (1-\lambda_1)\lambda_2$ and thus $C_3 = 1 - H_b(\lambda_1 * \lambda_2)$. From the markov chain $X - Y - Z$ we can see that

$$C_3 = I(X;Z) \leq I(X;Y,Z) = I(X;Y) = C_1$$

and additionaly

$$C_3 = I(X;Z) \leq I(X,Y;Z) = I(Y;Z) = C_2$$

and thus

$$C_3 \leq \min\{C_1, C_2\}$$

(d) If one can reencode Y then we obtain that

$$C_3 = \min\{C_1, C_2\}$$

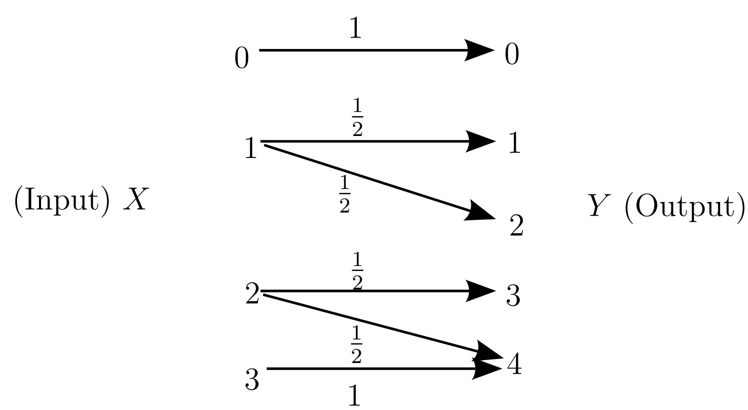
since only one of the channels will be the bottleneck between X and Z.

(e) From the Markov chain $X - Y - Z$ we can see that

$$C_3' = I(X; Y, Z) = I(X; Y) = C_1$$

8. **Channel capacity**

(a) What is the capacity of the following channel



(b) Provide a simple scheme that can transmit at rate $R = \log_2 3$ bits through this channel.

Solution for Channel capacity

(a) We can use the solution of previous home question:

$$C = \log\left(2^{C_1} + 2^{C_2} + 2^{C_3}\right)$$

Now we need to calculate the capacity of each channel:

$$C_1 = \max_{p(x)} I(X; Y) = H(Y) - H(Y|X) = 0 - 0 = 0$$

$$C_2 = \max_{p(x)} I(X;Y) = H(Y) - H(Y|X) = 1 - 1 = 0$$

$$\begin{aligned} C_3 = \max_{p(x)} I(X;Y) &= \max_{p(x)} \{H(Y) - H(Y|X)\} \\ &= \max_{p(x)} \left[-\frac{1}{2}p_2 \log\left(\frac{1}{2}p_2\right) - \left(\frac{1}{2}p_2 + p_3\right) \log\left(\frac{1}{2}p_2 + p_3\right)\right] - p_2 \end{aligned}$$

Assigning $p_3 = 1 - p_2$ and derive against p_2:

$$\frac{dI(X;Y)}{dp_2} = -\frac{p_2}{2} \cdot \frac{1}{2} \cdot \frac{1}{\frac{p_2}{2}} - \frac{1}{2}\log\left(\frac{p_2}{2}\right) + \frac{2-p_2}{2} \cdot \frac{1}{2} \cdot \frac{1}{\frac{2-p_2}{2}} + \frac{1}{2}\log\left(\frac{2-p_2}{2}\right) - 1 = 0$$

And as result $p_2 = \frac{2}{5}$:

$$C_3 \approx 0.322$$

And, finally:

$$C = \log(2^0 + 2^0 + 2^{0.322}) \approx 1.7$$

(b) Here is a simple code that achieves capacity.

Encoding: You just use ternary representation of the message and send using 0,1,2 but no 3 (or 0,1,3 but no 2) of the input channel.
Decoding: map the ternary output into the message.

9. **A channel with a switch.** Consider the channel that is depicted in Fig.1, there are two channels with the conditional probabilities $p(y_1|x)$ and $p(y_2|x)$. These two channels have common input alphabet $\mathcal{X}$ and two **disjoint** output alphabets $\mathcal{Y}_1, \mathcal{Y}_2$ (a symbol that appears in $\mathcal{Y}_1$ can't appear in $\mathcal{Y}_2$). The position of the switch is determined by R.V Z which is independent of X, where $\Pr(Z = 1) = \lambda$.

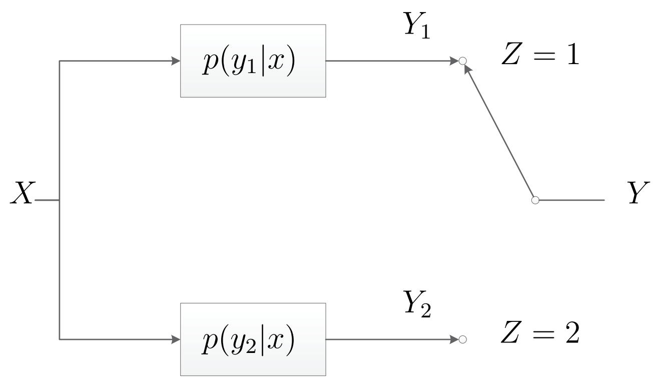


Figure 1: The channel.

(a) Show that

$$I(X;Y) = \lambda I(X;Y_1) + \bar{\lambda} I(X;Y_2). \tag{14}$$

(b) The capacity of this system is given by $C = \max_{p(x)} I(X;Y)$. Show that

$$C \leq \lambda C_1 + \bar{\lambda} C_2, \tag{15}$$

where $C_i = \max_{p(x)} I(X;Y_i)$.
When is equality achieved?

(c) The sub-channels defined by $p(y_1|x)$ and $p(y_2|x)$ are now given in the following figure, where $p = \frac{1}{2}$.
Find the input probability $p(x)$ that maximizes $I(X;Y)$.
For this case, does the equality $C = \lambda C_1 + \bar{\lambda} C_2$ stand? explain!

Solution: Channel with state

(a) Since the alphabet $\mathcal{Y}_1, \mathcal{Y}_2$ are disjoint the markov chain $X - Y - Z$ holds. Additionally, X and Z are independent and thus we have

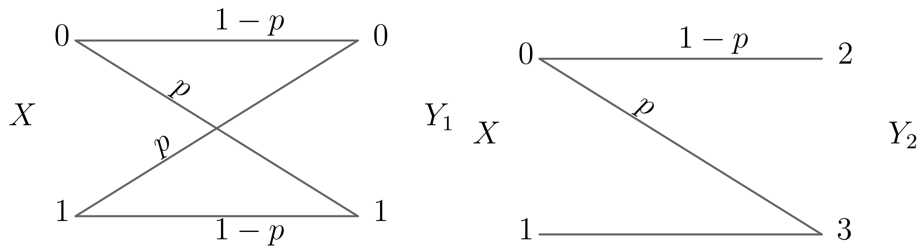


Figure 2: (a) describes channel 1 - BSC with transition probability p. (b) describes channel 2 - Z channel with transition probability p.

that

$$\begin{aligned} I(X;Y) &= I(X;Y|Z) \\ &= \Pr(z=1)I(X;Y|z=1) + \Pr(z=2)I(X;Y|z=2) \\ &= \lambda I(X;Y|z=1) + \bar{\lambda} I(X;Y|z=2) \\ &= \lambda I(X;Y_1) + \bar{\lambda} I(X;Y_2) \end{aligned}$$

(b) Consider the following:

$$\begin{aligned} C &= \max_{p(x)} I(X;Y) \\ &= \max_{p(x)} \{\lambda I(X;Y_1) + \bar{\lambda} I(X;Y_2)\} \\ &\leq \max_{p(x)} \{\lambda I(X;Y_1)\} + \max_{p(x)} \{\bar{\lambda} I(X;Y_2)\} \\ &= \lambda C_1 + \bar{\lambda} C_2 \end{aligned}$$

(c) Since $p = \frac{1}{2}$, the capacity of the first channel is 0 and thus we only need to maximize over the second channel. Thus, we would like to find $p(x)$ such that $\bar{\lambda} I(X;Y_2) = \bar{\lambda}(H_b(\frac{\alpha}{2}) - \alpha H_b(\frac{1}{2})$ is maximized where $\Pr(x=0) = \alpha \in [0,1]$. In this case the solution is $\alpha = 0.4$. We can see that the equality in holds.

10. **Channel with state**

A discrete memoryless (DM) state dependent channel with state space $\mathcal{S}$ is defined by an input alphabet $\mathcal{X}$, an output alphabet $\mathcal{Y}$ and a set of channel transition matrices $\{p(y|x,s)\}_{s\in\mathcal{S}}$. Namely, for each $s \in \mathcal{S}$

the transmitter sees a different channel. The capacity of such a channel where the state is know causally to both encoder and decoder is given by:

$$C = \max_{p(x|s)} I(X;Y|S). \tag{16}$$

Let $|\mathcal{S}| = 3$ and the three different channels (one for each state $s \in \mathcal{S}$) are as illustrated in the following figure

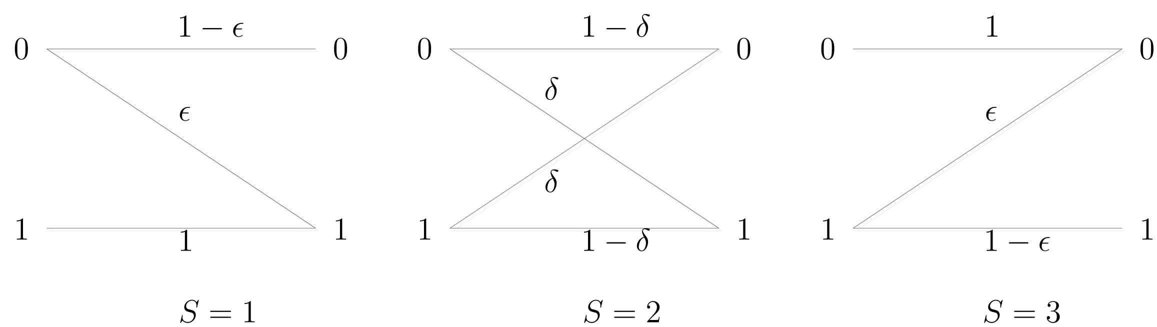


Figure 3: The three state dependent channel.

The state process is i.i.d. according to the distribution $p(s)$.

(a) Find an expression for the capacity of the S-channel (the channel the transmitter sees given $S = 1$) as a function of ϵ.

(b) Find an expression for the capacity of the BSC (the channel the transmitter sees given $S = 2$) as a function of δ.

(c) Find an expression for the capacity of the Z-channel (the channel the transmitter sees given $S = 3$) as a function of ϵ.

(d) Find an expression for the capacity of the DM state dependent channel (using formula (16)) for $p(s) = [\frac{1}{2} \ \ \frac{1}{3} \ \ \frac{1}{6}]$ as a function of ϵ and δ.

(e) Let us define a conditional probability matrix $P_{X|S}$ for two random variables X and S with $|\mathcal{X}| = \{0, 1\}$ and $|\mathcal{S}| = \{1, 2, 3\}$, by:

$$\left[P_{X|S}\right]_{i=1,j=1}^{3,2} = p(x = j-1|s = i). \tag{17}$$

What is the input conditional probability matrix $P_{X|S}$ that achieves the capacity you have found in (d)?

Solution: Channel with state

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

$$\begin{aligned}\mathcal{C}_S &= \max_{p(x|s=1)} I(X;Y|S=1)\\ &= \max_{p(x|s=1)} \left[H(Y|S=1) - H(Y|X,S=1)\right]\end{aligned}$$

Assume that the input X is distributed according to $X \sim Bernoulli(\alpha)$ for $s=1$, and let us calculate the entropy terms:

$$\begin{aligned}H(Y|X,S=1) =& \sum_{x\in\mathcal{X}} H(Y|X=x,S=1)\\ =&\alpha H(Y|X=0,S=1) + (1-\alpha)H(Y|X=1,S=1)\\ =&\alpha H_b(\epsilon) + (1-\alpha)\cdot 0 = \alpha H_b(\epsilon)\end{aligned}$$

Consider

$$\begin{aligned}p(y=0|s=1) =& p(y=0,x=0|s=1) + p(y=0,x=1|s=1)\\ =&\alpha(1-\epsilon)\end{aligned}$$

Then

$$H(Y|X,S=1) = H_b(\alpha(1-\epsilon)).$$

We want to maximize

$$I(X;Y|S=1) = H_b(\alpha(1-\epsilon)) - \alpha H_b(\epsilon).$$

Taking derivative with respect to α and find the roots of the derivation gives the capacity.

(b) For the binary symmetric channel (BSC) the capacity is given by $\mathcal{C}_{BSC} = 1 - H_b(\delta)$.

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

(d) The capacity of the state dependent channel is

$$\begin{aligned}\mathcal{C} &= \max_{p(x|s)} I(X;Y|S)\\ &= \max_{p(x|s)} \big[p(S=1)I(X;Y|S=1) + p(S=2)I(X;Y|S=2)\\ &\quad + p(S=3)I(X;Y|S=3)\big]\\ &= \frac{1}{2}\mathcal{C}_S + \frac{1}{3}\mathcal{C}_{BSC} + \frac{1}{6}\mathcal{C}_Z\end{aligned}$$

(e) The rows of the matrix $P_{X|S}$ are the conditional probability function which achieve capacity for each of the sub channel.

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

11. **Modulo channel**

(a) Consider the DMC defined as follows: Output $Y = X \oplus_2 Z$ where X, taking values in $\{0,1\}$, is the channel input, $\oplus_2$ is the modulo-2 summation operation, and Z is binary channel noise uniform over $\{0,1\}$ and independent of X. What is the capacity of this channel?

(b) Consider the channel of the previous part, but suppose that instead of modulo-2 addition $Y = X \oplus_2 Z$, we perform modulo-3 addition $Y = X \oplus_3 Z$. Now what is the capacity?

Solution: Modulo channel

(a) This is a simple case of a BSC with transition probability of $p = 1/2$ and thus the capacity in this case is $C_{BSC} = 0$.

(b) In this case we can model the channel as a BEC as studied in class. The input is of binary alphabet and the output is of a trenary alphabet. The probability of error is $\Pr(z=1) = p = \frac{1}{2}$ and thus the capacity for this channel is $C_{BEC} = 1 - p = \frac{1}{2}$.

12. **Cascaded BSCs**: Given is a cascade of k identical and independent binary symmetric channels, each with crossover probability α.

(a) In the case where no encoding or decoding is allowed at the intermediate terminals, what is the capacity of this cascaded channel as a function of k, α.

(b) Now, assume that encoding and decoding is allowed at the intermediate points, what is the capacity as a function of k, α.

(c) What is the capacity of each of the above settings in the case where the number of cascaded channels, k, goes to infinity?

Solution: Cascaded BSCs.

(a) Cascaded BSCs result a new BSC with a new parameter, β. Therefore, the capacity is $C_a = 1 - H_2(\beta)$ and the parameter β can be found as follows:

$$\begin{aligned}\beta &= \sum_{\{i \le k : i \text{ is odd}\}} \binom{k}{i} \alpha^i (1-\alpha)^{k-i} \\ &= \sum_{i=1}^{k} \binom{k}{i} \frac{1-(-1)^i}{2} \alpha^i (1-\alpha)^{k-i} \\ &= \frac{1}{2} - \frac{1}{2} \sum_{i=1}^{k} \binom{k}{i} (-\alpha)^i (1-\alpha)^{k-i} \\ &= \frac{1}{2}(1-(1-2\alpha)^k).\end{aligned}$$

Answers of β as the initial sum over odd indices or the binary convolutional of k identical parameters α have been accepted as well.

(b) We have seen in HW that in the case of encoding and decoding the capacity of the cascaded channel equals $C_b = \min\{C_i\}$. Since all channels are identical with capacity $1 - H_2(\alpha)$, we have that $C_b = 1 - H_2(\alpha)$.

(c) In (a), $\beta \to 0.5$ as $k \to \infty$ so $C_a \to 0$.

For (b), the number of cascaded channels does not change the capacity which remains $C_b = 1 - H_2(\alpha)$.