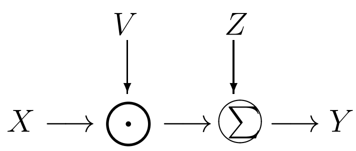
Homework Set #6
Source channel separation, Max entropy principle, and channel coding with side information

1. **Fading channel.**
   Consider an additive noise fading channel



$$Y = XV + Z,$$

   where Z is additive noise, V is a random variable representing fading, and Z and V are independent of each other and of X.

   (a) Argue that knowledge of the fading factor V improves capacity by showing
   $$I(X;Y|V) \geq I(X;Y).$$

   (b) Incidentally, conditioning does not always increase mutual information. Give an example of $p(u,r,s)$ such that $I(U;R|S) < I(U;R)$.

2. **Diversity System**

   For the following system, a message $W \in \{1, 2, \ldots, 2^{nR}\}$ is encoded into two symbol blocks $X_1^n = (X_{1,1}, X_{1,2}, ..., X_{1,n})$ and $X_2^n = (X_{2,1}, X_{2,2}, ..., X_{2,n})$ that are being transmitted over a channel. The average power constrain on the inputs are $\frac{1}{n}E[\sum_{i=1}^n X_{1,i}^2] \leq P_1$ and $\frac{1}{n}E[\sum_{i=1}^n X_{2,i}^2] \leq P_2$. The channel has a multiplying effect on X_1, X_2 by factor h_1, h_2, respectively, i.e., $Y = h_1X_1 + h_2X_2 + Z$, where Z is a white Gaussian noise $Z \sim N(0, \sigma^2)$.

   (a) Find the joint distribution of X_1 and X_2 that bring the mutual information $I(Y; X_1, X_2)$ to a maximum? (You need to find $\arg\max P_{X_1,X_2} I(X_1, X_2; Y)$.)

   (b) What is the capacity of the system ?

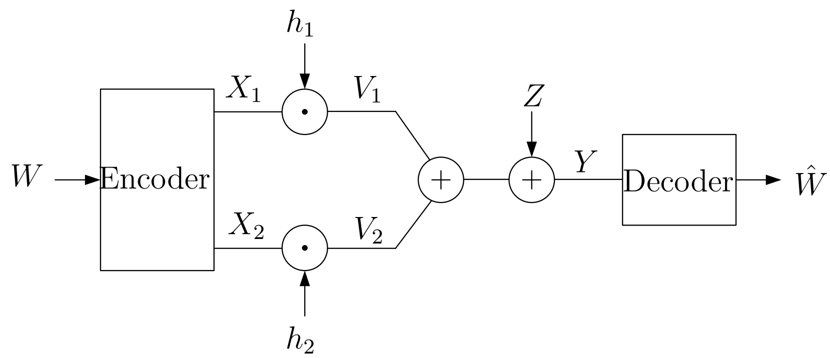


Figure 1: The communication model

(c) Express the capacity for the following cases:

i. $h_1 = 1$, $h_2 = 1$?

ii. $h_1 = 1$, $h_2 = 0$?

iii. $h_1 = 0$, $h_2 = 0$?

3. **AWGN with two noises**(15 points)

Figure 2 depicts a communication system with an AWGN (Additive white noise Gaussian) channel whith two i.i.d. noises $Z_1 \sim N(0, \sigma_1^2)$, $Z_2 \sim N(0, \sigma_2^2)$ that are independent of each other and are added to the signal X, i.e., $Y = X + Z_1 + Z_2$. The average power constrain on the input is P, i.e., $\frac{1}{n}E[\sum_{i=1}^{n} X_i^2] \leq P$. In the sub-questions below we consider the cases where the noise Z_2 may or may not be known to the encoder and decoder.

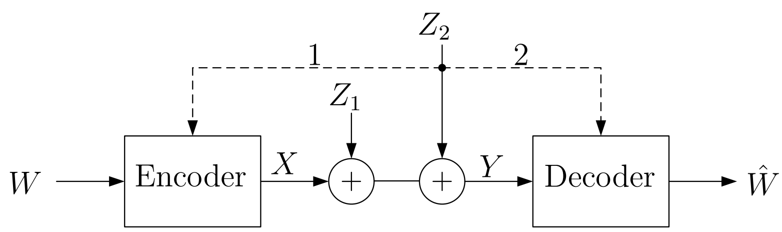


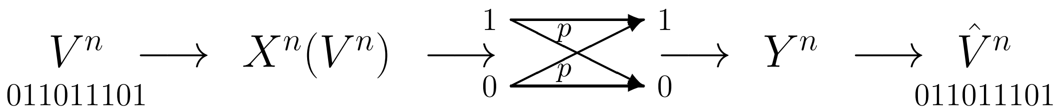
Figure 2: Two noise sources

(a) Find the channel capacity for the case in which the noise in not known to either sides (lines 1 and 2 are <u>disconnected</u> from the encoder and the decoder).

(b) Find the capacity for the case that the noise Z_2 is known to the encoder and decoder (lines 1 and 2 are <u>connected</u> to both the encoder and decoder). This means that the codeword X^n may depend on the message W and the noise Z_2^n and the decoder decision $\hat{W}$ may depend on the output Y^n and the noise Z_2^n. (**Hint**: Could the capacity be lager than $\frac{1}{2}\log(1+\frac{P}{\sigma_1^2})$?)

(c) Find the capacity for the case that the noise Z_2 is known only to the decoder. (line 1 is <u>disconnected</u> from the encoder and line 2 is <u>connected</u> to the decoder). This means that the codewords X^n may depend only on the message W and the decoder decision $\hat{W}$ may depend on the output Y^n and the noise Z_2^n.

4. **Source and channel. (Please read the relevant lecture on source channel separation)**
We wish to encode a Bernoulli(α) process $V_1, V_2, \ldots$ for transmission over a binary symmetric channel with error probability p.



Find conditions on α and p so that the probability of error $P(\hat{V}^n \neq V^n)$ can be made to go to zero as $n \longrightarrow \infty$.

5. **Maximum entropy.**
Find the maximum entropy density f satisfying $EX = \alpha_1, E\ln X = \alpha_2$. That is,

$$\text{maximize} \quad h(f)$$

subject to $\int xf(x)\,dx = \alpha_1, \int(\ln x)f(x)\,dx = \alpha_2$. What family of densities is this?

6. **Minimum relative entropy $D(P \parallel Q)$ under constraints on P.**
We wish to find the (parametric form) of the probability mass function $P(x), x \in \{1, 2, \ldots\}$ that minimizes the relative entropy $D(P \parallel Q)$ over all P such that $\sum P(x) g_i(x) = \alpha_i, i = 1, 2, \ldots$.

   (a) Use Lagrange multipliers to guess that

   $$P^*(x) = Q(x) e^{\sum_{i=1}^{\infty} \lambda_i g_i(x) + \lambda_0}$$

   achieves this minimum if there exist λ_i's satisfying the α_i constraints. This generalizes the theorem on maximum entropy distributions subject to constraints.

   (b) Verify that P^* minimizes $D(P \parallel Q)$.

7. **Maximum entropy with marginals.**
What is the maximum entropy probability mass function $p(x, y)$ with the following marginals? You may wish to guess and verify a more general result.

	y_1	y_2	y_3	
x_1	p_{11}	p_{12}	p_{13}	1/2
x_2	p_{21}	p_{22}	p_{23}	1/4
x_3	p_{31}	p_{32}	p_{33}	1/4
	2/3	1/6	1/6	