Mathematical methods in communication (361-2-5891) 22nd July 2009
(Dr. Permuter Haim and Mr. Yuval Carmel)
Material: Any hard-copy material (books and notes) is allowed.

Final Exam

Total time for the exam: 3 hours!

1. **True or False** (55 points)

   Let X, Y, Z be discrete random variable. Copy each relation to your notebook and write **true** or **false**. If it's true, prove it. If it is false give a counterexample or prove that the opposite is true.

   For instance:

   - $H(X) \geq H(X|Y)$ is **true**. Proof: In the class we showed that $I(X;Y) > 0$, hence $H(X) - H(X|Y) > 0$.
   - $H(X) + H(Y) \leq H(X,Y)$ is **false**. Actually the opposite is true, i.e., $H(X) + H(Y) \geq H(X,Y)$ since $I(X;Y) = H(X) + H(Y) - H(X,Y) \geq 0$.

   (a) If $H(X|Y) = H(X)$ if and only if X and Y are independent.

   (b) If X, Y, Z form the Markov chain, $X - Y - Z$, then $H(X|Y) \leq H(Y|Z)$.

   (c) For any two probability mass functions (pmf) P, Q,

   $$D\left(\frac{P+Q}{2}||Q\right) \leq \frac{1}{2}D(P||Q),$$

   where $D(||)$ is a divergence between two pmfs.

   (d) Let X and Y be two independent random variables. Then

   $$H(X,Y) \leq H(X+Y).$$

   (e) $|I(X;Y) - I(X;Y|Z)| \leq H(Z)$

   (f) Let X^n be i.i.d $\sim P_X$. Let A_n and B_n be two sets of sequences X^n such that $\lim_{n\to\infty} \Pr(A_n) = 1$ and $\lim_{n\to\infty} \Pr(B_n) = 1$. For instance, A_n could be the typical set with parameter ϵ, i.e., $A_n =$

$\{x^n : |\frac{1}{n}\log P(x^n) - H(X)| \leq \epsilon\}$. Then, there might be a case where

$$\lim_{n\to\infty} \frac{1}{n}\log|A_n \cap B_n| < H(X),$$

namely the normalized log of the cardinality of $A_n \cap B_n$ could be less than $H(X)$.

(g) There exists a compression software, that whenever we apply it on any file it compresses by 50% the file.

(h) Let X, Y be two random variables, jointly Gaussian, with mean zero and $E[X^2] = E[Y^2] = \sigma^2$ and $E[XY] = \rho\sigma^2$, where $|\rho| \leq 1$. Then,

$$h(Y|X) = \frac{1}{2}\log 2\pi e\sigma^2(1-\rho^2)$$

(i) Let X, Y be two random variables with continuous alphabet and density distribution $f_{X,Y}$. Then,

$$0 \leq h(Y|X),$$

where $h(Y|X)$ is differential entropy of Y given X.

(j) Let $R(D)$ be a rate distortion function. $R(D)$ is nonincreasing in D. I.e., if $D_1 \geq D_2$ then $R(D_1) \leq R(D_2)$.

(k) Let $R(D)$ be a rate distortion function where the source is memoryless X and the reconstruction is $\hat{X}$. For any distortion we have

$$R(D) \leq \log\Big(\min(|\hat{\mathcal{X}}|, |\mathcal{X}|)\Big)$$

where $|\hat{\mathcal{X}}|$ and $|\mathcal{X}|$ are the cardinality of the alphabets $\hat{X}$ and X, respectively.

2. **Compression** (15 points)

   (a) Give a Huffman encoding into an alphabet of size D = 2 of the following probability mass function:

   $$\left(\frac{1}{2}, \frac{1}{8}, \frac{1}{8}, \frac{1}{16}, \frac{1}{16}, \frac{1}{16}, \frac{1}{16}\right)$$

   (b) Assume you have a file of size 1,000 symbols where the symbols are distributed i.i.d. according to the pmf above. After applying the Huffman code, what would be the pmf of the compressed binary file ($P(0) =?$ and $P(1) =?$) and what would be the expected length?

3. **Diversity System** (15 points)

   For the following system, a message $W \in \{1, 2, \dots, 2^{nR}\}$ 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 $\arg\max P_{X_1,X_2} I(X_1, X_2; Y)$.)

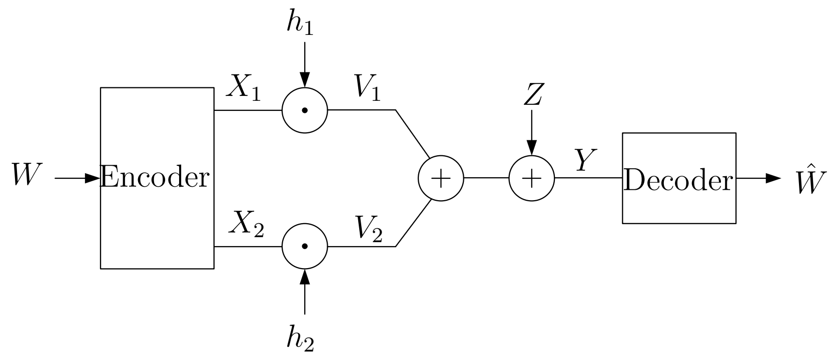

Figure 1: The communication model

(b) What is the capacity of the system ?

(c) Express the capacity for the following cases:

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

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

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

4. **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.

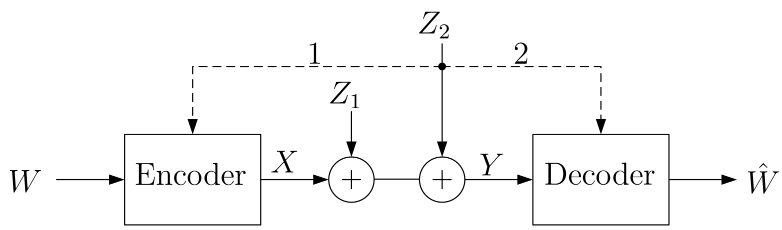

Figure 2: Two noise sources

(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 connected 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.