Random codes in communication 22 July 2012
(Dr. Permuter Haim)

Final Exam, MOED GIMEL
Total time for the exam: 3 hours!

1) **True/False Bound on each Huffman codeword** (15 points)
The following claim is sometimes found in the literature:

It can be shown that the length of the binary Huffman codeword of a symbol a_i with probability p_i is always less than or equal to $\lceil -\log_2 p_i \rceil$

Is this claim true? Justify your answer

2) **Deletion Channel** (15 points)
Consider a binary sequence of length n, denoted by $X^n = (X_1, \ldots, X_n)$. Consider another binary sequence of length n called deletion pattern, denoted by $D^n = (D_1, \ldots, D_n)$, which determines how X_n is to be deleted. Then, the output of the deletion process, denoted by $y(X^n, D^n)$, is derived from X^n by deleting the bits at those locations where the deletion pattern is 1.
Consider the following example for $n = 10$:

$$X^n = (0,1,1,0,0,0,1,1,1,0)$$
$$D^n = (0,1,0,1,1,0,0,0,1,0)$$
$$y(X^n, D^n) = (0,1,0,1,1,0)$$

The source sequence $X^n \in \{0,1\}^n$ is i.i.d. Bernoulli$(1/2)$, and the deletion pattern $D^n \in \{0,1\}^n$ is i.i.d. Bernoulli(d), independent of X^n.
We are interested in computing the mutual information between X^n and $y(X^n, D^n)$, $I(X^n; y(X^n, D^n))$. For each relation state if its true or false:

a) $I(X^n; y(X^n, D^n)) = H(y(X^n, D^n)) + H(D^n) - H(y(X^n, D^n)|X^n, D^n)$
b) $I(X^n; y(X^n, D^n)) = H(y(X^n, D^n)) - H(D^n) + H(D^n|X^n, y(X^n, D^n))$
c) $I(X^n; y(X^n, D^n)) = H(y(X^n, D^n)) - H(D^n)$
d) $I(X^n; y(X^n, D^n)) = H(X^n) - H(X^n|y(X^n, D^n)) - H(D^n|X^n, y(X^n, D^n))$

3) **Condition on length of Huffman codeword to be larger then 1** (20 points)
Consider a source of K symbols, with $p_1 \geq p_2 \geq \ldots \geq p_K$. Find the largest q s.t. $p_1 < q$ implies $l_1 > 1$, where l_i is the length of the binary Huffman codeword associated with symbol i.

4) **Power loading with a cost** (25 points)
Consider n parallel additive white Gaussian noise (AWGN) point-to-point channels. For the i-th channe (where $i \in \{1, 2, \ldots, n\}$) the input and output are denoted by X_i and Y_i respectively and the additive Gaussian noise, which is denoted by Z_i, is distributed according to $Z_i \sim N(0, N_i)$. The channels are additive in the sense that:

$$Y_i = X_i + Z_i, \ \forall \ i \in \{1, 2, \ldots, n\}. \tag{1}$$

The channel is illustrated in Fig. 1 The capacity of this setting is given by:

$$C = \frac{1}{2} \sum_{i=1}^{n} \log \left(1 + \frac{P_i}{N_i}\right), \tag{2}$$

which is achieved by taking the input vector $\mathbf{X} = (X_1, X_2, \ldots, X_n) \sim N(0, \mathbf{\Sigma})$, where the

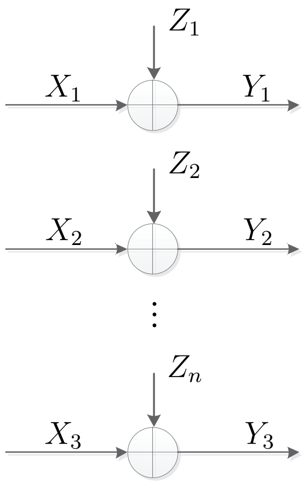

Fig. 1. The n parallel AWGN channels.

covariance is:

$$\Sigma = \begin{pmatrix} P_1 & 0 & \dots & 0 \\ 0 & P_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & P_n \end{pmatrix} \tag{3}$$

and is bound to the power constraint:

$$\sum_{i=1}^{n} P_i \leq P. \tag{4}$$

In order to find the optimal distribution of powers $\{P_i\}_{i=1}^n$ between the channels, the following optimization problem is solved:

$$\max_{\{p_i\}_{i=1}^n} \frac{1}{2} \sum_{i=1}^{n} \log\left(1 + \frac{P_i}{N_i}\right), \tag{5}$$

subject to

$$\begin{gathered} \textstyle\sum_{i=1}^{n} P_i \leq P, \\ P_i \geq 0,\ \forall\ i \in \{1, 2, \dots, n\}. \end{gathered}$$

Solving the optimization in (5) yields the famous ”Water-Filling” solution.
In this question we consider an extension of the classical Water-Filling problem. Here we wish to allocate power to the n parallel channels where the constraint is on the total price (and not power). Lets $\{\beta_i\}_{i=1}^n$ be known, non-negative constants which represent the cost per unit of power in channel i. The transmission is bound to a total cost constraint:

$$\sum_{i=1}^{n} \beta_i P_i \leq B. \tag{6}$$

As in the classical problem, the solution must take into account the non-negativity of the power allocated to each channel, i.e. we must have $P_i \geq 0$ for every $i \in \{1, 2, \dots, n\}$.

a) State the optimization problem (and the constraints) for the extended setting.
b) Construct the Lagrangian function for this maximization problem.

c) Solve the optimization problem using Lagrange multipliers technique in order to find P_i as a function of (β_i, N_i), for every $i \in \{1, 2, \ldots, n\}$.
You may use the function x^+ which is defined by:

$$x^+ = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

d) Explain the interpretation of the price allocation for each channel. How does it differs form the classical Water-Filling solution? How should one choose $\{\beta_i\}_{i=1}^n$ and B so that the solution to the extended problem will reduce into the classical one?

e) Now let us consider only two parallel channels, i.e. $n = 2$. Moreover, it is now given that $B = 10$, $N_1 = 3$, $N_2 = 2$, $\beta_1 = 1$, $\beta_2 = 2$. Find (P_1, P_2) and calculate the capacity of channel.

5) **Entropic Sources** (25 points)
Consider a (not necessarily memoryless) source **Z**. Let $H(Z^n)$ denote the entropy of an n-tuple from this source.
Definition 1: A source **Z** is said to be **subentropic** if the total mass of its most likely $\lfloor 2^{(1+\delta)H(Z^n)} \rfloor$ outcomes goes to 1 as $n \to \infty$ for any $\delta > 0$.
Definition 2: A source is **superentropic** if the total mass of its most likely $\lfloor 2^{(1-\delta)H(Z^n)} \rfloor$ outcomes does not go to 1 as $n \to \infty$ for any $\delta > 0$.
Definition 3: A source is **entropic** if it is both subentropic and superentropic.
Answer the following questions: I.

a) Any source is either subentropic or superentropic or both. **True or False ?**

b) For each of the following sources, classify the source as

   a. subentropic but not superentropic
   b. superentropic but not subentropic
   c. entropic
   d. none of the above.

   i) $Z^n = (X_1, \cdots, X_n)$ is an i.i.d. source with finite alphabet.
   ii) For $0 < q < 1$,

$$P[Z^n = z^n] \;=\; \begin{cases} 1 - q & \text{if } z^n = (0, \cdots 0) \\ q/(2^n - 1) & \text{if } z^n \neq (0, \cdots 0). \end{cases} \tag{7}$$

   iii) Z^n has four types of outcomes
      A) one mass with probability $\frac{1}{2}$;
      B) 2^n masses each with probability $(\frac{1}{2} - \frac{1}{n})2^{-n}$;
      C) $\lfloor \frac{1}{n} 2^{n^2/2} \rfloor$ masses each with probability $2^{-n^2/2}$;
      D) one mass with the remaining probability if $\frac{1}{n} 2^{n^2/2}$ is not integer valued.

Good Luck!