Introduction to Information and Coding Theory September 23, 2014
(Prof. Permuter Haim, Mr. Tal Kopetz, and Mr. Oron Sabag)

Final Exam - Moed Bet
Total time for the exam: 3 hours!

1) **Parallel Gaussian channels (25 Points)** Consider a channel consisting of 2 parallel Gaussian channels, with inputs X_1 and X_2 and outputs given by

$$Y_1 = X_1 + Z_1,$$
$$Y_2 = X_2 + Z_2.$$

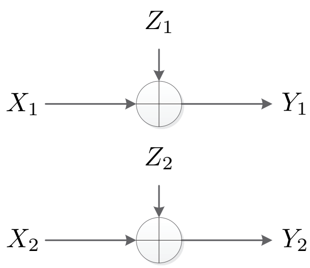

Fig. 1: Parallel Gaussian channels.

The random variables Z_1 and Z_2 are independent of each other and of the inputs, and have the variances σ_1^2 and σ_2^2 respectively, with $\sigma_1^2 < \sigma_2^2$.

a) Suppose $X_1 = X_2 = X$ and we have the power constraint $E[X^2] \leq P$. At the receiver, an output $Y = Y_1 + Y_2$ is generated. What is the capacity C_a of the resulting channel with X as the input and Y as the output?

b) Suppose that we still have to transmit the same signal on both channels, but we can now choose how to distribute the power between the channels, i.e. $X_1 = aX$ and $X_2 = bX$. The new constraint is $E[X_1^2] + E[X_2^2] \leq 2P$. What is the capacity, C_b, of this channel with X as the input and (Y_1, Y_2) as the output? Which a and b achieve that capacity?

c) We now assume that Z_1 and Z_2 are dependent, specifically, $Z_2 = 2Z_1$. As in subsection b, we can choose how to distribute the power between the channels, i.e. $X_1 = aX$ and $X_2 = bX$ under the power constraint $E[X_1^2] + E[X_2^2] \leq 2P$. The outputs of the channels are given by

$$Y_1 = aX + Z_1,$$
$$Y_2 = bX + 2Z_1.$$

What is the capacity, C_c, of this channel with X as the input and (Y_1, Y_2) as the output? Which a and b achieve that capacity?

2) **Erasure Channel with Feedback (25 Points)**

Let X be a random variable that is uniformly distributed in the interval $[0, 1]$.

a) Is it possible to generate from one realization of X a binary random variable that is distributed Bernoulli(p)? If yes, prove it.

Consider the erasure channel with feedback as depicted in Fig. 2.

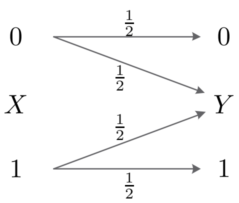

Fig. 2: Erasure Channel with erasure parameter $\epsilon = \frac{1}{2}$.

A student provided the following coding scheme for the erasure channel: The message M has a finite alphabet of size 2^{nR} and the points of the alphabet are distributed uniformly in the interval $[0, 1]$, i.e. $m \in \{k \cdot \frac{1}{2^{nR}}\}_{k=0}^{2^{nR}-1}$. Fix a parameter $p \in [0, 1]$. The interval $[0, 1]$ is divided into two parts, $[0, p)$ and $[p, 1)$. In the first transmission, if $m \in [0, p)$ the encoder transmits $'0'$ and if $m \in [p, 1)$ the encoder transmits $'1'$.

Upon a successful transmission, the decoder knows the interval where the message falls and this interval is divided again with the same parameter p. If the transmission failed, the encoder repeats the transmitted bit until a successful transmission is established.

b) What is the rate of the proposed coding scheme.

c) Can this coding scheme achieve the capacity of the erasure channel? If yes, prove it.

3) **Secure Network Coding (25 Points)** Consider the network depicted in Fig. 3.
The source S would like to transmit a message W to the terminal T. The message, W, is a random binary vector of length k, i.e. $W = [w_1, w_2, \ldots, w_k]$, where each element w_i is distributed $w_i \sim Bern(0.5)$. Each link in the network can carry only one bit, the bit b_1 is transmitted at the upper link and b_2 through the lower link. A spy acquires, E, which is a random observation of one of the links. We know that $E = b_1$ with probability p and $E = b_2$ with probability $1 - p$.
Our goal is to maximize the amount of information that is transmitted to the terminal, while preserving that $I(E; W) = 0$ which means zero information available to the spy. All codebooks are known to the encoder, decoder, and to the spy.

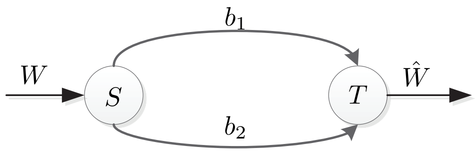

Fig. 3: Network with one source and one terminal.

a) Find $I(A; A \oplus B)$ for $A \sim Bern(\alpha)$ and $B \sim Bern(0.5)$.
b) What is the maximum number of bits (maximum k) that the source S can send to node T in one transmission assuming that the spy is NOT listening, i.e., $I(E; W)$ is NOT necessarily 0?
Provide an achievability scheme and a converse.
c) What is the maximum number of bits (maximum k) that the source S can send to node T in one transmission while preserving $I(E; W) = 0$ for any value of p?
Provide an achievability scheme and a converse. For the achievablitiy, you may use an additional RV which is distributed uniformly in the interval $[0, 1]$ and is drawn at the encoder S.
d) Is there a specific value of p which will allow us to send more bits? If yes, prove and if no give a counter example.

4) **Bhattacharyya distance (25 Points)** For two probability density functions, $f(x)$ and $g(x)$, define the *Bhattacharyya distance* between f and g as

$$D_b(f, g) = -\log\left(\int_{-\infty}^{\infty} \sqrt{f(x)g(x)}dx\right) \tag{1}$$

The Bhattacharyya distance is widely used in various fields such as machine learning, statistics, and more.
For this question, the base of the logarithm is 2.

a) Prove that $0 \leq D_b(f, g) \leq \infty$.
When does $D_b(f, g) = 0$? When does $D_b(f, g) = \infty$?
Hint: You can use the Cauchy Schwarz inequality: for any two real valued functions $f_1(x), f_2(x)$, we have:

$$\left|\int_{-\infty}^{\infty} f_1(x)f_2(x)dx\right|^2 \leq \int_{-\infty}^{\infty} |f_1(x)|^2 dx \int_{-\infty}^{\infty} |f_2(x)|^2 dx. \tag{2}$$

b) We define the differential divergence as follows:

$$D(f||g) = \int_{-\infty}^{\infty} f(x) \log \frac{f(x)}{g(x)} dx. \tag{3}$$

Let $h(x)$ be a third probability density function. Show that

$$D_b(f, g) \leq \frac{1}{2}\left(D(h||f) + D(h||g)\right). \tag{4}$$

c) Assume that $D_b(f, g) < \infty$. For what $h(x)$, there is an equality in Eq.(4)?
d) Does the following inequality holds?

$$2D_b(f, g) \leq \min\{D(g||f), D(f||g)\} \tag{5}$$

If yes, prove it, if not, give a counter example.

Good Luck!