MAE1 206 : Information Engineering

These exercises are destined to students in the final stage of their exam prepaparation. They are based on the Intended Learning Outcomes (ILO) listed in the course webpage, so prior to each exercise, make sure the ILO are well understood. All exercises can be treated separately, however, following the suggested order allows a better understanding. These exercises were not timed to fit in 2 hours. Solutions will be provided later on the course webpage.

Exercise 1 : On Gaussian mixtures

A very common probability distribution when dealing with information engineering (be it communications over noisy channels or even Inference in Machine Learning applications) is the Gaussian mixture distribution. In this exercise, we explore this probability distribution and characterize its properties through the information engineering tools of the course. Throughout the exercise, you will need each of the following tools.

Evaluated ILO :

- List and identify basic probability distributions (discrete and continuous)
- Define and compute the expectation and variance of a random variable
- Define conditional and marginal distributions from joint distributions
- Define the entropy, mutual information
- Compute the entropy of random sources
- Compute the mutual information for random channels
- Relate to the notion of bit-rates (bits/sec)

Let X be a binary random variable which can take two possible values $+1$ and -1 with pmf P_X defined as

$$P_X(x) = \begin{cases} p & \text{if } x = 1, \\ 1-p & \text{if } x = -1. \end{cases} \tag{1}$$

Let W be a Gaussian random variable with mean $\mu = 0$ and variance σ^2, which is independent of X and whose pdf P_W is given by

$$P_W(w) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{w^2}{2\sigma^2}\right). \tag{2}$$

Let Y be the random variable given by

$$Y = X + W. \tag{3}$$

Question 1

Which of the random variables X, W and Y are discrete random variables, and which of them are continuous random variables ?

In the following, we will the term probability distribution to denote either a pmf or a pdf.

Question 2

Show that the conditional distribution $P_{Y|X}(y|x)$ for all $y \in \mathbb{R}$ and $x \in \{-1, 1\}$ is given by

$$P_{Y|X}(y|x) = P_W(y - x). \tag{4}$$

This conditional distribution is often termed the channel distribution.

Question 3

Prove that the channel distribution $P_Y(y)$ writes as

$$P_Y(y) = pP_W(y-1) + (1-p)P_W(y+1). \tag{5}$$

Question 4

Justify the name Gaussian mixture.

Question 5

Give (without computing any integration) the mean and variance of the variable Y.

Question 6

Represent graphically this probability distribution $P_Y(y)$ as a function of y.

> This channel model is termed in literature the Binary Input Additive White Gaussian Noise (BI-AWGN) channel, and models the so-called BPSK transmission over a Gaussian channel. It is one of the most common probabilistic models for the design of error correction codes, and the dimensioning of satellite transmission links for instance.

Question 7

Assume in the following that $p = 0.5$, i.e., X is a uniform binary random variable. Prove that the mutual information $I(X;Y)$ writes as :

$$I(X;Y) = 1 - \int_{\mathbb{R}} P_Y(y) H_2(P_{X|Y}(0|y))\, dy, \tag{6}$$

where H_2 is the binary entropy function defined by

$$H_2(p) = -p\log_2(p) - (1-p)\log_2(1-p). \tag{7}$$

Question 8

What is the unit of the mutual information ? what does it correspond to ?

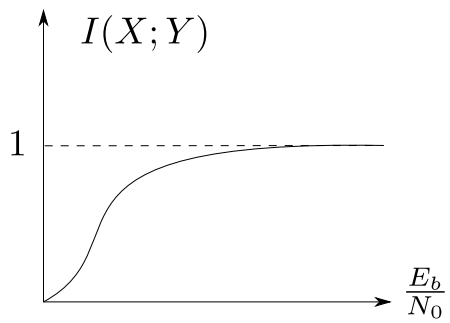
Question 9

Show that the previous mutual information verifies that $I(X;Y) \leq 1$, and interpret this result.

One can show that the mutual information $I(X;Y)$ depends only on the signal to noise ratio $\frac{E_b}{N_0}$ defined by

$$\frac{E_b}{N_0} = \frac{1}{\sigma^2}. \tag{8}$$

This is due to the fact that $P_{Y|X}(y|x)$ and $P_Y(y)$ are both functions of only y and σ^2. In the following, we give the graphical representation of this mutual information as a function of the ratio $\frac{E_b}{N_0}$.



Question 10

Justify (qualitatively) the value at high $\frac{E_b}{N_0}$ given by

$$I(X;Y) = 1. \tag{9}$$

Question 11

Justify (qualitatively) the value at low $\frac{E_b}{N_0}$ given by

$$I(X;Y) = 0. \tag{10}$$

Question 12

Justify why the mutual information is increasing in $\frac{E_b}{N_0}$.

Question 13

If X was a four-valued random variable $\mathcal{X} = \{-3, -1, 1, 3\}$, what would be the maximum value achievable by the mutual information.

Exercise 2 : On compression of m-addic sources

In this exercise, we aim at proving that Huffman codes are strictly optimal for a family of random srouces called the m-addic sources.

Evaluated ILO :

- Define lossy and lossless compression
- Compute the entropy of a random source
- Describe the principle of lossless compression
- Define an entropy code
- Describe Huffman coding
- Show that Huffman coding achieves the entropy

Question 1

What type of compression would you recommend for images (lossless or lossy) ? Justify your answer.

In this exercise, we will focus on lossless compression applied to grascale images. To this end, assume that we have a (64×64) grayscale image in which each pixel is represented with an intensity (integer valued) coded over 3 bits.

Question 2

How may intensity levels are possible for each pixel ?

Assume that we can model the image by a random variable I whose probability distribution P_I is given by the following table :

Intesity level i	0	1	2	3	4	5	6	7
Probability $P_I(i)$	$\frac{1}{32}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$

Question 3

Compute the entropy of this source $H(I)$.

Question 4

What is the unit of entropy ? Interpret the notion of entropy.

In order to compress this source losslessly, we will resort to entropy coding, and more specifically, to Huffman coding.

Question 5

Construct the Huffman tree of this source.

Question 6

Construct an encoding codebook based on the obtained Huffman tree, as follows :

Intesity level i	0	1	2	3	4	5	6	7
Binary code $c(i)$								

Question 7

Comment on the respective lengths of the obtained codewords.

Question 8

Compute the average length of the obtained code.

Question 9

What do you notice ? Comment your finding.

Exercise 3 : On linear block codes

In this exercise, we investigate a family of linear block codes, and compare their performances.

Evaluated ILO :
- Define Forward Error Correction (FEC) and its purpose
- Define linear block codes using a generator matrix or a parity matrix
- Assess the performance of a linear block code (BER and BLER)
- Compare different linear block codes (un-coded, Hamming, ...)

Consider a first code $\mathcal{C}_1$ with $k = 3$ input bits, and $n = 6$ output bits.

Question 1

Given the following generator matrix of the code, what is the code rate R_1 ?

$$G_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

Question 2

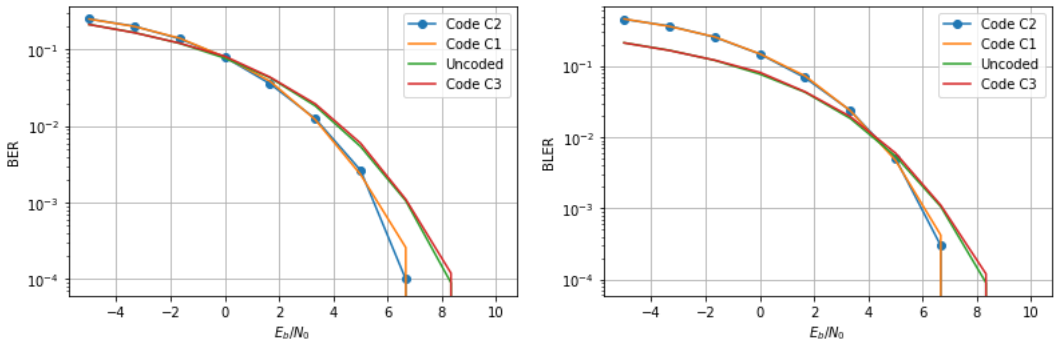
To which interval does the rate of a code belong ?

Question 3

What is the effect of decreasing the code rate on the probability of error ?

Question 4

What is the number of possible output codewords ?



Question 5

Consider the following parity check matrix of the code.

$$H_1 = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 \end{pmatrix}$$

Which are, among the given choices, valid codewords ?

- ◯ Codeword $c = [1\ 0\ 1\ 0\ 1\ 0]$
- ◯ Codeword $c = [0\ 0\ 0\ 0\ 0\ 0]$
- ◯ Codeword $c = [0\ 0\ 0\ 1\ 1\ 1]$
- ◯ Codeword $c = [1\ 1\ 1\ 0\ 1\ 0]$

We wish to desing an error correction code to communicate over a BI-AWGN channel as defined in Exercise 1, with an $\frac{E_b}{N_0}$ close to 10 dB. To this end, we consider three disctinct error correction codes $\mathcal{C}_1$ as defined previously, and two other codes $\mathcal{C}_2$ and $\mathcal{C}_3$ defined by the following generator matrices

$$G_2 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix} \text{ and } G_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Consider the Bit Error Rate (BER) and BLock Error Rate (BLER) curves given in the following.

Question 6

Interpret the results of the BER and BLER curves of both codes, with respect to the $\frac{E_b}{N_0}$.

Question 7

Which code would you suggest to use for the required value of $\frac{E_b}{N_0}$? Comment on your choice.

Exercise 4 : On security proofs of the OTP

In this exercise, we aim at proving the perfect secrecy of the one-time pad, and explain why violating the one-time pad leads to security breaches.

Evaluated ILO :
- Define conditional and marginal distributions from joint distributions
- Compute mutual information of discrete sources
- Describe the principle of a stream cipher
- Describe and justify the one-time pad principle
- Assess the effect of using the same cipher twice

Let M be a binary random variable with a Bernoulli(p) distribution which models a binary message to be transmitted, i.e.,

$$P_M(m) = \begin{cases} p & \text{if } m = 1, \\ 1-p & \text{if } m = 0. \end{cases} \tag{11}$$

Let K be a uniform binary random variable, independent of M, which models an encryption key and whose probability distribution is given by

$$P_K(0) = P_K(1) = \frac{1}{2}. \tag{12}$$

Let C be a binary random variable which models the encryption of the message M with the key K as follows :

$$C = M \oplus K. \tag{13}$$

The purpose of the first part of the exercise is to prove that $I(M;C) = 0$. To this end, the following questions need solving.

Question 1

To what set of values belongs the random variable C ?

Question 2

Prove that the probability distribution $P_C(c)$ verifies

$$P_C(0) = P_C(1) = \frac{1}{2}. \tag{14}$$

Question 3

Prove that the joint probability distribution $P_{M,C}$ verifies that

$$P_{M,C}(m., c) = P_M(m)P_K(m \oplus c) = \frac{1}{2}P_M(m). \tag{15}$$

Question 4

Justify hence why

$$I(M;C) = 0. \tag{16}$$

This proves why the OTP is indeed perfectly secure. We will see in the following how violating the OTP yields inevitably an information leakage. To this end, let M_1 and M_2 be two binary random variables distributed botih following a distribution P_M which is a Bern(p) (they both have the same distribution but are not equal). And let K be a uniform binary encryption key.

Let C_1 and C_2 be the encrypted messages using the same key K, given by :

$$C_1 = K \oplus M_1 \text{ and } C_2 = K \oplus M_2. \tag{17}$$

In order to assess the information leakage, our aim is to prove that the mutual information $I(M_1M_2; C_1C_2)$, to be read the mutual information between the pair of messages (M_1, M_2) and the pair of messages (C_1, C_2), is non-zero, $I(M_1M_2; C_1C_2) > 0$. To this end, you will need to solve the following questions.

Question 1

Show that the mutual information $I(M_1M_2; C_1C_2)$ writes as

$$I(M_1M_2; C_1C_2) = 2H(M) - H(M_1|C_1C_2) - H(M_2|C_1C_2M_1). \tag{18}$$

Question 2

Show that M_2 is a function of C_1, C_2 and M_1. Hence, conclude that :

$$H(M_2|C_1C_2M_1) = 0. \tag{19}$$

Question 3

By admitting that $H(M_1|C_1C_2) = H(M_1|C_1)$, and using the conclusion of the first part of the exercice, $I(M_1, C_1) = 0$, show that

$$H(M_1|C_1C_2) = H(M_1) = H(M). \tag{20}$$

Question 4

Once all these equalities are proved, prove that $I(M_1M_2; C_1C_2) > 0$.

An interesting fact here, is that even though eahc of the messages M_1 and M_2 is secured individually, if an eavesdropper intercepts both encrypted messages C_1 and C_2, then there exists an information leakage. Hence, individual security does not guarantee zero information leakage, i.e.,

$$I(M_1, C_1) = 0, I(M_2, C_2) = 0 \nRightarrow I(M_1M_2; C_1C_2) = 0, \tag{21}$$

and this is exactly what you noticed during the PBL on cryptography.