ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
School of Computer and Communication Sciences

Handout 34 | Information Theory and Coding
Final Exam | January. 27, 2023

4 problems, 64 points
180 minutes
2 sheets (4 pages) of notes allowed.

Good Luck!

PLEASE WRITE YOUR NAME ON EACH SHEET OF YOUR ANSWERS.

PLEASE WRITE THE SOLUTION OF EACH PROBLEM ON A SEPARATE SHEET.

PROBLEM 1. (12 points) Suppose W is a channel with input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$, and with capacity $C(W)$.

Suppose that p_1 and p_2 are two probability distributions on $\mathcal{X}$, and let q_1 and q_2 denote the corresponding output distributions, i.e., $q_k(y) = \sum_x W(y|x)p_k(x)$. Let I_1 and I_2 denote the mutual information $I(X;Y)$ between the channel input and channel output when X has distribution p_1 and p_2 respectively.

(a) (4 pts) Show that

$$I_1 = \sum_x p_1(x) D\big(W(\cdot|x) \;\|\; q_2\big) - D(q_1 \| q_2).$$

[Here $W(\cdot|x)$ denotes the probability distribution q on $\mathcal{Y}$ with $q(y) = W(y|x)$.]

(b) (4 pts) Suppose that $I_2 = C(W)$, i.e., p_2 is a capacity achieving input distribution. Show that

$$\sum_x p_1(x) D\big(W(\cdot|x) \;\|\; q_2\big) \le C(W).$$

(c) (2 pts) Suppose further that p_1 is also capacity achieving. Show that $q_1 = q_2$.

(d) (2 pts) Consider a deterministic channel with input $x \in \{0, 1, 2\}$ and output $y = \mathbb{1}\{x \neq 0\}$. What is the capacity and which input distributions achieve it?

Problem 2. (16 points) Suppose W_1 and W_2 are two channels with the same input alphabet $\mathcal{X}$ and the same output alphabet $\mathcal{Y}$. Both the channels are discrete and memoryless.

We are asked to design an encoder enc : $\{1,\dots,M\} \to \mathcal{X}^n$ and decoder dec : $\mathcal{Y}^n \to \{1,\dots,M\}$ so that the error probability is small for both of the following two cases: (1) the channel is W_1 for the duration of the transmission, (2) the channel is W_2 for the duration of the transmission.

For this task we adopt a random coding technique (as in class). That is, we pick a distribution p_X, and choose

$$\{\mathrm{Enc}(m)_i : m = 1\dots,M,\ i = 1\ \dots,n\}$$

as i.i.d. random variables each with distribution p_X.

Let $p^{(k)}_{XY}(x,y) = p_X(x)W_k(y|x)$, $k = 1,2$. Let $q^{(k)}_{XY}$ denote the distibution with the same marginals as $p^{(k)}$, but is in product form, i.e., $q^{(k)}(x,y) = p_X(x)p^{(k)}_Y(y)$.

Fix $\epsilon > 0$. Let $T_k = T(n, p^{(k)}_{XY}, \epsilon)$ be the set of ϵ-typical (x^n, y^n) pairs with respect to the distribution $p_X(x)W_k(y|x)$.

(a) (4 pts) Two colleagues suggest two different ideas for the decoder:

- α. decode m if $\mathrm{Enc}(m)$ is the only codeword for which $(\mathrm{Enc}(m), y^n) \in T_1 \cap T_2$.
- β. decode m if $\mathrm{Enc}(m)$ is the only codeword for which $(\mathrm{Enc}(m), y^n) \in T_1 \cup T_2$.

Explain why 'idea α' should *not* be used.

Define the quantity

$$\delta_1 = \Pr\big((X^n,Y^n) \notin T_1\big) + (M-1)\Pr\big((\tilde{X}^n,Y^n) \in T_1\big) + (M-1)\Pr\big((\tilde{X}^n,Y^n) \in T_2\big)$$

where $\{(X_i, Y_i, \tilde{X}_i) : i = 1,\dots,n\}$ are i.i.d. with distribution $p_X(x)W_1(y|x)p_X(\tilde{x})$. Define δ_2 analogously.

(b) (4 pts) Show that when W_1 is the channel, the expected error probability of the random code decoded by 'idea β' is upper bounded by δ_1.

Let I_k denote $I(X;Y)$ when (X,Y) has distribution $p^{(k)}$. Note that $I_k = D(p_k\|q_k)$.

(c) (4 pts) Show that $D(p_2\|q_1) \ge I_2$.

(d) (4 pts) Show that whenever R is strictly less than both I_1 and I_2, for any $\delta > 0$ there exists an encoder and decoder of rate at least R and error probability at most δ under both the cases (of the channel being W_1 or W_2).

[Hint: consider the three terms in the expression for δ_1. We know from class that, ignoring ϵ's, $\Pr((\tilde{X}^n,Y^n) \in T_1) \approx 2^{-nD(p_1\|q_1)}$. What can you say about $\Pr((\tilde{X}^n,Y^n) \in T_2)$ that appears in the third term.]

PROBLEM 3. (18 points) Suppose $G \in \{1, 2, \dots\}$ is a positive integer valued random variable with distribution p_G.

(a) (4 pts) Suppose N is a random variable with $p_N(k) = (1-q)q^{k-1}$, $k = 1, 2, \dots$, with q chosen such that $E[N] = E[G]$. Show that $H(N) - H(G) = D(p_G \| p_N)$.

(b) (2 pts) Fact: for a random variable N as above $H(N) = f(E[N])$ with $f(\mu) = \mu \log \mu - (\mu - 1)\log(\mu - 1)$. Show that $H(G) \le f(E[G])$.

Suppose U is a random variable taking values in the finite alphabet $\mathcal{U} = \{1, \dots, K\}$. Let $p_k = \Pr(U = k)$. We learn the value of U by asking a sequence of questions of the form "Is U equal to u?", until the answer is 'yes'. Let G be the number of questions asked. (Note that the phrase "until the answer is 'yes'" may require us to ask a question whose answer we already know.)

A 'questioning strategy' is a deterministic map from $(p_1, \dots, p_K)$ to the order in which we pose the questions "Is $U = 1$?", ..., "Is $U = K$?".

(c) (4 pts) Suppose $p_1 \ge \cdots \ge p_K$. What questioning strategy minimizes $E[G]$? Justify your answer. [Hint: suppose $p_i < p_j$ but the question "Is $U = i$?" is asked before "Is $U = j$". Show that such a strategy can't be optimal.]

(d) (2 pts) Show that the strategy you found in (a) not only minimizes $E[G]$ but also $\Pr(G > k)$ for every k.

(e) (2 pts) What is the relationship between $H(G)$ and $H(U)$?

Suppose $U_1, U_2, \dots$ is a stationary process with entropy rate H. We use a questioning strategy as above to learn U_1 in G_1 questions. Having learned that $U_1 = u_1$, we know that U_2 is distributed according to $p_{U_2|U_1=u_1}$. We use a questioning stragegy based on this distribution to learn U_2 with G_2 questions. Continuning in this fashion, having already learned $(u_1, \dots, u_{n-1})$, we use the strategy based on the distribution $p_{U_n|U^{n-1}=u^{n-1}}$ to learn U_n in G_n questions.

(f) (4 pts) With the function f as in part (b), show that

$$f(E[G_n|U^{n-1} = u^{n-1}]) \ge H(U_n|U^{n-1} = u^{n-1}),$$

and conclude that $f(E[G_n]) \ge H$. [Hint: f is a concave function.]

Problem 4. (18 points)

Recall that the minimum distance of a binary code $\mathcal{C}$ is defined as

$$d_{\min}(\mathcal{C}) := \min_{\substack{x,x' \in \mathcal{C} \\ x \neq x'}} d_H(x, x').$$

Suppose we are told that for any binary code with blocklength n and minimum distance d or larger, the number of codewords satisfies $M \leq \text{bound}_0(n, d)$.

(a) (4 pts) Show that we can improve such a bound to

$$M \leq \text{bound}_1(n, d) := \min_{0 \leq n' \leq n} 2^{n-n'} \text{bound}_0(n', d).$$

[Hint: classify the M codewords of the code according to their $(n - n')$ bit prefixes.]

(b) (2 pts) Consider the function $\text{bound}_0(n, d)$ that equals to 1 if $d > n$, and equals to ∞ if $d \leq n$. Why is this an upper bound to the number of codewords of a binary code of blocklength n with minimum distance d or more?

(c) (2 pts) What is the improved bound constructed by the method in (a) starting from the bound in (b)? How does this compare with the Singleton bound?

(d) (4 pts) Suppose a binary code $\mathcal{C}$ is of blocklength n and has M codewords. Let $\mathbf{c}_m = (c_{m1}, \ldots, c_{mn})$ denote the m'th codeword ($m = 1, \ldots, M$). Fix an index $i \in \{1, \ldots, n\}$, and let M_i be the number of codewords with $c_{mi} = 1$. Show that

$$\sum_{m=1}^{M} \sum_{m'=1}^{M} \mathbb{1}\{c_{mi} \neq c_{m'i}\} = 2M_i(M - M_i) \leq M^2/2.$$

(e) (4 pts) For $\mathcal{C}$ as in (d) and d denoting its minimum distance, show that $M(M-1)d \leq nM^2/2$.

The inequalty in (e) is equivalent to, $2d \leq \frac{nM}{M-1}$. We can re-arrange this inequality to upper bound M in terms of n and d, but since $\frac{nM}{M-1} > n$, this can only be done when $2d > n$. We thus find $M \leq \lfloor 2d/(2d - n) \rfloor$, valid for $2d > n$.

(f) (2 pts) For $d = 7$ and $n = 14$ compare the bound above to its improvement via (a), and also to the Singleton and Sphere packing bounds. [Possibly useful numerics: $V_3 := \sum_{i=0}^{3} \binom{14}{i} = 470$, $V_6 := \sum_{i=0}^{6} \binom{14}{i} = 6476$, $2^{14}/V_3 = 34.859\ldots$, $2^{14}/V_6 = 2.52\ldots$.]