Problem 1 – Source and Channel Coding (30 Points)

(a) **(10 Points)** Consider a fixed-length source coding setting with rate R, block length n, alphabet $\mathcal{X}$, source sequence $\mathbf{X} \in \mathcal{X}^n$ assumed to be discrete and memoryless according to P_X, estimate $\hat{\mathbf{X}} \in \mathcal{X}^n$, and error probability $P_{\rm e}$.

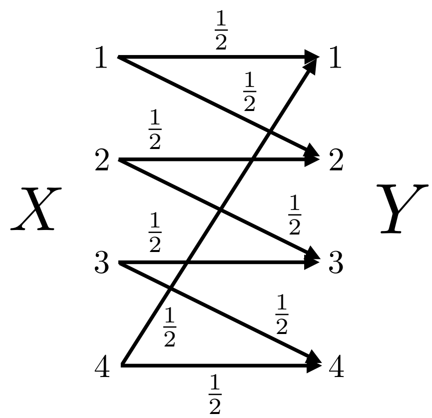
Consider the following chain of inequalities:

$$
\begin{aligned}
nR &\overset{(i)}{\geq} H(\hat{\mathbf{X}}) && \text{\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_} \\
&\overset{(ii)}{\geq} H(\hat{\mathbf{X}}) - H(\hat{\mathbf{X}}|\mathbf{X}) && \text{\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_} \\
&\overset{(iii)}{=} I(\mathbf{X};\hat{\mathbf{X}}) && \text{\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_} \\
&\overset{(iv)}{=} H(\mathbf{X}) - H(\mathbf{X}|\hat{\mathbf{X}}) && \text{\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_} \\
&\overset{(v)}{\geq} H(\mathbf{X}) - nP_{\rm e}\log_2|\mathcal{X}| - 1 && \text{\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_} \\
&\overset{(vi)}{=} nH(X) - nP_{\rm e}\log_2|\mathcal{X}| - 1 && \text{\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_}
\end{aligned}
$$

In the space to the right of each of steps (i)–(vi), write down the most suitable explanation (one only) among the following:

- Definition of entropy
- Definition of mutual information
- Non-negativity of entropy
- Non-negativity of mutual information
- The source is memoryless
- Data processing inequality
- Fano's inequality
- Uniform distribution maximizes entropy
- Conditioning reduces entropy
- Chain rule for mutual information

(b) **(20 Points)** Consider the discrete memoryless channel with alphabets $\mathcal{X} = \mathcal{Y} = \{1, 2, 3, 4\}$ and transition probabilities depicted in the following diagram:



(i) Compute the channel capacity $C = \max_{P_X} I(X;Y)$, showing your working.

(ii) Write down two different capacity-achieving input distributions P_X^* that both attain the maximum in $\max_{P_X} I(X;Y)$.

(iii) Briefly describe a simple method (i.e., encoder and decoder) for transmitting at a positive rate while achieving an error probability of exactly zero.

(Extra space for answering Problem 1(b))

Problem 2 – Discrete and Continuous Information Measures (25 Points)

(a) **(18 Points)** Let X and Y be discrete real-valued random variables with joint probability mass function P_{XY}, and let U and V be continuous real-valued random variables with joint probability density function f_{UV}. Recall that $H(\cdot)$ denotes the entropy for discrete random variables, and $h(\cdot)$ denotes the differential entropy for continuous random variables.

For each of the following, either explain why the given statement is always true, or explain why it is sometimes false. In your answers, you may make use of any statement proved in the lectures, unless it is the exact statement in the question.

(Hint: If done well, each of these can be answered correctly in 1 or 2 sentences.)

(i) $I(X;Y) \le H(X)$

(ii) $I(U;V) \le h(U)$

(iii) $H(X) = H(cX)$ for any fixed constant $c > 0$

(iv) $h(U) = h(cU)$ for any fixed constant $c > 0$

(v) $H(X) \le \frac{1}{2}\log_2\big(2\pi e\,\mathbb{E}[X^2]\big)$

(vi) $h(U) \le \frac{1}{2}\log_2\big(2\pi e\,\mathbb{E}[U^2]\big)$

(Extra space for answering Problem 2(a))

(b) **(7 Points)** The exponential distribution has a probability density function given by

$$f_{\exp}(u) = \begin{cases} \lambda e^{-\lambda u} & u \geq 0 \\ 0 & \text{otherwise,} \end{cases}$$

where $\lambda > 0$ is a parameter. The mean of this distribution is $\frac{1}{\lambda}$ (you do not need to prove this). Show that any non-negative valued random variable with some density function $f_U(u)$ and mean $\frac{1}{\lambda}$ must have a differential entropy no higher than that of the exponential distribution.

Problem 3 – Linear Codes (25 Points)

Consider the linear channel code with generator matrix

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

All of questions (a)–(c) below concern this particular code.

(a) **(10 Points)** Write down all of the codewords of the code, as well as the parity check matrix $\mathbf{H}$ and the minimum distance $d_{\min}$.

(b) **(5 Points)** Let $d_{\mathrm{H}}(\cdot,\cdot)$ denote the Hamming distance, and again let $d_{\min}$ be the minimum distance of the code. Is it possible to find three different codewords $\mathbf{x}$, $\mathbf{x}'$, and $\mathbf{x}''$ of this code such that both $d_{\mathrm{H}}(\mathbf{x},\mathbf{x}') = d_{\min}$ and $d_{\mathrm{H}}(\mathbf{x},\mathbf{x}'') = d_{\min}$? Explain.

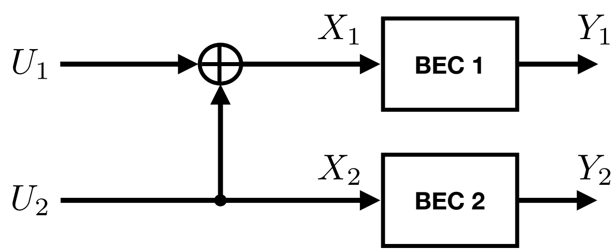
(Note: It is important to observe that both of these $d_{\mathrm{H}}(\cdot,\cdot)$ expressions have the same first argument $\mathbf{x}$.)

(c) **(10 Points)** Let $\mathbf{u} = (u_1, u_2, u_3)$ be a triplet of information bits, and let $\mathbf{x} = \mathbf{uG}$ be the resulting codeword. Suppose that a noise vector $\mathbf{z} = (z_1, z_2, z_3, z_4, z_5, z_6)$ is generated by drawing an index $i \in \{1, 2, 3, 4, 5, 6\}$ uniformly at random and setting $z_i = 1$, then setting all other z_j $(j \neq i)$ to zero. The resulting output vector $\mathbf{y} = \mathbf{x} \oplus \mathbf{z}$ (with modulo-2 addition) is passed to a decoder, who also knows $\mathbf{G}$ and $\mathbf{H}$. Describe a decoder based on syndrome decoding that is able to recover $\mathbf{u}$ with success probability $5/6$.

Problem 4 – A Challenging Calculation (20 Points)

Consider the setup shown in the following illustration, where:

- The random variables U_1, U_2, X_1, X_2 take values on $\{0, 1\}$, whereas Y_1 and Y_2 take values on $\{0, e, 1\}$ with e representing an "erasure";
- U_1 and U_2 are independent, and equal 0 or 1 with probability $\frac{1}{2}$ each;
- We have $X_2 = U_2$, and $X_1 = U_1 \oplus U_2$, with $\oplus$ denoting modulo-2 addition;
- "BEC 1" and "BEC 2" are binary erasure channels, each having transition law $\mathbb{P}[Y_i = X_i] = 1 - \epsilon$ and $\mathbb{P}[Y_i = e] = \epsilon$ (for some $\epsilon \in (0, 1)$) with independence between the two channels.



We can express the joint mutual information $I(U_1, U_2; Y_1, Y_2)$ using the chain rule as

$$I(U_1, U_2; Y_1, Y_2) = I(U_1; Y_1, Y_2) + I(U_2; Y_1, Y_2 | U_1),$$

By carefully using the assumptions in the above four dot points, find exact expressions for both $I(U_1; Y_1, Y_2)$ and $I(U_2; Y_1, Y_2 | U_1)$, writing your answer in terms of the erasure probability ϵ.

(Extra space for answering Problem 4)

[Use this page for any extra working if you run out of space. You must clearly write "See final pages" for any question continued here, and here you must clearly indicate each exact question and part (e.g., 2(c)).]

[Use this page for any extra working if you run out of space. You must clearly write "See final pages" for any question continued here, and here you must clearly indicate each exact question and part (e.g., 2(c)).]

END OF PAPER