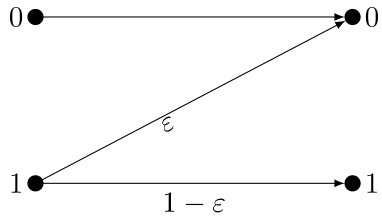
EE5139/EE6139
Homework 9

Semester I, AY2021/2022
Marco Tomamichel

Exercise 9.1 Z-channel (EE5139)

A Z channel is a binary channel with conditional pmf $p(0|0) = 1$, $p(0|1) = \epsilon$.



Suppose $\epsilon = 1/2$, compute the channel mutual information.

Exercise 9.2 Type Classes (EE5139)

Let X be a random variable on $\mathcal{X}$ with pmf p_X. The set of sequences of *type* $\lambda \in \mathcal{P}(\mathcal{X})$ is defined as

$$\mathcal{T}^{(n)}(\lambda) := \{\boldsymbol{x} \in \mathcal{X}^n : f_{\boldsymbol{x}} = \lambda\},$$

where $\mathcal{P}(\mathcal{X})$ stands for the set of all distributions over $\mathcal{X}$, and for a given sequence $\boldsymbol{x}$, $f_{\boldsymbol{x}}$ stands for the induced empirical distribution, *i.e.*, $f_{\boldsymbol{x}}(x) := n^{-1} \cdot \sum_{i=1}^{n} \delta_{x_i,x}$. Let X^n be n i.i.d. copies of X, *i.e.*, $p_{X^n} = p_X^{\otimes n}$. Show that the probability that X^n being any sequence $\boldsymbol{x} \in \mathcal{X}^n$ depends only on its type and p_X, namely

$$p_{X^n}(\boldsymbol{x}) = 2^{-n(H(f_{\boldsymbol{x}})+D(f_{\boldsymbol{x}}\|p_X))}.$$

Exercise 9.3 Channel Coding and List Decoding (EE6139)

In class, we saw that for all rates R below capacity C, there exists a sequence of $(2^{nR}, n)$-codes such that the average error probability tends to zero. Now, suppose we allow the decoder to output a list of 2^{nL} number of messages (instead of one), and decoding is considered successful if and only if the transmitted message is in the list. Show that for all rates $R < C$, there exists a sequence of $(\lceil 2^{n(R+L)} \rceil, n)$-codes[1] such that the average probability of error tends to zero.
Hint: Consider the joint typical set as follows

$$\mathcal{A}_\epsilon^{(n)}(X,Y) = \left\{ (\boldsymbol{x}^n, \boldsymbol{y}^n) \in \mathcal{X}^n \times \mathcal{Y}^n \;\middle|\; \begin{array}{r} \left| \frac{1}{n} \sum_{i=1}^{n} \log \frac{1}{p(x_i)} - H(X) \right| \le \epsilon, \\ \left| \frac{1}{n} \sum_{i=1}^{n} \log \frac{1}{p(y_i)} - H(Y) \right| \le \epsilon, \\ \left| \frac{1}{n} \sum_{i=1}^{n} \log \frac{1}{p(x_i, y_i)} - H(X,Y) \right| \le \epsilon \end{array} \right\}.$$

For any $\epsilon > 0$, the jointly typical sequences satisfy the following properties: if $\tilde{X}^n, \tilde{Y}^n$ are independent, $\tilde{X}^n \sim p^n(x)$, $\tilde{Y}^n \sim p^n(y)$, we have

- $\Pr[(\tilde{X}^n, \tilde{Y}^n) \in \mathcal{A}_\epsilon^{(n)}(X,Y)] \le 2^{-n(I(X;Y)-3\epsilon)}$,

---

[1] In this case, a (M, n)-code is comprised of an encoder $e : \mathcal{M} \to \mathcal{X}^n$ and decoder $d : \mathcal{Y}^n \to \mathcal{P}(\mathcal{M})$ where $|\mathcal{M}| = M$.

- $\Pr[(\tilde{X}^n, \tilde{Y}^n) \in \mathcal{A}_\epsilon^{(n)}(X,Y)] \geq (1-\epsilon)2^{-n(I(X;Y)+3\epsilon)}$.

One may consider a random encoder $e : w \mapsto X^n(w) \in \mathcal{X}^n$; and a decoder, upon receiving $\boldsymbol{y} \in \mathcal{Y}^n$, outputs a list of $\tilde{w}$'s such that $(e(\tilde{w}), \boldsymbol{y})$ is jointly typical. (Question: What is/are the error event(s)?)

Exercise 9.4 Independently generated codebooks (EE6139)

Let $(X,Y) \sim p(x,y)$, and let $p(x)$ and $p(y)$ be their marginals. Consider two randomly and independently generated codebooks $\mathcal{C}_1 = \{X^n(1), \ldots, X^n(2^{nR_1})\}$ and $\mathcal{C}_2 = \{Y^n(1), \ldots, Y^n(2^{nR_2})\}$. The codewords of $\mathcal{C}_1$ are generated independently each according to $\prod_{i=1}^n p_X(x_i)$, and the codewords for $\mathcal{C}_2$ are generated independently according to $\prod_{i=1}^n p_Y(y_i)$. Define the set

$$\mathcal{C} = \{(x^n, y^n) \in \mathcal{C}_1 \times \mathcal{C}_2 : (x^n, y^n) \in \mathcal{A}_\epsilon^{(n)}(X,Y)\},$$

where $\mathcal{A}_\epsilon^{(n)}$ has been defined in the hint for Exercise 9.3. Show that

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}\left[|\mathcal{C}|\right] = R_1 + R_2 - I(X;Y).$$

Exercise 9.5 Shared Randomness does not increase capacity (EE5139)

Suppose that in the definition of the $(2^{nR}, n)$ code for the DMC $p(y|x)$, we allow the encoder and the decoder to use random mappings. Specifically, let W be an arbitrary random variable independent of the message M and the channel, *i.e.*, $p(y_i|x^i, y^{i-1}, m, w) = p_{Y|X}(y_i|x_i)$ for $i \in [1:n]$. The encoder generates a codeword $x^n(m, W), m \in [1 : 2^{nR}]$, and the decoder generates an estimate $\hat{m}(y^n, W)$. Show that this randomization does not increase the capacity of the DMC.