EE 231A
Information Theory
Instructor: Rick Wesel

Handout #9, Problem Set 4
Tuesday, April 21 2020
Due: Tuesday, April 28, 2020

105 pts
Reading: Chapter 7 of *Elements of Information Theory*

Lectures 7: Definition and Computation of Channel Capacity

1. (9 pts) *An additive noise channel.* Find the channel capacity of the following discrete memoryless channel:

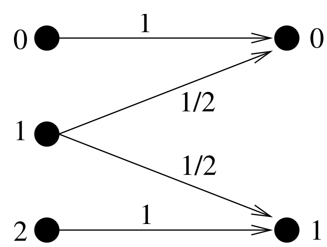
$$Y = X + Z\,. \tag{1}$$

where $\Pr\{Z = 0\} = \Pr\{Z = a\} = \frac{1}{2}$. The alphabet for x is $\mathbf{X} = \{0, 1\}$. Assume that Z is independent of X.

Observe that the channel capacity depends on the value of a.

2. (8 pts) *Inverse Erasure Channel.*

Find the capacity of the inverse erasure channel shown below.



3. (10 pts) *Cyclic Symmetry*

In class we defined a channel to have cyclic symmetry if the mutual information is invariant to cyclic shifts of the input distribution. This problem gives a matrix form of this definition.

Consider the channel transition matrix P where the entry in the xth row and the yth column denotes the conditional probability $p(y|x)$.

Matrix Conditions for Cyclic Symmetry The channel described by P has cyclic symmetry if the following two conditions are satisfied:

- All the rows of P are permutations of each other.
- The set C of columns of P can be separated into a certain collection of mutually exclusive, collectively exhaustive subsets S_i. (i.e. The equations below are satisfied:)

$$\bigcup_i S_i = C \tag{2}$$

$$\text{if } S_i \neq S_j \text{ then } S_i \bigcap S_j = \emptyset\,, \tag{3}$$

  such that each subset S_i may be completely constructed from any one element of S_i as follows: S_i contains exactly one instance of each cyclic shift of that element, and nothing else. Note that the elements of both C and S_i are columns.

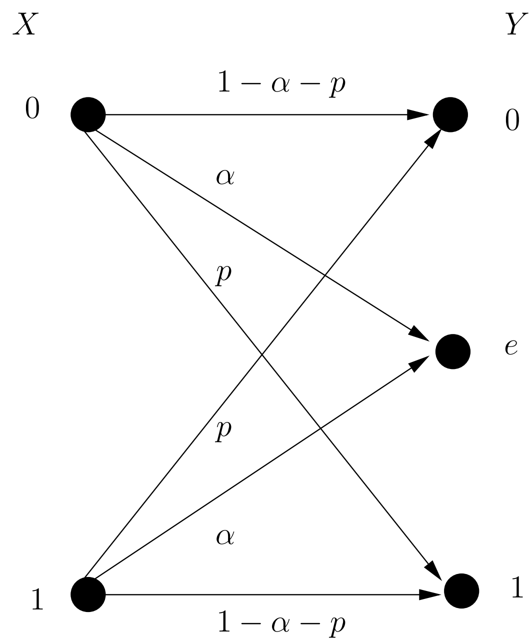
(a) (1 pt) Give P for the binary erasure channel.

(b) (2 pts) For the binary erasure channel, decompose the columns of P into the subsets S_i described above.

(c) (7 pts) Prove that the **Matrix conditions for Cyclic Symmetry** described above are indeed sufficient for a channel to have cyclic symmetry. While the first two parts were specific to the binary erasure channel, this proof needs to work for *any* discrete memoryless channel.

4. (15 pts) *Errors, Erasures, and Symmetry.*

Consider a binary channel that has *both* erasures and symmetric bit errors. This channel is illustrated below.



(a) (3 pts) Write down the transition matrix for this channel.

(b) (3 pts) Does this channel satisfy the matrix conditions for cyclic symmetry? Explain your answer.

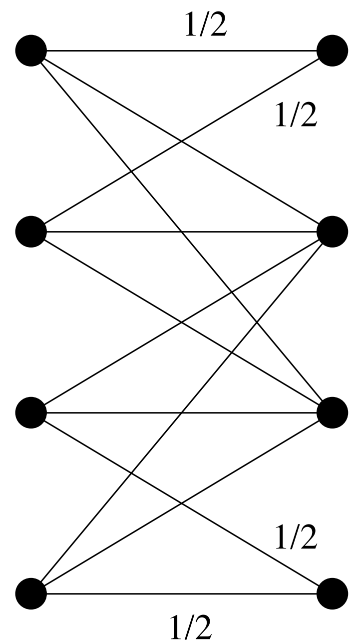
(c) (3 pts) Does this channel satisfy the matrix conditions for weak symmetry? Explain your answer.

(d) (3 pts) Find the capacity of this errors and erasures channel. Feel free to use the notation $h_3(p_1, p_2, p_3)$ for the entropy of a discrete random variable that takes on three values with probabilities p_1, p_2, p_3.

(e) (3 pts) Show how your capacity expression simplifies to other channel capacities that we have studied when α or p are set to zero.

5. (9 pts) *Symmetric Channel?*

Consider the channel illustrated below where branches have probabilities of $1/2$ or $1/4$. The eight unlabeled branches have probability $1/4$.



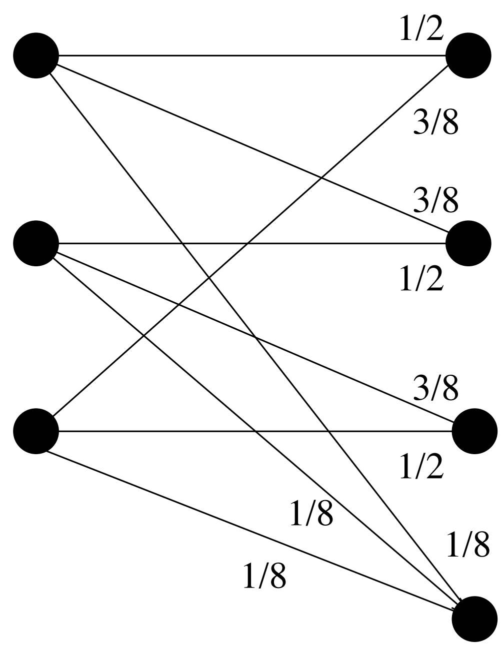
(a) (3 pts) State the conditions for a channel to be weakly symmetric. Write down the matrix of conditional probabilities for this channel. Does this channel meet the conditions to be weakly symmetric?

(b) (3 pts) Is the mutual information of this channel invariant to shifts in the input probability distribution? If yes, give a brief justification. If no, give a counterexample.

(c) (3 pts) What is the capacity of this channel?

6. (9 pts) *Symmetric Channel? (Part II)*

Consider the channel illustrated below where branches have probabilities as labeled.



PLEASE LEAVE ANY CAPACITIES YOU COMPUTE IN TERMS SUCH AS $H(p_1, p_2, p_3)$. NO NEED TO REDUCE RESULTS WITH A CALCULATOR.

(a) (3 pts) State the conditions for a channel to be weakly symmetric. Write down the matrix of conditional probabilities for this channel. Does this channel meet the conditions to be weakly symmetric? If so compute the capacity using weak symmetry.

(b) (3 pts) Is the mutual information of this channel invariant to cyclic shifts in the input probability distribution? If yes, give a brief justification and compute the capacity. If no, give a counterexample.

(c) (3 pts) Compute the capacity using the technique of finding an upper bound and then achieving it.

Lectures 8: Proof of Achievability of Channel Capacity

7. (15 pts) *Joint Typicality.* MS S98

The sequence pair (x_1^n, y_1^n) is drawn i.i.d. according to the p.m.f.

$$p(x^n, y^n) = \prod_{i=1}^{n} p(x_i, y_i)\,. \tag{4}$$

The sequence pair (x_2^n, y_2^n) is drawn in the same way, but independently of (x_1^n, y_1^n).

(a) What is the limit as $n \to \infty$ of the probability that (x_1^n, y_1^n) is in the set $A_\epsilon^{(n)}$ of *jointly* typical sequences?

(b) Give an upper bound on the probability that $(x_1^n, y_2^n) \in A_\epsilon^{(n)}$. The bound should be a function of n and ϵ.

(c) For X equally likely to be 0 or 1 and for the $p(y|x)$ of a BSC with transition probability $p = 0.03113$, evaluate this upper bound for $n = 100$ and $\epsilon = 0.1$.

(d) What is the limit as $n \to \infty$ of the probability that (x_1^n, y_2^n) is in the set $A_\epsilon^{(n)}$ of jointly typical sequences?

(e) How is this related to the probability of error with typical set decoding?

8. (10 pts) *Triple Typicality.* The triple sequence $(x^n,\ y^n,\ z^n)$ is a member of the triple typical set $A_\epsilon^{(n)}$ for the joint distribution $p(x,y,z)$ if the following inequalities are satisfied:

$$\left| -\frac{1}{n}\log p(x^n) - H(X) \right| < \epsilon \tag{5}$$

$$\left| -\frac{1}{n}\log p(y^n) - H(Y) \right| < \epsilon \tag{6}$$

$$\left| -\frac{1}{n}\log p(z^n) - H(Z) \right| < \epsilon \tag{7}$$

$$\left| -\frac{1}{n}\log p(x^n, y^n) - H(X,Y) \right| < \epsilon \tag{8}$$

$$\left| -\frac{1}{n}\log p(x^n, z^n) - H(X,Z) \right| < \epsilon \tag{9}$$

$$\left| -\frac{1}{n}\log p(y^n, z^n) - H(Y,Z) \right| < \epsilon \tag{10}$$

$$\left| -\frac{1}{n}\log p(x^n, y^n, z^n) - H(X,Y,Z) \right| < \epsilon \tag{11}$$

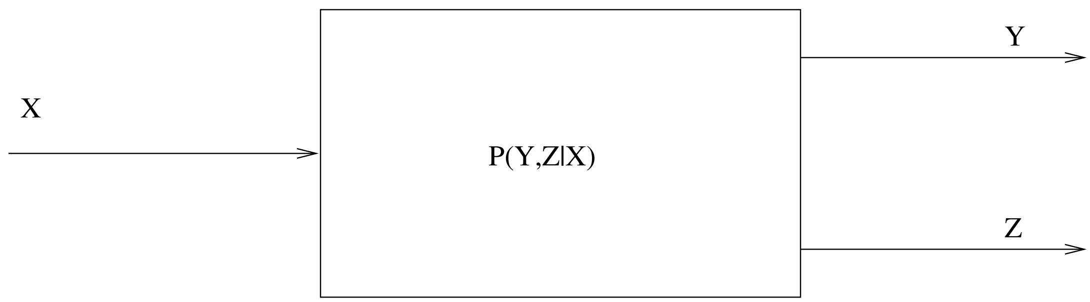
The marginal distributions above are those implied by the joint distribution $p(x,y,z)$.

(a) Assuming (x^n, y^n, z^n) are drawn according to $p(x,y,z)$, What does $Pr(A_\epsilon^{(n)})$ converge to as $n \to \infty$? Justify your answer briefly.

(b) Suppose that the pair of sequences (x^n, y^n) are drawn according to $p(x,y)$ and that the sequence z^n is drawn according to $p(z)$ **independently** of (x^n, y^n). Derive an upper bound on the probability that the triple (x^n, y^n, z^n) generated in this way is a member of the triple typical set $A_\epsilon^{(n)}$. Justify each step in your derivation. *Hint:* Note that $I(X,Y;Z) = H(X,Y) + H(Z) - H(X,Y,Z)$.

9. (10 pts) *Joint Typicality on the Three-Way Channel.*

Consider a discrete memoryless channel with one input X, but *two* outputs, Y, and Z, that both contain information about X and are in general dependent. This channel is illustrated below.



(a) Find (and demonstrate) an upper bound on the probability that X^n, Y^n, Z^n is a member of the triple typical set if X^n and (Y^n, Z^n) are chosen independently, X^n according to $P(X)$ and (Y^n, Z^n) according to $P(Y, Z)$. Your upper bound will involve a term ϵ that converges to zero as $n \longrightarrow \infty$

Recall that the triple typical set is defined as follows:

The triple sequence (x^n, y^n, z^n) is a member of the triple typical set $A_\epsilon^{(n)}$ for the joint distribution $p(x, y, z)$ if the following inequalities are satisfied:

$$\left| -\frac{1}{n} \log p(x^n) - H(X) \right| < \epsilon \tag{12}$$

$$\left| -\frac{1}{n} \log p(y^n) - H(Y) \right| < \epsilon \tag{13}$$

$$\left| -\frac{1}{n} \log p(z^n) - H(Z) \right| < \epsilon \tag{14}$$

$$\left| -\frac{1}{n} \log p(x^n, y^n) - H(X, Y) \right| < \epsilon \tag{15}$$

$$\left| -\frac{1}{n} \log p(x^n, z^n) - H(X, Z) \right| < \epsilon \tag{16}$$

$$\left| -\frac{1}{n} \log p(y^n, z^n) - H(Y, Z) \right| < \epsilon \tag{17}$$

$$\left| -\frac{1}{n} \log p(x^n, y^n, z^n) - H(X, Y, Z) \right| < \epsilon \tag{18}$$

The marginal distributions above are those implied by the joint distribution $p(x, y, z)$.

(b) Prove via a random coding argument that with input distribution $P_i(X)$ we can achieve the rate $I(X; Y, Z)$ where the joint distribution used to compute the mutual information is $P(X, Y, Z) = P_i(X)P_c(Y, Z|X)$. In your proof of achievability, you need only show that the average probability of error converges to zero.

10. (10 pts) *All codes are good.*
Prove Markov's inequality:
For X a positive random variable and $\delta > 0$,

$$P(X \geq \delta) \leq \frac{EX}{\delta}\,, \tag{19}$$

and use it to show that the channel coding theorem we proved in class implies the following:

For a fixed target probability of error, the probability that a randomly selected code will exceed that target goes to zero as the blocklenth goes to infinity.