ECE 534 Information Theory - MIDTERM 2

11/01/2010, LH 305.

- This exam has 4 questions, each of which is worth 25 points.
- You will be given the full 1.25 hours. **Use it wisely!** Many of the problems have short answers; try to find shortcuts. Do questions that you think you can answer correctly first.
- You may bring and use two 8.5x11" double-sided crib sheets.
- No other notes or books are permitted.
- No calculators are permitted.
- Talking, passing notes, copying (and all other forms of cheating) is forbidden.
- Make sure you explain your answers in a way that illustrates your understanding of the problem. Ideas are important, not just the calculation.
- Partial marks will be given.
- Write all answers directly on this exam.

Your name: ______________________________

Your UIN: ______________________________

Your signature: ______________________________

The exam has 4 questions, for a total of 100 points.

Question:	1	2	3	4	Total
Points:	25	25	25	25	100
Score:					

1. *Cascaded Z channels.*

Consider the Z-channel, of capacity C_Z, with input distribution $p(X = 0) = q$, $p(X = 1) = 1 - q$ given in Figure 1 below.

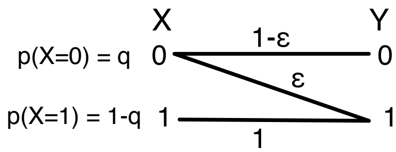


Figure 1: Z-channel with capacity C_Z.

(a) (10 points) Find an equation to be solved for the optimal (capacity achieving) value of q. Simplify this equation as much as possible, but you do not need to solve the equation.

Solution: We find q that maximizes $I(X;Y) = H(Y) - H(Y|X)$ by finding an expression for $I(X;Y)$ in terms of q and ϵ, setting its derivative with respect to q to 0 and solving for q in terms of ϵ. We have

$$\begin{aligned} H(Y) &= H(q(1-\epsilon)) \\ H(Y|X=1) &= 0 \\ H(Y|X=0) &= H(\epsilon) \\ H(Y|X) &= qH(\epsilon) \\ I(X;Y) &= H(Y) - H(Y|X) \\ &= H(q(1-\epsilon)) - qH(\epsilon) \end{aligned}$$

Setting $\frac{dI(X;Y)}{dq} = 0$ would be the equation to solve for q.

(b) (7 points) Consider the cascade of two Z channels, as shown in Figure 2 below. The output of the first channel is input directly into the second channel (no re-encoding). Find the transition probabilities of the overall channel between X and Z, called the XZ channel.

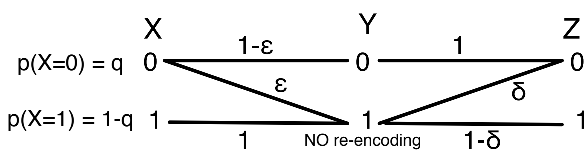


Figure 2: Two cascaded, and non-re-encoded Z-channels with capacity C_{XZ}.

Solution: The effective channel between X and Z is described by

$$p(Z=0|X=0) = 1 - \epsilon + \epsilon \cdot \delta,\ p(Z=1|X=0) = \epsilon(1-\delta),$$

$$p(Z=0|X=1) = \delta,\ p(Z=1|X=1) = 1-\delta.$$

(c) (4 points) Find the value of δ (in terms of ϵ) that makes the overall XZ channel symmetric and find the corresponding capacity (called C_{XZ}) of the symmetric XZ channel.

Points earned: ______________ out of a possible 25 points

Solution: Setting the crossover probabilities to be equal, we obtain $\delta = \frac{\epsilon}{1+\epsilon}$. The capacity is then given by $C_{XZ} = log_2(2) - H(\delta)$.

(d) (4 points) Now assume that at Y, we CAN decode and re-encode the received sequence. What is the capacity of the system now? *HINT: you can express it in terms of C_Z and C_{YZ} and C_{XZ} (or subsets thereof) - no need to go further into details than that.*

Solution: The capacity in this case is $C = \min(C_Z, C_{YZ})$.

.

Points earned: __________ out of a possible 25 points

2. *True or false (T/F) and short answer - back up your T/F statements with a reason.*

(a) (7 points) What is the capacity of the m-input discrete memoryless channel in which $Y = (X + Z) \bmod m$, where $X \in \{0, 1, 2, \cdots, m-1\}$ and

$$Z = \begin{cases} 1 & \text{with probability } \frac{3}{4} \\ 0 & \text{with probability } \frac{1}{4} \end{cases}$$

Solution: Noisy typewriter since channel transition matrix is symmetric. So, by the symmetric channel capacity theorem, $C = \log_2(m) - H(1/4)$.

(b) (6 points) Does feedback increase capacity? If so, state an example of when it may do so, if not, prove why not.

Solution: It does not increase capacity of a discrete memoryless channel (proven in class) but it may increase the capacity of a channel with memory.

(c) (6 points) Is 1110100 a codeword of the $(7, 4)$ Hamming code?

Solution: No, you can check this by multiplying it by the Hamming parity check matrix – the product is not the all 0 vector.

(d) (6 points) What is the entropy of $\mathbf{BY}$ if

$$\mathbf{B} = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}, \quad \mathbf{Y} \sim \mathcal{N}\left(\begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \right), \quad 0 \le \rho \le 1$$

Solution:

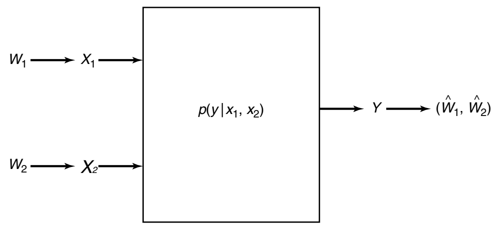
$$\begin{aligned} h(bfBY) &= h(\mathbf{Y}) + \log|\det(\mathbf{B})| \\ &= \frac{1}{2}\log(2\pi e)^2((\sigma_1^2\sigma_2^2(1-\rho^2)) + \log(6) \end{aligned}$$

.

Points earned: ____________ out of a possible 25 points

3. *Multiple-access channel.*

A *discrete memoryless multiple access channel* consists of a channel with three users (two senders, one receiver) and three alphabets, $\mathcal{X}_1, \mathcal{X}_2$ and $\mathcal{Y}$ and a probability transition matrix $p(y|x_1, x_2)$, as shown below. In this channel model, two independent messages W_1 and W_2 are encoded as $X_1^n(W_1)$ and $X_2^n(W_2)$ respectively for blocklength n at two independent encoders (they do not cooperate). From the received vector Y^n the receiver wishes to decode **both** W_1 and W_2, and forms estimates $\hat{W}_1$ and $\hat{W}_2$.



(a) (5 points) Based on our regular definitions of codes for a point-to-point channel, how would you define a code for this multiple access channel?

Solution: Pg. 525 Cover+Thomas

(b) (5 points) In point to point channels we talk about a rate R being "achievable". How would you define the achievability of a rate **pair** (R_1, R_2)?

Solution: Pg. 525 Cover + Thomas

(c) (15 points) We will show later on that the capacity region of the multiple access channel is the closure of the convex hull of all rate pairs (R_1, R_2) satisfying

$$R_1 \leq I(X_1; Y|X_2) \quad (1)$$
$$R_2 \leq I(X_2; Y|X_1) \quad (2)$$
$$R_1 + R_2 \leq I(X_1, X_2; Y) \quad (3)$$

for some product distribution $p(x_1)p(x_2)$ on $\mathcal{X}_1 \times \mathcal{X}_2$.
Suppose that our multiple access channel is such that

$$Y = X_1 + X_2 + Z, \quad Z \sim \mathcal{N}(0, N).$$

We no longer constrain X_1 and X_2 to lie in discrete alphabets, but we do put power constraints on them, i.e. $E[X_1^2] \leq P_1$ and $E[X_2^2] \leq P_2$. Find the capacity region for this Gaussian multiple access channel by finding the input distributions that maximize (1)–(3) (with explanations why) and plot this region (in the (R_1, R_2) plane).

Solution: Pg. 544 - 545 Cover + Thomas

.

Points earned: ______________ out of a possible 25 points

4. (25 points) *Time-varying channel* A train pulls out of the station at constant velocity. The received signal energy thus falls off with time as $1/i^2$. The total received signal at time i is

$$Y_i = \left(\frac{1}{i}X_i + Z_i\right)$$

where $Z_1, Z_2, \cdots$ are i.i.d. $\sim \mathcal{N}(), N)$. The transmitter constraint for block length n is

$$\frac{1}{n}\sum_{i=1}^{n} x_i^2(w) \le P, \quad w \in \{1, 2, \cdots, 2^{nR}\}.$$

Using Fano's inequality, show that the capacity C is equal to zero for this channel.

Solution:

Just as in the proof of the converse for the Gaussian channel

$$\begin{aligned}
nR = H(W) &= I(W;\hat{W}) + H(W|\hat{W}) && (9.69)\\
&\le I(W;\hat{W}) + n\epsilon_n && (9.70)\\
&\le I(X^n;Y^n) + n\epsilon_n && (9.71)\\
&= h(Y^n) - h(Y^n|X^n) + n\epsilon_n && (9.72)\\
&= h(Y^n) - h(Z^n) + n\epsilon_n && (9.73)\\
&\le \sum_{i=1}^{n} h(Y_i) - h(Z^n) + n\epsilon_n && (9.74)\\
&= \sum_{i=1}^{n} h(Y_i) - \sum_{i=1}^{n} h(Z_i) + n\epsilon_n && (9.75)\\
&= \sum_{i=1}^{n} I(X_i;Y_i) + n\epsilon_n. && (9.76)
\end{aligned}$$

Now let P_i be the average power of the ith column of the codebook, i.e.,

$$P_i = \frac{1}{2^{nR}}\sum_w x_i^2(w). \qquad (9.77)$$

Then, since $Y_i = \frac{1}{i}X_i + Z_i$ and since X_i and Z_i are independent, the average power of Y_i is $\frac{1}{i^2}P_i + N$. Hence, since entropy is maximized by the normal distribution,

$$h(Y_i) \le \frac{1}{2}\log 2\pi e(\frac{1}{i^2}P_i + N). \qquad (9.78)$$

Continuing with the inequalities of the converse, we obtain

$$\begin{aligned}
nR &\le \sum (h(Y_i) - h(Z_i)) + n\epsilon_n && (9.79)\\
&\le \sum \left(\frac{1}{2}\log(2\pi e(\frac{1}{i^2}P_i + N)) - \frac{1}{2}\log 2\pi e N\right) + n\epsilon_n && (9.80)\\
&= \sum \frac{1}{2}\log\left(1 + \frac{P_i}{i^2 N}\right) + n\epsilon_n. && (9.81)
\end{aligned}$$

Since each of the codewords satisfies the power constraint, so does their average, and hence

$$\frac{1}{n}\sum_i P_i \le P. \qquad (9.82)$$

This corresponds to a set of parallel channels with increasing noise powers. Using waterfilling, the optimal solution is to put power into the first few channels which have the lowest noise power. Since the noise power in the channel i is $N_i = i^2 N$, we will put power into channels only where $P_i + N_i \le \lambda$. The height of the water level in the water filling is less than $N + nP$, and hence the for all channels we put power, $i^2 N < nP + N$, or only $o(\sqrt{n})$ channels. The average rate is less than $\frac{1}{n}\sqrt{n}\frac{1}{2}\log(1 + nP/N)$ and the capacity per transmission goes to 0. Hence there capacity of this channel is 0.

.

Points earned: ____________ out of a possible **??** points