

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING  
EXAMINATIONS 2018

MSc and EEE/EIE PART IV: MEng and ACGI

**Corrected copy**

**INFORMATION THEORY**

Thursday, 10 May 10:00 am

Time allowed: 3:00 hours

**There are FOUR questions on this paper.**

**Answer ALL questions.**

*All questions carry equal marks*

**Any special instructions for invigilators and information for candidates are on page 1.**

Examiners responsible      First Marker(s) :      C. Ling  
   Second Marker(s) :      D. Gunduz

## Information for students

### Notation:

- (a) Random variables are shown in Tahoma font.  $x$ ,  $\mathbf{x}$ ,  $\mathbf{X}$  denote a random scalar, vector and matrix respectively.
- (b) The size of a set  $A$  is denoted by  $|A|$ .
- (c) By default, the logarithm is to the base 2.
- (d)  $\oplus$  denotes the exclusive-or operation, or modulo-2 addition.
- (e) “i.i.d.” means “independent identically distributed”.
- (f)  $H(\cdot)$  is the entropy function.
- (g)  $C(x) = \frac{1}{2} \log_2(1+x)$  is the capacity function for the Gaussian channel in bits/channel use.

## The Questions

### I. Basics of information theory.

- a) Let the joint distribution of two random variables  $X$  and  $Y$  be given by

$p(x, y)$	$y=0$	$y=1$
$x=0$	1/4	1/4
$x=1$	1/4	1/4

Compute the following quantities:

- i)  $H(X), H(Y)$
  - ii)  $H(X|Y), H(Y|X)$
  - iii)  $H(X, Y)$
  - iv)  $I(X, Y)$
  - v) Draw a Venn diagram for the above quantities. [10]
- b) Let  $X_i$  ( $i = 1, 2$ ) be i.i.d. Bernoulli ( $p = 1/4$ ) and  $Y_1 = X_2$  and  $Y_2 = X_1$ . Calculate  $I(X_i; Y_i)$  and  $I(X_{1:2}; Y_{1:2})$ . [7]
- c) Given two binary distributions  $\mathbf{p} = \{p, 1 - p\}$  and  $\mathbf{q} = \{r, 1 - r\}$ . Calculate relative entropies  $D(\mathbf{p}||\mathbf{q})$  and  $D(\mathbf{q}||\mathbf{p})$  for  $p = \frac{1}{4}$  and  $r = \frac{1}{2}$ . [8]

2. Gaussian sources and channels.

- a) Rate-distortion of Gaussian sources. Assume  $X \sim N(0, \sigma^2)$  and  $E(X - \hat{X})^2 \leq D$ . Justify each step of the following derivations.

- i) Lower bound on mutual information.

$$\begin{aligned}
 I(X; \hat{X}) &\stackrel{(1)}{=} h(X) - h(X | \hat{X}) \stackrel{(2)}{=} \frac{1}{2} \log 2\pi e \sigma^2 - h(X - \hat{X} | \hat{X}) \\
 &\stackrel{(3)}{\geq} \frac{1}{2} \log 2\pi e \sigma^2 - h(X - \hat{X}) \stackrel{(4)}{\geq} \frac{1}{2} \log 2\pi e \sigma^2 - \frac{1}{2} \log (2\pi e \text{Var}(X - \hat{X})) \\
 &\stackrel{(5)}{\geq} \frac{1}{2} \log 2\pi e \sigma^2 - \frac{1}{2} \log 2\pi e D \\
 &\stackrel{(6)}{\Rightarrow} I(X; \hat{X}) \geq \max \left( \frac{1}{2} \log \frac{\sigma^2}{D}, 0 \right)
 \end{aligned}$$

- ii) Achievability. To show that we can find a distribution  $p(\hat{x}, x)$  that achieves the lower bound, we construct a test channel that introduces distortion  $D < \sigma^2$  shown in Fig. 2.1.

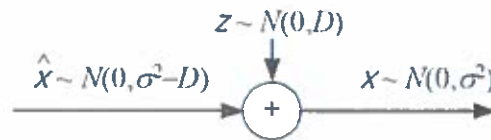


Fig. 2.1. Test channel.

$$\begin{aligned}
 I(X; \hat{X}) &= h(X) - h(X | \hat{X}) = \frac{1}{2} \log 2\pi e \sigma^2 - h(X - \hat{X} | \hat{X}) \\
 &\stackrel{(7)}{=} \frac{1}{2} \log 2\pi e \sigma^2 - h(Z | \hat{X}) \stackrel{(8)}{=} \frac{1}{2} \log \frac{\sigma^2}{D} \\
 &\stackrel{(9)}{\Rightarrow} I(X; \hat{X}) = \max \left( \frac{1}{2} \log \frac{\sigma^2}{D}, 0 \right) \\
 &\stackrel{(10)}{\Rightarrow} D(R) = \frac{\sigma^2}{2^{2R}}
 \end{aligned}$$

[10]

- b) Waterfilling. Consider three parallel Gaussian channels, i.e.,

$$\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix} = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} + \begin{pmatrix} Z_1 \\ Z_2 \\ Z_3 \end{pmatrix}$$

where

$$\begin{pmatrix} Z_1 \\ Z_2 \\ Z_3 \end{pmatrix} \sim N \left( 0, \begin{bmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{bmatrix} \right)$$

and there is a power constraint  $E(X_1^2 + X_2^2 + X_3^2) \leq 3P$ . Assume that  $\sigma_1^2 \geq \sigma_2^2 \geq \sigma_3^2$ . At what power does the channel behaves like

- i) a single channel with noise variance  $\sigma_3^2$ ? Find the channel capacity in this case.
- ii) a pair of channels with noise variances  $\sigma_3^2$  and  $\sigma_2^2$ ? Find the channel capacity in this case.
- iii) three channels with noise variances  $\sigma_3^2$ ,  $\sigma_2^2$ , and  $\sigma_1^2$ ? Find the channel capacity in this case.

[10]

- c) Denote by  $\phi(\mathbf{x})$  a multivariate Gaussian distribution with zero-mean and covariance matrix  $\mathbf{K}$ . Given any continuous probability density  $f(\mathbf{x})$  with the same mean and covariance matrix, prove that their relative entropy is given by

$$D(f||\phi) = h_\phi(\mathbf{x}) - h_f(\mathbf{x})$$

where  $h_\phi(\mathbf{x})$  and  $h_f(\mathbf{x})$  are their differential entropies, respectively.

[5]

### 3. Channel capacity.

- a) The transition matrix  $\mathbf{Q}$  of a *weakly symmetric channel* has the following properties: all columns of  $\mathbf{Q}$  have the same sum  $= |\mathbf{X}||\mathbf{Y}|^{-1}$ , and all rows are permutations of each other. Justify each step of the following derivations.

$$H(Y|X) \stackrel{(1)}{=} \sum_{x \in \mathbf{X}} p(x) H(Y|X=x) \stackrel{(2)}{=} H(\mathbf{Q}_{1,:}) \sum_{x \in \mathbf{X}} p(x) \stackrel{(3)}{=} H(\mathbf{Q}_{1,:})$$

where  $\mathbf{Q}_{1,:}$  is the entropy of the first row of the  $\mathbf{Q}$  matrix. Thus

$$\begin{aligned} I(X;Y) &\stackrel{(4)}{=} H(Y) - H(Y|X) \stackrel{(5)}{=} H(Y) - H(\mathbf{Q}_{1,:}) \stackrel{(6)}{\leq} \log |\mathbf{Y}| - H(\mathbf{Q}_{1,:}) \\ &\Rightarrow C \stackrel{(7)}{=} \log |\mathbf{Y}| - H(\mathbf{Q}_{1,:}) \end{aligned}$$

[7]

- b) In fact, the property that the uniform input distribution achieves the capacity of weakly symmetric channels can be extended to *generally symmetric channels* (to be proved in part c). In these channels, the set of outputs can be partitioned into subsets in a way, for each subset of the transition matrix: all rows are permutations of each other, and all columns are permutations of each other (similar to symmetric channels).
- Show that the binary erasure channel is a generally symmetric channel, but is not a weakly symmetric channel.
  - Which one of the following two channels is generally symmetric? Compute the capacity of the generally symmetric channel.

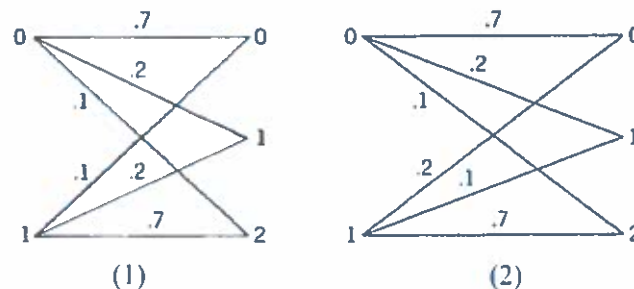


Fig. 3.1. Transition probabilities of two discrete memoryless channels (DMC).

[8]

- c) Now, let us prove the uniform input distribution achieves the capacity of *generally symmetric channels*, in two steps.

- Firstly, using the method of Lagrange multipliers, show that for any DMC, if there is an input distribution  $Q(x) > 0$  for all  $x$ , such that
 
$$I(x;Y) = C \quad \text{for all } x \quad (3.1)$$

where

$$I(x;Y) = \sum_y P(y|x) \log \frac{P(y|x)}{\sum_x Q(x)P(y|x)}.$$

then  $C$  is the channel capacity.

- Show that Equation (3.1) holds for generally symmetric channels if the input distribution is uniform.

[10]

4. Network information theory.

a) Multi-access channel.

- i) Describe the capacity region of a two-user multiple access Gaussian channel. Interpret the corner points (i.e., why can one of the users achieve the full capacity of a single-user channel as if the other user were absent?)

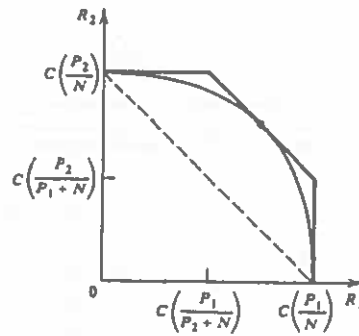


Fig. 4.1. Capacity region of multiple access channel.

- ii) Verify the following equality for the corner point:

$$C\left(\frac{P_1}{N}\right) + C\left(\frac{P_2}{P_1 + N}\right) = C\left(\frac{P_1 + P_2}{N}\right)$$

where  $C(x) = (\log(1+x))/2$  is the capacity function for the Gaussian channel.

- iii) Now consider a multi-access channel of  $m$  users, each user with the same power  $P$ . Define the degrees of freedom (DoF) as

$$d = \lim_{P \rightarrow \infty} \frac{C\left(\frac{mP}{N}\right)}{C\left(\frac{P}{N}\right)}$$

Calculate the DoF for fixed  $m$ . Discuss the DoF per user as  $m$  increases.

[15]

- b) Slepian-Wolf coding. Let  $(X, Y)$  have the joint probability mass function

$p(x,y)$	0	1
0	$\frac{1}{2}$	$\frac{1}{4}$
1	0	$\frac{1}{4}$

Calculate and sketch the Slepian-Wolf rate region for this source pair.

[10]

