Version RV/5

EGT3
ENGINEERING TRIPOS PART IIB

Monday 25 April 2016 2 to 3.30

Module 4F5

ADVANCED COMMUNICATIONS AND CODING

Answer not more than ***three*** *questions.*

All questions carry the same number of marks.

The ***approximate*** *percentage of marks allocated to each part of a question is indicated in the right margin.*

Write your candidate number ***not*** *your name on the cover sheet.*

STATIONERY REQUIREMENTS
Single-sided script paper

SPECIAL REQUIREMENTS TO BE SUPPLIED FOR THIS EXAM
CUED approved calculator allowed
Attachment: 4F5 Advanced Communications and Coding data sheet (4 pages)
Engineering Data Book

10 minutes reading time is allowed for this paper.

You may not start to read the questions printed on the subsequent pages of this question paper until instructed to do so.

1 (a) Consider a probability mass function (pmf) p on the set $\{1,2,3,4\}$ with

$$p(1) = p(2) = p(3) = p(4) = \frac{1}{4}.$$

Let X and Y be independent random variables, each drawn according to p, and let $Z = X + Y$.

(i) Compute the pmf of Z. [15%]

(ii) Compute the entropies $H(X)$ and $H(Z)$. [15%]

(iii) Compute the conditional entropy $H(Y \mid Z)$.
Hint: Expand $H(Y,Z)$ in two different ways. [20%]

(iv) Suppose $X_1,\dots,X_n$ and $Y_1,\dots,Y_n$ are each drawn independently according to p. Let $Z_i = X_i + Y_i$, for $1 \le i \le n$. For large n, briefly describe the typical set for $Z^n = (Z_1,\dots,Z_n)$. Your answer should specify what the sequences in the typical set look like, as well as the approximate size of the set. [10%]

(b) Let X,Y,Z be jointly distributed discrete random variables. Prove the following inequalities, and give the conditions for equality to hold in each case. (State the conditions for equality in terms of the joint probability mass functions.)

(i) $H(X,Y,Z) - H(X,Y) \le H(X,Z) - H(X)$. [15%]

(ii) $H(X,Z \mid Y) + H(Y) - H(Z) \le H(X \mid Z) + H(Y \mid Z)$. [25%]

2 (a) A channel takes a two-bit input $\underline{X} \in \{00, 01, 10, 11\}$ and erases exactly one bit, marking the erased bit by ε. Either of the two input bits is equally likely to be erased. Therefore, if $\underline{X} = 00$, the output $\underline{Y}$ is either 0ε or $\varepsilon 0$, with equal probability.

(i) Write down a channel transition matrix, whose entries are the probabilities $P(\underline{Y} \mid \underline{X})$, for $\underline{X} \in \{00, 01, 10, 11\}$ and $\underline{Y} \in \{0\varepsilon, \varepsilon 0, 1\varepsilon, \varepsilon 1\}$. [10%]

(ii) Find the capacity of the channel. [25%]

(iii) Describe a scheme to communicate reliably over the channel. Your scheme should have simple encoding and decoding rules. [15%]

(b) A simple linear code can be constructed as follows. Place d^2 binary source digits in a d by d grid. Add a column to the right and a row at the bottom of the grid, resulting in a $d+1$ by $d+1$ grid. Fill the last column and the last row with binary digits so that the number of ones per row and the number of ones per column are even. Then read out the codeword from the grid in any agreed order. The picture below illustrates the encoding procedure for $d = 2$, encoding 4 source digits into a codeword of length 9, read row-wise from the grid:

0	1	
1	1	

Source digits $[0, 1, 1, 1]$

$\longrightarrow$

0	1	1
1	1	0
1	0	1

Code digits $[0, 1, 1, 1, 1, 0, 1, 0, 1]$

(i) Express the rate of the code as a function of d. [10%]

(ii) For $d = 2$, there are 9 code digits, 3 grid row parity constraints and 3 grid column parity constraints. If you were to draw a factor graph of this constraint pattern, how many variable nodes and constraint nodes would you have and what would the degrees of the nodes in the graph be? [15%]

(iii) Compute the design rate implied by the factor graph in part (b)(ii). You should find a discrepancy with respect to the true rate. Explain the discrepancy. [25%]

3 A key distribution system for access to a sensitive database is based on a Reed-Solomon code. Five directors are each provided with one code digit from a Reed-Solomon codeword of length 5 over GF(11). The codeword was obtained by taking the inverse Discrete Fourier Transform (DFT) of a length 5 sequence consisting of two zeros followed by the length 3 secret key that unlocks the database. The aim of the key distribution system is to allow a subset of the five directors to agree to access the database by combining their code digits and recovering the secret key.

The Reed Solomon code is based on the following length 5 DFT matrix

$$\mathbf{F} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 4 & 5 & 9 & 3 \\ 1 & 5 & 3 & 4 & 9 \\ 1 & 9 & 4 & 3 & 5 \\ 1 & 3 & 9 & 5 & 4 \end{bmatrix}.$$

(a) What is the parity-check matrix of the Reed-Solomon code? [10%]

(b) What is the minimum distance of the Reed-Solomon code? [10%]

(c) What is the probability that a hacker who guesses the secret key uniformly at random among all its possible values will succeed in unlocking the database in the first attempt? [10%]

(d) Four directors can agree to unlock the database by assuming any value for the fifth director's digit. If their assumption is wrong, they can apply the Reed-Solomon decoder to recover the correct digit. Explain why this is true. [10%]

(e) Director 5 is captured by pirates at sea during a sailing holiday, and Directors 1 to 4 agree to unlock the database in her absence. Their code digits are 6,5,5 and 5. They initially assume Director 5's digit to be zero. Is their assumption correct? [15%]

(f) Recover the secret key using the four code digits in part (e). [30%]

(g) Given its minimum distance, how many erasures can the Reed-Solomon code recover? Hence specify how many directors are needed in theory to unlock the database. [15%]

4 Consider a QPSK constellation, with four constellation symbols given by $\{(A,A),\ (A,-A),\ (-A,-A),\ (-A,A)\}$.

(a) The constellation is used to signal over the discrete-time fading channel $Y = hX + N$, where the fading coefficient $h \sim \mathcal{CN}(0,1)$ and the noise $N \sim \mathcal{CN}(0,N_0)$ are complex Gaussian random variables. The fading coefficient h is known at the receiver.

(i) Specify the optimal detector and the decision rule to recover the transmitted symbol from Y, assuming that the constellation symbols are equally likely. [15%]

(ii) Compute a bound for the probability of detection error conditioned on h. Your answer should be in terms of the Q-function and the ratio $\frac{|h|^2 E_b}{N_0}$, where E_b is the average energy per bit of the constellation. [20%]

(iii) The probability density function of $|h|^2$ is given by $f_{|h|^2}(x) = \mathrm{e}^{-x},\ \ x \geq 0$. Using this, compute a bound for the probability of detection error averaged over all realisations of h. You may use the bound $Q(x) \leq \frac{1}{2}\mathrm{e}^{-x^2/2}$ for $x \geq 0$. [10%]

(b) To improve the performance over the fading channel, we use a system with two transmit antennas and one receive antenna. The channel is now given by

$$Y = h_a X_a + h_b X_b + N,$$

where X_a, X_b are the symbols transmitted from the two antennas. The corresponding channel fading coefficients from the two antennas are h_a, h_b, which are i.i.d. $\mathcal{CN}(0,1)$. The noise N is $\mathcal{CN}(0,N_0)$. The channel is used over two time-periods to transmit two symbols u, v drawn from the constellation above. In the first period, the two antennas transmit symbols u, v, respectively; in the second period, they transmit $v^*, -u^*$, respectively. The fading coefficients stay the same over the two time-periods and are known to the receiver. The constellation symbols are equally likely.

(i) Specify a detection procedure to recover the symbols u, v from $Y[1], Y[2]$, the channel outputs of the two time-periods. [20%]

(ii) Compute a bound for the probability of error for the detection of (u,v) averaged over all realisations of h_a, h_b. Note that $|h_a|^2$ and $|h_b|^2$ are independent, with the probability density function given in part (a)(iii) above. As before, use the bound $Q(x) \leq \frac{1}{2}\mathrm{e}^{-x^2/2}$ for $x \geq 0$. [20%]

(iii) Compare the transmission schemes in parts (a) and (b), in terms of rate, transmission power, and the average probability of detection error. [15%]

END OF PAPER

THIS PAGE IS BLANK

Module 4F5: Engineering Tripos 2015/16 – Numerical Answers

1. a) i) $p_z(2) = p_z(8) = \frac{1}{16}$, $p_z(3) = p_z(7) = \frac{2}{16}$, $p_z(4) = p_z(6) = \frac{3}{16}$, $p_z(5) = \frac{4}{16}$.

   ii) $H(Z) = 2.656$, $H(X) = 2$ bits. iii) $H(Y|Z) = 1.344$ bits

   b) i) Equality if $P_{Z|X,Y} = P_{Z|X}$, i.e., $Z - X - Y$ form a Markov chain ii) Equality if $P_{X|Z,Y} = P_{X|Z}$, i.e., $X - Z - Y$ form a Markov chain.

2. a) i)

		$\underline{Y}$			
	$P(\underline{Y}\mid\underline{X})$	0ϵ	$\epsilon 0$	1ϵ	$\epsilon 1$
$\underline{X}$	00	$\frac{1}{2}$	$\frac{1}{2}$	0	0
	10	0	$\frac{1}{2}$	$\frac{1}{2}$	0
	01	$\frac{1}{2}$	0	0	$\frac{1}{2}$
	11	0	0	$\frac{1}{2}$	$\frac{1}{2}$

   ii) 1 bit/channel use

   b) i) $\frac{d^2}{(d+1)^2}$ ii) 9 variable nodes, each of degree 2, 6 constraint nodes, each of degree 3.
   iii) Design rate =1/3, true rate = 4/9.

3. a) $\mathbf{H} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 4 & 5 & 9 & 3 \end{bmatrix}$.

   b) $d_{min} = 3$

   c) $1/11^3 = 0.0075$

   d) e)

   f) Secret key = $[9, 3, 7]$

   g) Can recover upto 2 erasures, therefore 3 directors required to unlock the database

4. a) i) ii) $P_{e|h} \leq 2\mathcal{Q}\left(\sqrt{\frac{2|h|^2 E_b}{N_0}}\right)$ iii) $\left(1 + \frac{E_b}{N_0}\right)^{-1}$.

   b) i) ii) $2\left(1 + \frac{E_b}{N_0}\right)^{-2}$ iii)