Version RV/4

EGT3
ENGINEERING TRIPOS PART IIB

Friday 24 April 2015 9.30 to 11

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) A discrete source emits independent and identically distributed (i.i.d.) binary symbols with the distribution $P(1) = 0.01,\ P(0) = 0.99$.

(i) The symbols are taken fifty at a time, and a binary codeword is provided for each sequence of fifty symbols containing two or fewer 1's. Assuming that all codewords have the same length, find the minimum length (in bits) required to provide codewords for all sequences with two or fewer 1's. [15%]

(ii) If we assign codewords as in (a)(i) above, calculate the probability of observing a source sequence for which no codeword has been assigned. [15%]

(iii) What is the minimum expected number of bits per source symbol required by *any* compression code which assigns a codeword to every source sequence? [5%]

(iv) Briefly describe how you would construct a compression code for which the expected number of bits per source symbol is close to the minimum value in (iii). The code should assign a unique codeword to every source sequence. Assume you do not have to worry about the complexity of encoding and decoding. [20%]

(b) A discrete source emits n i.i.d. symbols from the set $\{a,b,c\}$ with the distribution $P(a) = \frac{1}{2},\ P(b) = P(c) = \frac{1}{4}$.

(i) Let $(x_1,\ldots,x_n)$ be a sequence with k 'a' symbols, where $k \leq n$. What is the probability of observing this sequence? [5%]

(ii) Let $A_{\varepsilon,n}$ denote the set of ε-typical source sequences of length n. Using your answer for (b)(i) above, describe precisely what an element in the set $A_{0.05,20}$ looks like. For example, could the sequence '*aaabbbacabbaccacaacc*' be in this set?
Hint: The definition of $A_{\varepsilon,n}$ is given in the data sheet. [20%]

(iii) Use properties of typical sequences together with your answer for (b)(ii) to show that [20%]

$$\binom{20}{9}2^{11} + \binom{20}{10}2^{10} + \binom{20}{11}2^{9} \leq 2^{20(1.5+0.05)}$$

2 Consider the four-symbol constellation $\{-3A, -A, A, 3A\}$, shown in Fig. 1.

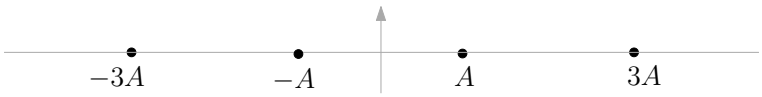

Fig. 1

(a) Suppose that this constellation is used to signal over the discrete-time Additive White Gaussian Noise (AWGN) channel

$$Y = X + N$$

where the noise N is distributed as $\mathcal{N}(0, \frac{N_0}{2})$, i.e., it is Gaussian with zero mean and variance $\frac{N_0}{2}$.

(i) Assuming that all the constellation symbols are equally likely, what are the decision regions for the optimal detector? [10%]

(ii) Compute the average probability of detection error for the decision regions obtained in (i). Your answer should be in terms of the ratio $\frac{A^2}{N_0}$ and the Q-function. [10%]

(iii) Find the decision regions for the optimal detector when the constellation symbols are chosen according the following probability distribution:
$P(X = -3A) = P(X = 3A) = 1/6$, and $P(X = -A) = P(X = A) = 1/3$. [25%]

(iv) Compute the probability of detection error for the scenario in (a)(iii) above. [20%]

(b) Now suppose that the constellation in Fig. 1 is used to signal over the fading channel $Y = hX + N$, where $h \sim \mathcal{CN}(0,1)$ and $N \sim \mathcal{CN}(0, N_0)$ are complex Gaussian random variables. Assume that the fading coefficient h is known at the receiver, and the constellation symbols are equally likely.

(i) At the receiver, we project Y in the direction of h by multiplying it by $\frac{h^*}{|h|}$, and then perform detection. (h^* is the complex conjugate of h.) What is the probability of detection error conditioned on h? [15%]

(ii) Use the approximation $Q(x) \approx \frac{1}{2}e^{-x^2/2}$ for the Q-function, and compute the probability of detection error averaged over all realisations of h. Note that the probability density function of $|h|^2$ is given by $f_{|h|^2}(x) = e^{-x}, \ x \geq 0$. [10%]

(iii) Compare the average probability of error for the fading channel with the probability of error for the AWGN channel in part (a). Explain qualitatively why one decreases much more slowly than the other as $\frac{A^2}{N_0}$ increases. [10%]

3 (a) An LDPC code is defined by the following parity-check matrix

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

A codeword of this code was transmitted over a binary erasure channel. There were five erasures denoted by ε, resulting in a received sequence

$$\mathbf{y} = [\ 0 \quad \varepsilon \quad \varepsilon \quad 1 \quad \varepsilon \quad 0 \quad \varepsilon \quad 1 \quad \varepsilon\ ].$$

Determine the transmitted codeword. [20%]

(b) A binary LDPC code has edge perspective polynomials

$$\begin{cases} \lambda(x) = 0.5x^2 + 0.3x^4 + 0.2x^5 \\ \rho(x) = x^5 \end{cases}$$

(i) What is the rate of the code? [10%]

(ii) If the code length is $N = 3900$, how many ones does its parity-check matrix have? [10%]

(c) A constraint node of degree 6 in the factor graph of a binary LDPC code receives the log-likelihood ratios $[3.2, -0.6, -7.2, 4.4, 2.8, 5.7]$ along its 6 edges. Determine the output messages it needs to compute to pass along its *third* edge in the sum-product algorithm, and in the min-sum algorithm.

Note: The sum-product and min-sum message passing equations are given in the data sheet. [25%]

(d) Let the function g_α with parameter α be defined as shown in Fig. 2.

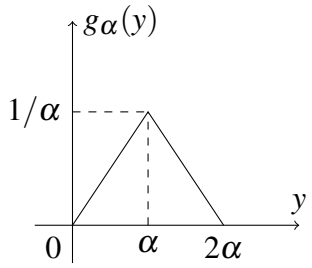

Fig. 2

(i) Consider a binary-input memoryless channel with input variable X taking values over the alphabet $\{0,1\}$, and a continuous output variable Y, with conditional densities

$$\begin{cases} f_{Y|X}(y|0) = g_\alpha(y), & \text{with } \alpha = 1 \\ f_{Y|X}(y|1) = g_\alpha(y), & \text{with } \alpha = 2 \end{cases}$$

Compute the mapping of y to the log-likelihood ratio $L(y)$ for $y \in [0,4]$.
Hint: Divide the interval $[0,4]$ into three regions and treat each region separately. [25%]

(ii) A length 5 repetition code is used at the input of the channel described in (d)(i) above. The channel observations are $[0.5, 1.2, 1.3, 2.2, 0.1]$. What is the transmitted codeword? [10%]

4 (a) Consider the cascade channel shown in Fig. 3, in which the output of a binary symmetric channel with crossover probability $p < \frac{1}{2}$ is the input to an erasure channel with erasure probability α.

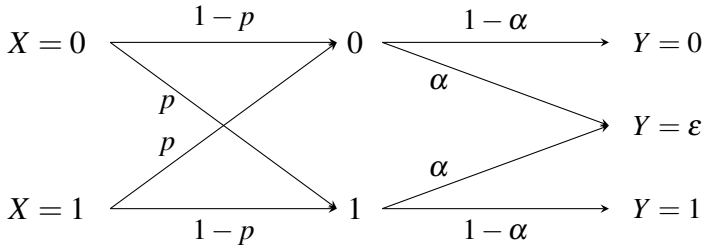

Fig. 3

(i) Write down the channel transition matrix for the cascade channel, whose entries are the probabilities $P(Y|X)$, for $X \in \{0,1\}$, $Y \in \{0,\varepsilon,1\}$. [15%]

(ii) Find the capacity of the cascade channel. Your answer should be in terms of p and α. [25%]

(b) Consider arithmetic over the Galois Field GF(5).

(i) What is the multiplicative order of the elements 2 and 3 in GF(5)? Why can't there be an element of multiplicative order 3 in GF(5)? [10%]

(ii) Specify a parity-check matrix $\mathbf{H}$ of a Reed-Solomon code $\mathcal{C}$ of block length $N = 4$ over GF(5) that can correct any error pattern of up to one error. [10%]

(iii) What is the dimension of the code $\mathcal{C}$? How many codewords does it contain? [10%]

(iv) Is the parity-check matrix $\mathbf{H}$ the only parity-check matrix of $\mathcal{C}$? If not, give an example of an alternative parity-check matrix for $\mathcal{C}$. [10%]

(v) Is $[1,2,4,3]$ a codeword of $\mathcal{C}$? Justify your answer. [5%]

(vi) Is $\mathcal{C}$ the only Reed-Solomon code over GF(5) with block length $N = 4$ that can correct any pattern of up to one error? Justify your answer. For example, if you answer no, specify a parity-check matrix $\mathbf{H}'$ of another Reed-Solomon code $\mathcal{C}'$ over GF(5) that can correct any pattern of up to one error. [15%]

END OF PAPER

Module 4F5: Engineering Tripos 2014/15 – Numerical Answers

1. a) i) $\lceil \log_2 1276 \rceil = 11$ bits; ii) 0.0138; iii) $H(X) = 0.0808$ bits/symbol

   b) i) $\left(\frac{1}{2}\right)^k \left(\frac{1}{4}\right)^{n-k}$; ii) $A_{0.05,20}$ consists of length 20 sequences with 9,10, or 11 'a' symbols.

2. a) i) Decision boundaries at $-2A, 0, 2A$ ii) $\frac{3}{2}\mathcal{Q}\left(\sqrt{\frac{2A^2}{N_0}}\right)$ iii) Decision boundaries at $-2A-\frac{N_0 \ln 2}{4A}, 0, 2A+\frac{N_0 \ln 2}{4A}$ iv) $\frac{1}{3}\mathcal{Q}\left(\sqrt{\frac{2(A-\frac{N_0 \ln 2}{4A})^2}{N_0}}\right)+\frac{2}{3}\left[\mathcal{Q}\left(\sqrt{\frac{2(A+\frac{N_0 \ln 2}{4A})^2}{N_0}}\right)+\mathcal{Q}\left(\sqrt{\frac{2A^2}{N_0}}\right)\right]$

   b) i) $\frac{3}{2}\mathcal{Q}\left(\sqrt{\frac{2|h|A)^2}{N_0}}\right)$ ii) $\frac{3}{4}\left(1+\frac{A^2}{N_0}\right)^{-1}$

3. a) [0 1 0 1 0 0 1 1 0]

   b) i) 14/39; ii) 15,000 ones

   c) Sum-product message= -0.47, Min-sum, message= -0.6

   d) $L(y) = \log 4$ for $y \in [0,1]$, $\log 4 + \log(\frac{2}{y} - 1)$ for $y \in [1,2]$, $-\infty$ for $y \in [2,4]$.

4. a) i)

		Y		
	$P_{Y\|X}$	0	ϵ	1
X	0	$(1-p)(1-\alpha)$	α	$p(1-\alpha)$
	1	$p(1-\alpha)$	α	$(1-p)(1-\alpha)$

   ii) $(1-\alpha)(1-H_2(p))$ bits/channel use

   b) i) Multiplicative order of both 2 and 3 = 4;

   ii) Either

$$(*) : \mathbf{H} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 3 \end{bmatrix} \text{ or } (**) : \mathbf{H} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 3 & 4 & 2 \end{bmatrix}$$

   iii) Dimension=2, 25 codewords; v) Codeword if $\mathbf{H}$ is given by (*) , not a codeword if (**); vi) Choose the other parity matrix in ii).

Advanced Communications and Coding Data Sheet

1 Information Theory

- The *Joint Entropy* of random variables $X_1, \ldots, X_n$ with joint pmf $P_{X_1 \ldots X_n}$ is

$$H(X_1, X_2, \ldots, X_n) = \sum_{x_1, \ldots, x_n} P_{X_1 \ldots X_n}(x_1, \ldots, x_n) \log \frac{1}{P_{X_1 \ldots X_n}(x_1, \ldots, x_n)}$$

- The *Conditional Entropy* of Y given X is

$$H(Y|X) = \sum_x P_X(x) \underbrace{\sum_y P_{Y|X}(y|x) \log \frac{1}{P_{Y|X}(y|x)}}_{H(Y|X=x)} = \sum_x P_X(x) H(Y|X=x)$$

  Note that a similar formula holds if we condition on a collection of rvs $(X_1, \ldots, X_n)$ instead of a single random variable X.

- *Chain rule for entropy*: The joint entropy of $X_1, \ldots, X_n$ can be written as

$$\begin{aligned} H(X_1, X_2 \ldots, X_n) &= H(X_1) + H(X_2|X_1) + H(X_3|X_2, X_1) + \ldots + H(X_n|X_{n-1}, \ldots, X_1) \\ &= \sum_{i=1}^{n} H(X_i|X_{i-1}, \ldots, X_1), \quad \text{where} \end{aligned}$$

  $H(X_i|X_{i-1}, \ldots, X_1) = -\sum_{x_1, \ldots, x_i} P_{X_1, \ldots, X_i}(x_1, \ldots, x_i) \log P_{X_i|X_1, \ldots, X_{i-1}}(x_i|x_1, \ldots, x_{i-1})$

- The *typical set* $A_{\epsilon,n}$ with respect to P is the set of sequences $(x_1, \ldots, x_n) \in \mathcal{X}^n$ which satisfy the property

$$2^{-n(H(X)+\epsilon)} \leq P(x_1, \ldots, x_n) \leq 2^{-n(H(X)-\epsilon)},$$

  where $P(x_1, \ldots, x_n) = \prod_{i=1}^{n} P(x_i)$.

- The *relative entropy* or the K-L distance between two pmfs P and Q (defined on the same alphabet) is

$$D(P||Q) = \sum_{x \in \mathcal{X}} P(x) \log \frac{P(x)}{Q(x).}$$

- The *mutual information* between random variables X and Y with joint pmf P_{XY} is

$$\begin{aligned} I(X;Y) &= H(X) - H(X|Y) = H(Y) - H(Y|X) \\ &= H(X) + H(Y) - H(X,Y) = D(P_{XY}||P_X P_Y) \end{aligned}$$

- *Chain rule for Mutual Information*:

$$\begin{aligned} I(X_1, X_2, \ldots, X_n; Y) &= I(X_1; Y) + I(X_2; Y|X_1) + \ldots + I(X_n; Y|X_{n-1}, \ldots, X_1) \\ &= \sum_{i=1}^{n} I(X_i; Y|X_{i-1}, X_{i-2}, \ldots, X_1) \end{aligned}$$

- Random variables X, Y, Z are said to form a *Markov chain* (denoted $X - Y - Z$) if their joint pmf can be written as $P_{XYZ} = P_X P_{Y|X} P_{Z|Y}$.
- *Data-Processing inequality*: If X, Y, Z form a Markov chain, then $I(X;Y) \geq I(X;Z)$
- *Fano's inequality*: Let $\hat{X}$ be *any* estimator of X from Y, i.e., $X - Y - \hat{X}$. Then, the probability of error $P_e = \Pr(\hat{X} \neq X)$ satisfies
$$1 + P_e \log |\mathcal{X}| \geq H(X|Y)$$
- The *differential entropy* of a continuous random variable X with pdf f_X is
$$h(X) = \int_{-\infty}^{\infty} f_X(u) \log \frac{1}{f_X(u)} du$$
Joint differential entropy, conditional differential entropy, mutual information, relative entropy are all defined similarly to the discrete case.
- The differential entropy of a Gaussian random variable $X \sim \mathcal{N}(0, \sigma^2)$ (mean zero and variance σ^2) is
$$h(X) = \frac{1}{2} \log(2\pi e \sigma^2) \text{ bits}$$

2 Modulation & Wireless Communication

- *White Gaussian Noise* $N(t)$: A random process in which $N(t)$ is a Gaussian rv for each t, and
$$\mathbb{E}[N(t)] = 0, \quad \mathbb{E}[N(t)N(s)] = \frac{N_0}{2}\delta(t-s), \quad \text{for all } t, s$$
- The $\mathcal{Q}$-*function* characterises the tail probability of a standard Gaussian distribution:
$$\mathcal{Q}(y) = \int_y^{\infty} \tfrac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$
- *Complex Gaussians*: $h \sim \mathcal{CN}(0, \sigma^2)$ means that h is a complex random variable whose real and imaginary parts are i.i.d Gaussians, each distributed as $\mathcal{N}(0, \frac{\sigma^2}{2})$
- If $h \sim \mathcal{CN}(0, \sigma^2)$, then the squared-magnitude $|h|^2$ is exponentially distributed, i.e., its pdf is
$$f_{|h|^2}(x) = \frac{1}{\sigma^2} \exp\left(\frac{-x}{\sigma^2}\right), \quad x \geq 0$$
- The *Delay Spread* T_d of a multipath fading channel is the maximum difference between delays of the paths from transmitter to receiver. The number of channel taps is $\lceil 2WT_d \rceil$, where W is the one-sided baseband bandwidth of the transmitted signal.
- If $T_d \ll \frac{1}{2W}$, the channel is said to have *flat fading* (no inter-symbol interference). If $T_d > \frac{1}{2W}$, the fading channel has multiple taps and is said to be *frequency selective*.
- The *coherence bandwidth* of the channel is $1/(2T_d)$. If the one-sided baseband bandwidth of the transmitted signal is less than the coherence bandwidth, there will be only one channel tap, i.e., flat fading.

3 Coding

Finite Fields, Linear Codes

Finite fields:

- A Galois Field GF(q) for $q = p^m$ where p is any prime number consists of a multiplicative group of order $q-1$ and an additive group of order q
- The order of an element α in a group is the smallest power n such that $\alpha^n = 1$, where 1 is the neutral element of the group
- *Lagrange Theorem:* the order of a subgroup (and thus the order of any element in a group) divides the order of the group

Linear codes:

- A code can correct t or less errors if and only if $2t < d_{\min}$ where $d_{\min}$ is the minimum Hamming distance between any two codewords
- *Singleton Bound:* for an (N, K) q-ary code, $d_{\min} \leq N - K + 1$ with equality for Maximum Distance Separable (MDS) codes
- A K-dimensional linear code of codeword length N (an (N, K) linear code) has $K \times N$ encoder matrices, $(N-K) \times N$ parity-check matrices, and a rate $R = K/N$
- An (N, K) linear code has a systematic encoder matrix of the form $\mathbf{G} = [\mathbf{I}_K, \mathbf{P}]$, to which corresponds a parity-check matrix of the form $\mathbf{H} = [-\mathbf{P}^T, \mathbf{I}_{N-K}]$
- The minimum distance $d_{\min}$ of a linear code is the minimum Hamming weight $w_{\min}$ of any non-zero codeword

Reed Solomon Codes

- The Discrete Fourier Transform (DFT) of a vector $\mathbf{x} = [x_0, \ldots, x_{N-1}]$ with elements over a finite field $\mathcal{F}$ is defined by $X_k = \sum_{i=0}^{N-1} x_i \alpha^{ik}$ where α must be an element of multiplicative order N in $\mathcal{F}$
- The inverse DFT is $x_i = \frac{1}{N^\star} \sum_{i=0}^{N-1} X_k \alpha^{-ik}$ where $N^\star$ is N if $\mathcal{F} = \mathrm{GF}(p)$ for p prime, and $N^\star = N \pmod p$ if $\mathcal{F} = \mathrm{GF}(p^m)$ for $m > 1$
- *Blahut's theorem:* the linear complexity of the DFT of a sequence of length N equals the Hamming weight of the sequence, provided the Hamming weight is less than $N/2$
- *Reed Solomon code:* An (N, K) linear code over GF(q) with a parity-check matrix $\mathbf{H} = [\alpha^{ij}]$ for $i = 0, \ldots, (N-K-1)$ and $j = 0 \ldots, N-1$, where α is an element of multiplicative order N is GF(q)
- A Reed Solomon code has rate $R = K/N$, has minimum distance $d_{\min} = N - K + 1$ and is hence MDS

LDPC Codes

- Log likelihood ratios
$$L(y_k) = \log \frac{P_{Y|X}(y_k|0)}{P_{Y|X}(y_k|1)} = \log \frac{P_{X|Y}(0|y_k)}{P_{X|Y}(1|y_k)}$$
where the second equality holds if and only if the channel has equi-probable inputs
- Log-likelihood ratio for a binary input Additive White Gaussian channel with inputs $\{-1, +1\}$ and noise variance σ^2
$$L(y) = \frac{2}{\sigma^2} y$$
- Extrinsic decoding rules for nodes in the factor graph for a binary code (sum product algorithm)
$$L_i^{ex} = \sum_{j=\backslash i} L(y_j)$$
for variable nodes, and
$$L_i^{ex} = 2 \tanh^{-1} \left(\prod_{j=\backslash i} \tanh \frac{L(y_j)}{2} \right)$$
for constraint nodes.
- Min-sum simplified decoding rule for a constraint node
$$\text{sign}(L_i^{ex}) = \prod_{j \neq i} \text{sign}(L(y_j)) \text{ and } |L_i^{ex}| = \min_{j \neq i} |L(y_j)|$$
- Degree polynomials from a node perspective:
$$L(x) = \sum_{i=1}^{d_v^{\max}} L_i x^i, R(x) = \sum_{i=1}^{d_c^{\max}} R_i x^i$$
- Degree polynomials from an edge perspective:
$$\lambda(x) = \sum_{i=1}^{d_v^{\max}} \lambda_i x^{i-1}, \rho(x) = \sum_{i=1}^{d_c^{\max}} \rho_i x^{i-1},$$
- Average degrees $\bar{d}^{\ell} = \left(\int_0^1 \lambda(x) dx \right)^{-1}$ and $\bar{d}^{r} = \left(\int_0^1 \rho(x) dx \right)^{-1}$
- "Design" rate of an LDPC code
$$R = 1 - \frac{L'(1)}{R'(1)} = 1 - \frac{\int_0^1 \rho(x) dx}{\int_0^1 \lambda(x) dx}$$
- Density evolution for binary erasure channels with erasure probability δ:
$$\begin{cases} P_{ve} = \delta \lambda(P_{vi}) \\ P_{ce} = 1 - \rho(1 - P_{ci}) \end{cases}$$
or in one step, $P_{ve,k+1} = \delta \lambda \left(1 - \rho(1 - P_{ve,k})\right)$