EGT3
ENGINEERING TRIPOS PART IIB

Tuesday 2 May 2017 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: 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 binary source emits independent and identically distributed (i.i.d.) binary symbols $X_1, X_2, \ldots$ with distribution $P(1) = p$ and $P(0) = 1 - p$.

(a) For $p = 0.11$, consider a sequence of length 100 that contains 11 ones and 89 zeros. For what range of ε is this sequence in the typical set $A_{\varepsilon,n}$ for $n = 100$ with respect to P? [10%]

(b) Again for $p = 0.11$, consider a sequence that contains 10 ones and 90 zeros. For what range of ε is this sequence in the typical set $A_{\varepsilon,n}$ for $n = 100$ with respect to P? [20%]

(c) Starting from the definition of a typical sequence, prove the upper bound

$$|A_{\varepsilon,n}| \leq 2^{n(H_2(p)+\varepsilon)}$$

for the number of sequences in the typical set $A_{\varepsilon,n}$, where $H_2(p)$ is the binary entropy function evaluated at p. [10%]

(d) If we wish to assign a distinct codeword of fixed length L_T to every typical sequence in $A_{\varepsilon,n}$ for $\varepsilon = 10^{-2}$ and $n = 100$, use the bound proved in part (c) to determine a numerical upper bound on L_T for $p = 0.11$. [20%]

(e) Describe a data compression system that can compress and successfully decompress all source sequences and whose expected codeword length is close to the length L_T computed above. [20%]

(f) Use the upper bound proved in part (c) to show the following upper bound on a sum of binomial coefficients

$$\sum_{k=a}^{b} \binom{n}{k} \leq 2^{n(H_2(\alpha)+\beta)}, \text{ where } \alpha = \frac{a+b}{2n} \text{ and } \beta = \frac{a-b}{2n} \log \frac{1-\alpha}{\alpha}$$

for any integers a and b such that $0 \leq a \leq b \leq n$. [20%]

2 Two prisoners barred from communicating are allowed to play chess, where the guard carries a chessboard from one cell to another between moves. The prisoners pass secret information encoded in the movements of their knights. The chess moves allowed for a knight are shown in Fig. 1.

(a) Neglecting other chess pieces and assuming that the knight is nowhere near the edge of the board, what is the maximum number of bits that a prisoner can transmit by moving a knight? [10%]

(b) The guard becomes suspicious and decides to shift the knight by exactly one square either north, south, east or west after each move with a probability of $1/4$ for each direction. The guard does so in full view of the prisoner who made the move, before carrying the chessboard to the other prisoner. A noisy channel is induced, where the input random variable X corresponds to the move performed by one prisoner, and the output random variable Y is the knight's position observed by the other prisoner, assuming he knows the knight's original position. What is the capacity of this channel? [40%]

(c) The prisoners use a Reed-Solomon code to communicate over the noisy chess knight channel, where the code alphabet is the 8-ary channel input alphabet. For what block length N does a Reed-Solomon code exist over this alphabet? [10%]

(d) Assuming that any output position Y for which the input move X is ambiguous is dismissed as an erasure, let p be the probability of erasure. Choose an (N,K) Reed-Solomon code such that $0 < K < N(1-p)$, state its rate and compute the probability of decoding failure. [40%]

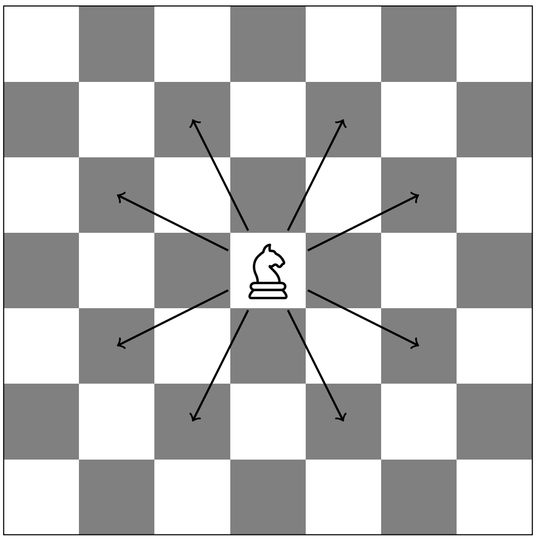

Fig. 1

3 Consider a design for binary low-density parity-check (LDPC) codes with edge perspective degree polynomials

$$\begin{cases} \lambda(x) = \frac{1}{12}x + \frac{3}{4}x^2 + \frac{1}{6}x^3 \\ \rho(x) = x^8. \end{cases}$$

(a) Determine the design rate corresponding to the degree polynomials. [20%]

(b) The threshold computed by density evolution for this code design is for an erasure probability of $p = 0.287$. What is the relative rate loss of this code design with respect to the channel capacity? [10%]

(c) Suppose we construct a parity-check matrix with $N = 900$ columns, with proportions of non-zero entries corresponding to the degree polynomials, and with linearly independent rows.

(i) How many rows does the matrix have? [10%]

(ii) How many variable nodes in its associated factor graph have degree 3? [10%]

(iii) How many codewords does the corresponding code have? [10%]

(iv) A codeword is transmitted and received with 250 erased symbols. Is the iterative decoder guaranteed to recover the transmitted codeword? Justify your answer. [20%]

(d) Suppose we design a family of parity-check matrices with linearly independent rows, with proportions of non-zero entries corresponding to the degree polynomials given, and with numbers of columns $N_1, N_2, N_3, \ldots$ where $N_1 < N_2 < N_3 < \ldots$ Let $d_1, d_2, d_3, \ldots$ be the minimum distances of the corresponding codes in the family.

(i) By considering the proportion of erasures that can be decoded by iterative decoding as the code length increases, state a lower bound on the limit

$$\lim_{k \to \infty} \frac{d_k}{N_k}.$$

[10%]

(ii) For a Binary Symmetric Channel (BSC), the proportion of errors that are guaranteed to be correctable corresponds to roughly half the minimum distance of a code. It would be tempting to argue that the iterative decoding threshold for the sum-product algorithm on the BSC is at least half of the lower bound obtained in part d(i). Why is this argument incorrect? [10%]

4 Consider the irregular four-symbol constellation $\{-\kappa A, -A, A, \kappa A\}$ with $\kappa > 1$ shown in Fig. 2.

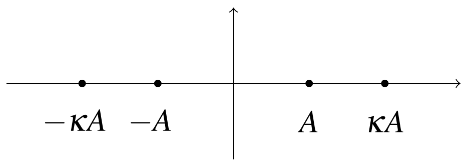

Fig. 2

(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 $\mathcal{N}(0, N_0/2)$, i.e. Gaussian with zero mean and variance $\frac{N_0}{2}$.

(i) Find the average signal energy E_s of the constellation as a function of κ and A. [10%]

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

(iii) Compute the average probability of detection error as a function of E_s, κ and N_0, and state the expression when $\kappa = 2$. [30%]

(b) Now suppose that the constellation in Fig. 2 with $\kappa = 2$ 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 coherent detection is performed, i.e. the fading coefficient h is known at the receiver.

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

(ii) Use the approximation $\mathcal{Q}(x) \approx \frac{1}{2}e^{-x^2/2}$ for the $\mathcal{Q}$-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 $f_{|h|^2}(x) = e^{-x}, x \geq 0$. [15%]

(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 this probability of error comparison is misleading and give an example of how reliable transmission can be achieved on the fading channel. [20%]

END OF PAPER

THIS PAGE IS BLANK