Dr R Venkataramanan

4F5 Advanced Communications and Coding Engineering Tripos 2014/15 – Solutions

Question 1

(a) i) The number of length fifty sequences with two or fewer ones is

$$\binom{50}{0} + \binom{50}{1} + \binom{50}{2} = 1276.$$

The number of bits needed to provide codewords for 1276 sequences = $\lceil \log_2 1276 \rceil = 11$ bits. [15%]

ii) The required probability is $1 - \sum_{k=0}^{2} \binom{50}{k} (0.01)^k (0.99)^{50-k} = 0.0138$ [15%]

iii) The minimum expected number of bits per source symbol is the entropy of the source, which is $H_2(0.99) = 0.0808$ bits. [5%]

iv) Fix $\epsilon > 0$ be very small. Consider the set of ϵ-typical length n sequences, given by

$$A_{\epsilon,n} = \left\{ x^n \in \{0,1\}^n : \; 2^{-n(H(X)+\epsilon)} \leq P(x^n) \leq 2^{-n(H(X)-\epsilon)} \right\}. \quad (1)$$

Here $H(X)$ is the source entropy, equal to 0.0808 bits. For all sequences $x^n \in A_\epsilon^n$, assign a codeword of length $\lceil n(H(X) + \epsilon) \rceil$, where $H(X) = 0.0808$. Append a flag bit, say 0, to each of these codewords. For the remaining codewords, assign codewords of length n (by taking the entire source sequence to be the codeword). Append a flag bit of 1 to each of these.

For any fixed ϵ, the probability of observing a sequence from A_ϵ^n goes to 1 as $n \to \infty$. This implies that the average number of bits per source symbol is $H(X) + \epsilon'$, where ϵ' is very small. (See lecture notes for the detailed calculation, which is not required here.) [20%]

(b) i) The probability of a *given* sequence $(x_1, x_2, \ldots, x_n)$ with k "*a*" symbols, j "*b*" symbols and $n - k - j$ "*c*" symbols (for $0 \leq j \leq n - k$) is

$$(1/2)^k \cdot (1/4)^j \cdot (1/4)^{n-k-j} \;=\; (1/2)^k \cdot (1/4)^{n-k},$$

i.e., it does not depend on j. [5%]

ii) Denoting the source random variable as X, we find that $H(X) = 1.5$ bits. Using the expression for $P(x^n)$ from (i) in the definition of $A_{\epsilon,n}$ in (1), we see that

$$A_{\epsilon,n} = \left\{ x^n \in \{0,1\}^n : \; 2^{-n(1.5+\epsilon)} \leq \; (1/2)^{k(x^n)} (1/4)^{n-k(x^n)} \leq 2^{-n(1.5-\epsilon)} \right\},$$

where $k(x^n)$ denotes the number of "a" symbols in the sequence x^n. Plugging in $n = 20$ and $\epsilon = 0.05$, we have [20%]

$$\begin{aligned} A_{\epsilon,n} = \left\{ x^n \in \{0,1\}^n : \; 2^{-31} \leq \; 2^{-(40-k(x^n))} \leq 2^{-29} \right\} &= \{ x^n \in \{0,1\}^n : \; -31 \leq \; -40 + k(x^n) \leq -29 \} \\ &= \{ x^n \in \{0,1\}^n : \; 9 \leq \; k(x^n) \leq 11 \}. \end{aligned} \quad (2)$$

Therefore $A_{0.05,20}$ *consists of length* 20 *sequences with* 9, 10, *or* 11 *a symbols.*

iii) The number of length 20 sequences with k a's is $\binom{20}{k}2^{20-k}$. (Choose k positions for the $a's$; each of the remaining $(20-k)$ positions can have either b or c.) Therefore the size of the set $A_{0.05,20}$ is

$$|A_{0.05,20}| = \binom{20}{9}2^{11} + \binom{20}{10}2^{10} + \binom{20}{11}2^{9}.$$

From the lecture notes, we know that $|A_{\epsilon,n}| \leq 2^{n(H(X)+\epsilon)}$. Hence, using $\epsilon = 0.05$ we obtain [20%]

$$|A_{0.05,20}| = \binom{20}{9}2^{11} + \binom{20}{10}2^{10} + \binom{20}{11}2^{9} \leq 2^{20(1.5+0.05)}.$$

Question 2

(a) i) As all the symbols are equally likely, the optimal MAP detector reduces to the minimum distance rule: The decision regions are [10%]

$$\hat{X} = \begin{cases} -3A & \text{if } Y \leq -2A \\ -A & \text{if } -2A < Y \leq 0 \\ A & \text{if } 0 < Y \leq 2A \\ 3A & \text{if } Y > 2A \end{cases}$$

ii) The probability of error is

$$\begin{aligned} P(\hat{X} \neq X) &= P(X = -3A)P(\hat{X} \neq -3A|X = -3A) + P(X = A)P(\hat{X} \neq A|X = A) \\ &+ P(X = A)P(\hat{X} \neq A|X = A) + P(X = 3A)P(\hat{X} \neq 3A|X = 3A). \end{aligned} \quad (3)$$

For $X = -3A$, we have

$$\begin{aligned} &P(\hat{X} \neq -3A \mid X = -3A) = P(Y > -2A \mid X = -3A) \\ &= P(\{-3A + N > -2A\} \mid X = -3A) = P(\{N > A\}) = \mathcal{Q}\left(\sqrt{\tfrac{2A^2}{N_0}}\right). \end{aligned} \quad (4)$$

For $X = -A$, we have [10%]

$$\begin{aligned} &P(\hat{X} \neq -A \mid X = -A) = P(\{Y < -2A\} \cup \{Y > 0\} \mid X = -A) \\ &= P(\{N < -A\} \cup \{N > A\}) = 2\mathcal{Q}\left(\sqrt{\tfrac{2A^2}{N_0}}\right). \end{aligned} \quad (5)$$

By symmetry, we have $P(\hat{X} \neq -3A|X = -3A) = P(\hat{X} \neq 3A|X = 3A)$, and $P(\hat{X} \neq -A|X = -A) = P(\hat{X} \neq A|X = A)$. Substituting in (3) and noting that all the symbols are equally likely, we obtain $P(\hat{X} \neq X) = \frac{3}{2}\mathcal{Q}\left(\sqrt{\frac{2A^2}{N_0}}\right)$.

iii) Now the optimal MAP detector is no longer the minimum-distance rule. For each y, we have to determine:

$$\hat{X}(y) = \underset{X \in \{-3A,-A,A,3A\}}{\arg\max} \; P(X)f(y|X)$$

Since symbol $-A$ has higher probability than $-3A$, we expect the decision boundary between them to be to the left of $-2A$. Let us verify this. The optimal detection rule to decide between $-3A$ and $-A$ is

$$\hat{X} = \underset{x \in \{-3A,-A\}}{\arg\max} \; P(Y = y|X = x)\; P(X = x) = \underset{x \in \{-3A,-A\}}{\arg\max} \; e^{-(y-x)^2/N_0} P(X = x) \quad (6)$$

For $x = -3A$, the test statistic in (6) is $\frac{1}{6}e^{-(y+3A)^2/N_0}$. For $x = -A$, the test statistic is $\frac{1}{3}e^{-(y+A)^2/N_0}$. Therefore, $\hat{X} = -3A$ when

$$\begin{aligned} \frac{1}{6}e^{-(y+3A)^2/N_0} \geq \frac{1}{3}e^{-(y+A)^2/N_0} &\Leftrightarrow -\ln 6 - (y + 3A)^2/N_0 \geq -\ln 3 - (y + A)^2/N_0 \\ &\Leftrightarrow y \leq -2A - \frac{N_0 \ln 2}{4A}. \end{aligned} \quad (7)$$

By symmetry, the decision boundary between A and $3A$ is now $2A + \frac{N_0 \ln 2}{4A}$. Since $-A$ and A have equal probability, their decision boundary will remain at zero. Therefore: [25%]

$$\hat{X} = \begin{cases} -3A & \text{if } Y \leq -2A - \frac{N_0 \ln 2}{4A} \\ -A & \text{if } -2A - \frac{N_0 \ln 2}{4A} < Y \leq 0 \\ A & \text{if } 0 < Y \leq 2A + \frac{N_0 \ln 2}{4A} \\ 3A & \text{if } Y > 2A + \frac{N_0 \ln 2}{4A} \end{cases}$$

Assessor's comment: Most students had the right idea that the decision boundaries would shift closer to the symbols with lower probability, but did not apply the MAP rule correctly. Some guessed (incorrectly) that the shift would be proportional to the symbol probabilities.

iv) To find $P(\hat{X} \neq X)$ using the general formula (3), we calculate $P(\hat{X} \neq -3A \mid X = -3A)$ and $P(\hat{X} \neq -A \mid X = -A)$ as follows.

$$P(\hat{X} \neq -3A \mid X = -3A) = P(N > A - \frac{N_0 \ln 2}{4A}) = \mathcal{Q}\left(\sqrt{\frac{2(A - \frac{N_0 \ln 2}{4A})^2}{N_0}}\right), \tag{8}$$

$$P(\hat{X} \neq -A \mid X = -A) = P(N < A + \frac{N_0 \ln 2}{4A}) + P(N > A) = \mathcal{Q}\left(\sqrt{\frac{2(A + \frac{N_0 \ln 2}{4A})^2}{N_0}}\right) + \mathcal{Q}\left(\sqrt{\frac{2A^2}{N_0}}\right). \tag{9}$$

Since $P(\hat{X} \neq -3A|X = -3A) = P(\hat{X} \neq 3A|X = 3A)$ and $P(\hat{X} \neq -A|X = -A) = P(\hat{X} \neq A|X = A)$, substituting the above in (3) gives [20%]

$$P(\hat{X} \neq X) = \frac{1}{3}\mathcal{Q}\left(\sqrt{\frac{2(A - \frac{N_0 \ln 2}{4A})^2}{N_0}}\right) + \frac{2}{3}\left[\mathcal{Q}\left(\sqrt{\frac{2(A + \frac{N_0 \ln 2}{4A})^2}{N_0}}\right) + \mathcal{Q}\left(\sqrt{\frac{2A^2}{N_0}}\right)\right].$$

(b) i) Multiplying the output by $h^*/|h|$, we obtain

$$\bar{Y} = |h|X + \bar{N} \tag{10}$$

where $\bar{N} \sim \mathcal{CN}(0, N_0)$. the effective signal is now $|h|X$, which can one of four (real) values $\{-|h|3A, -|h|A, |h|A, 3|h|A\}$. Since the effective signal is real-valued, we need only the real part of $\bar{Y}$ for detection. Therefore, from (10), we have [15%]

$$\Re(\bar{Y}) = |h|X + \Re(\bar{N}) \tag{11}$$

where $\Re(\bar{N}) \sim \mathcal{N}(0, \frac{N_0}{2})$. This is identical to the detection problem in part (a).(i) and (ii), except that the symbols values are now multiplied by $|h|$. Therefore the error probability conditioned on h is

$$P_{e|h} = \frac{3}{2}\mathcal{Q}\left(\sqrt{\frac{2(|h|A)^2}{N_0}}\right).$$

ii) Using the Q-approximation, $P_{e|h} \approx \frac{3}{4}e^{\frac{-(|h|A)^2}{N_0}}$. The probability of error averaged over all realisations of h is [10%]

$$P_e = \int_0^\infty \frac{3}{4}e^{\frac{-xA^2}{N_0}x}e^{-x}\,dx = \frac{3}{4}\left(\frac{1}{1 + \frac{A^2}{N_0}}\right). \tag{12}$$

iii) From part (a).(ii), the probability of error for the AWGN channel is $\frac{3}{2}\mathcal{Q}\left(\sqrt{\frac{2A^2}{N_0}}\right) \approx \frac{3}{4}e^{-A^2/N_0}$, which decreases *exponentially* with A^2/N_0. The probability of error for the fading channel is $\frac{1}{1+\frac{A^2}{N_0}}$, i.e., decays polynomially with A^2/N_0. The significantly slower decay of error for the fading channel is due to the probability of the channel being in *deep fade*, i.e., the channel strength $|h|^2 \ll (A^2/N_0)^{-1}$. For such an h, (12) implies that $P_{e|h}$ is close to $3/4$ — very high! [10%]

Question 3

(a) The first parity-check equation has only one erasure, allowing us to resolve the 7th symbol to a 1. The sixth parity-check equation also contains a single erasure, allowing us to resolve the third symbol to a zero. The third parity-check equation, which initially contained two erasures, now only contains one after resolving the third symbol, allowing us to resolve the 9th symbol to a zero. The fifth and second parity- check equation now allow us to resolve the 2nd and 5th symbol in turn to 1 and 0, respectively, yielding the solution [20%]

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

(b) i) The rate is [10%]

$$\begin{aligned} R &= 1 - \frac{\int_0^1 \rho(x)dx}{\int_0^1 \lambda(x)} \\ &= 1 - \frac{\left[\frac{1}{6}x^6\right]_0^1}{\left[\frac{.5}{3}x^3 + \frac{.3}{5}x^5 + \frac{.2}{6}x^6\right]_0^1} = 1 - \frac{1/6}{5/30 + 3/50 + 2/60} = \frac{14}{39} = 0.359 \end{aligned}$$

ii) If the code length is $N = 3900$ and we know that the rate is $R = K/N = K/3900 = 14/39$, we conclude that $K = 1400$. Hence, the parity-check matrix has $N - K = 2500$ rows and 3900 columns.

To find the number of ones in the parity check matrix, there are two ways to proceed. One is to note that 100 % of edges are connected to degree-6 constraint nodes because $\rho(x) = x^5$. Equivalently, all constraint nodes have degree 6, hence there are $6 \times 2500 = 15000$ ones in the parity-check matrix. [10%]

The other way is to calculate the average degree of a variable node using the formula (given in the data sheet): $\bar{d}^\ell = \left(\int_0^1 \lambda(x)dx\right)^{-1} = \frac{50}{13}$, and then compute the number of ones as $N \times \bar{d}^\ell = 15000$.

Assessor's comment: Most students made an error in calculating the number of ones in the parity check matrix, by using the edge perspective polynomial $\lambda(x)$ directly, instead of using it to first find the average variable node degree.

(c) (A calculator is needed to answer the first part of this question.) For the sum-product algorithm,

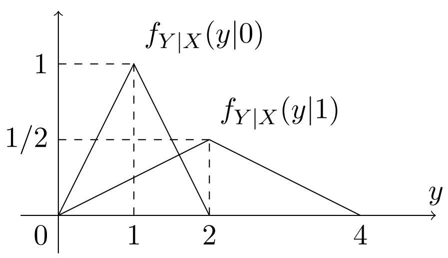
$$\begin{aligned} L_o^{sp}(3) &= 2\tanh^{-1} \prod_{i \neq 3} \tanh \frac{L_i}{2} \\ &= 2\tanh^{-1}\left[\tanh\left(\frac{3.2}{2}\right) \cdot \tanh\left(\frac{-0.6}{2}\right) \cdot \tanh\left(\frac{4.4}{2}\right) \cdot \tanh\left(\frac{2.8}{2}\right) \cdot \tanh\left(\frac{5.7}{2}\right)\right] \\ &= -0.47 \end{aligned}$$

while for the min-sum algorithm,

$$\begin{cases} \text{sign}(L_o^{ms}(3)) = \prod_{i \neq 3} \text{sign}(L_i) = -1 \\ |L_o^{ms}(3)| = \min_{i \neq 3} |L_i| = 0.6 \end{cases}$$

hence $L_o^{ms}(3) = -0.6$. [25%]

(d) i) The following graph shows both densities and may be helpful (but not essential) for visualising the answer:



We divide the interval $[0,4]$ into three intervals $[0,1]$, $[1,2]$ and $[2,4]$ and consider each separately:

For $y \in [0,1]$, we have

$$\begin{cases} f_{Y|X}(y|0) = g_1(y) = y \\ f_{Y|X}(y|1) = g_2(y) = \frac{1}{4}y \end{cases}$$

so the log-likelihood ratio is

$$L(y) = \log \frac{y}{\frac{1}{4}y} = \log 4.$$

For $y \in [1,2]$, we have

$$\begin{cases} f_{Y|X}(y|0) = g_1(y) = 2 - y \\ f_{Y|X}(y|1) = g_2(y) = \frac{1}{4}y \end{cases}$$

hence

$$L(y) = \log \frac{2-y}{\frac{1}{4}y} = \log 4 + \log(\tfrac{2}{y} - 1).$$

For $y \in [2,4]$, we have [25%]

$$\begin{cases} f_{Y|X}(y|0) = g_1(y) = 0 \\ f_{Y|X}(y|1) = g_2(y) = 1 - \frac{1}{4}y \end{cases}$$

hence

$$L(y) = -\infty.$$

ii) The 4th received symbol was definitely a 1 since $f_{Y|X}(2.2|0) = 0$, and since it is a repetition code, the transmitted codeword is $[1,1,1,1,1]$. [10%]

Question 4

(a) i) The transition matrix is:

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

[15%]

ii) The mutual information for the cascade channel is

$$I(X;Y) = H(Y) - H(Y|X) = H(Y) - \sum_{x\in\{0,1\}} P(x)H(Y|X=x). \qquad (13)$$

Next, we find that

$$H(Y|X=0) = H(Y|X=1) = H(\,\{(1-p)(1-\alpha), \alpha, p(1-\alpha)\}\,),$$

where $H(\{p_1,p_2,p_3\})$ denotes the entropy of the pmf $\{p_1,p_2,p_3\}$. Therefore,

$$I(X;Y) = H(Y) - H(\,\{(1-p)(1-\alpha), \alpha, p(1-\alpha)\}\,). \qquad (14)$$

Since the channel is symmetric in inputs 0 and 1, the mutual information is maximised by choosing $P(X=0) = P(X=1) = \frac{1}{2}$. (This can also be explicitly checked by using an arbitrary input distribution of the form $\{x, 1-x\}$ and optimising.) [25%]

With $P(X=0) = P(X=1) = \frac{1}{2}$, we get $P(Y=1) = P(Y=0) = \frac{1-\alpha}{2}$, $P(Y=\epsilon) = \alpha$. Therefore, from (14), the capacity is

$$\mathcal{C} = \max_{P_X} I(X;Y) = H\left(\left\{\frac{1-\alpha}{2}, \alpha, \frac{1-\alpha}{2}\right\}\right) - H\left(\{(1-p)(1-\alpha), \alpha, p(1-\alpha)\}\right),$$

which after simplification yields $\mathcal{C} = (1-\alpha)(1-H_2(p))$. Note that this is the product of the capacities of the two channels in the cascade. (In general, the capacity of the cascade channel may not be the product of the individual capacities.)

(b) i)

$$2^0 = 1, 2^1 = 2, 2^2 = 4, 2^3 = 3, 2^4 = 1$$

and

$$3^0 = 1, 3^1 = 3, 3^2 = 4, 3^3 = 2, 3^4 = 1$$

so the multiplicative order of 2 and 3 is 4.

There cannot be an element of multiplicative order 3 because the order of the multiplicative group is 4 and, according the Lagrange's theorem, the order of any element in the group must divide 4. [10%]

ii) Take two rows of the Fourier matrix with primitive element $\alpha = 2$ or $\alpha = 3$. Two possible answers depending on whether 2 or 3 is used as a primitive element (the two solutions are denoted (*) and (**) in the rest of the crib): [10%]

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

iii) The code is of dimension 2. It has $5^2 = 25$ codewords. [10%]

iv) Any elementary row operations on a parity-check matrix yields another parity-check matrix for the same code. For example, we can obtain an equivalent parity-check matrix by replacing the second row by the sum of the two rows, i.e., [10%]

$$(*) : \mathbf{H}' = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 3 & 0 & 4 \end{bmatrix} \text{ or } (**) : \mathbf{H}' = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 4 & 0 & 3 \end{bmatrix}$$

(note that this answer is not unique and any equivalent parity-check matrix to the **H** speficied in quesion (b) is a valid answer here.)

v) We multiply the sequence with the parity-check matrix to check if it is a codeword,

$$(*) : [1, 2, 4, 3] \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 3 \end{bmatrix}^T = [0, 0],$$

yes it is a codeword, or

$$(**) : [1, 2, 4, 3] \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 3 & 4 & 2 \end{bmatrix}^T = [0, 1],$$

no it is not a codeword (note that the correct answer varies here according to the path chosen in part (ii)). [5%]

vi) Here the student should notice that there were two possible choices in part (ii) and state the remaining solution, i.e. (**) if (*) was chosen in (b), or (*) if (**) was chosen in (b).

Assessor's comment: Some students gave a parity check matrix consisting of the first 3 rows of the DFT matrix. This is also ok, since such a code will have lower dimension $k = 1$, but will still correct one error, since it has $d_{min} = 4$. [15%]