4F5 Advanced Information Theory and Coding 2022 Crib

Question 1

(a) The proof relies on the fact that $\mathrm{R}_8(2^n) = 0$ for $n \geq 3$ and hence $\mathrm{R}_8(2^n)$ can only be non-zero for $n = 0, 1, 2$.

- If $m = 2k$ is even, $\mathrm{R}_8(3^m + 1) = \mathrm{R}_8(\mathrm{R}_8(3^2)^k + 1) = \mathrm{R}_8(1^k + 1) = 2$, but $\mathrm{R}_8(2^n) = 2$ implies $n = 1$, in which case $3^m = 2^n - 1 = 1$, i.e., $m = 0$.
- If $m = 2k + 1$ is odd, $\mathrm{R}_8(3^m + 1) = \mathrm{R}_8(3\,\mathrm{R}_8(3^2)^k + 1) = 4$ but $\mathrm{R}_8(2^n) = 4$ implies $n = 2$ and hence $3^m = 2^n - 1 = 3$, i.e., $m = 1$.

[10%]

(b) $385 = 5 \times 7 \times 11$ hence by the Chinese Remainder Theorem only numbers whose remainders with respect to $5, 7$ and 11 aren't zero have multiplicative inverses. $\mathrm{R}_7(28) = 0$ so 28 is not invertible. [10%]

(c)

$$\boldsymbol{X} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \end{bmatrix}$$

[10%]

(d) By Lagrange's theorem, the order of any element in a group must divide the order of the group, but the order of the multiplicative group of GF(32) is 31 which is a prime number hence the order of $1 + X + X^2 + X^4$ must be 31.

[10%]

(e) The code has dimension $K = 2$ and contains $q^K = 4^2 = 16$ codewords. It is clearly MDS (any subset of 2 columns is an invertible matrix) so it satisfies the Singleton Bound with equality $d_{\min} = N - K + 1 = 3 - 2 + 1 = 2$.

[10%]

(f) (i) We compute Euler's function $\varphi(247) = (13-1) \times (19-1) = 216$ and find the multiplicative inverse of our public exponent $e = 97$ in $\mathbb{Z}_{216}$ using Euclid's extended algorithm

n_1	n_2	q	r	(a_1, b_1)	(a_2, b_2)
216	97	2	22	$(1, 0)$	$(0, 1)$
22	97	4	9	$(1, -2)$	$(0, 1)$
22	9	2	4	$(1, -2)$	$(-4, 9)$
4	9	2	1	$(9, -20)$	$(-4, 9)$
4	1			$(9, -20)$	$(-22, \mathbf{49})$

to yield our secret exponent $d = 49$, or $d = 110001$ in binary. To decrypt, we need to compute $R_{247}(175^d)$. We prepare the ground by computing the powers of 175 in $\mathbb{Z}_{247}$ as $(175, 175^2, 175^4, 175^8, 175^{16}, 175^{32}) = (175, 244, 9, 81, 139, 55)$ and hence recover the encrypted message

$$X = R_{247}(24^{49}) = R_{247}(175 \times 175^{16} \times 175^{32}) = R_{247}(175 \times 139 \times 55) = 123$$

[40%]

(ii) We compute $R_m(10^e)$ where (m, e) is the public modulus and exponent of our possible correspondents. For Juliet, we obtain $R_{187}(10^7) = 175$ and for Romeo, $R_{391}(10^5) = 295$ so clearly the message is from Juliet.

[10%]

Generally well-answered question. It was a slightly new format where 1 1/2 question in previous years have been compressed into one question with two halves. The first half on mathematics was done fairly well. It was surprising to see a few students attempt to calculate the order of $1+X^2+X^5$ *in GF(32) the "hard" way without noticing that the order of the field 31 is a prime number and hence, by Lagrange's theorem, the order of every element has to be 31 since it has to divide the order of the field. This sort of question had been asked in the past so those who failed to notice probably did not attempt enough past tripos questions when preparing for the exam. The second part of the question on cryptography was done with a variable level of success as it did require quite a few calculations which many stumbled. Mere calculation errors were barely penalised but many did not quite lay out the correct operations and struggled to get points as a result.*

Question 2

(a) The multiplicative group has order 28 and hence the possible lengths are divisors of 30, excluding 1 and 2 for which Reed-Solomon codes cannot be defined as they would either have rate 0 or not correct any errors, i.e., 4,7,14,28

[10%]

(b) $\alpha^2 = 28, \alpha^3 = 17, \alpha^4 = 1$, hence the length is $N = 4$.

[10%]

(c)

$$\boldsymbol{F} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & \alpha & \alpha^2 & \alpha^3 \\ 1 & \alpha^2 & 1 & \alpha^2 \\ 1 & \alpha^3 & \alpha^2 & \alpha \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 12 & 28 & 17 \\ 1 & 28 & 1 & 28 \\ 1 & 17 & 28 & 12 \end{bmatrix}$$

[10%]

(d) The parity-check matrix verifies that the first two coefficients of the DFT of a codeword are zero, hence it consists of the first two columns of the DFT matrix, transposed, i.e.,

$$\boldsymbol{H} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 12 & 28 & 17 \end{bmatrix}$$

[10%]

(e) We compute the DFT matrix, using the hint to compute $1/N = 1/4 = 22$ in GF(29), as

$$F^{-1} = \frac{1}{N} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & \alpha^3 & \alpha^2 & \alpha \\ 1 & \alpha^2 & 1 & \alpha^2 \\ 1 & \alpha & \alpha^2 & \alpha^3 \end{bmatrix} = \begin{bmatrix} 22 & 22 & 22 & 22 \\ 22 & 26 & 7 & 3 \\ 22 & 7 & 22 & 7 \\ 22 & 3 & 7 & 26 \end{bmatrix}$$

The encoder matrix consists of the last 2 rows of the inverse DFT matrix, i.e.,

$$\boldsymbol{G} = \begin{bmatrix} 22 & 7 & 22 & 2 \\ 22 & 3 & 7 & 26 \end{bmatrix}$$

[10%]

(f) Examining the first two digits of $\boldsymbol{R}$, we note that the DFT of the error sequence is generated by the recurrence relation $E_{k+1} = 28 \times E_k$ and hence the full error sequence in the frequency domain is $\boldsymbol{E} = [28, 1, 28, 1]$. We substract the error sequence from the received word in the frequency domain to obtain $\boldsymbol{R} - \boldsymbol{E} = [0, 0, 1, 11]$ and cut out the information digits $\boldsymbol{U} = [1, 11]$.

[20%]

(g) The D transform (equivalent to the z transform with $D = z^{-1}$) of the recurrence relation is determined by the locations 0 and 1 of the erasures (or potential errors), i.e.,

$$C(D) = (1 - \alpha^0 D)(1 - \alpha^1 D) = (1 - D)(1 - 12D) = 1 - 13D + 12D^2$$

resulting in the recurrence relation

$$E_k - 13E_{k-1} + 12E_{k-2} = 0$$

or, equivalently,

$$E_k = 13E_{k-1} - 12E_{k-2}.$$

[20%]

(h) We apply the recurrence relation to the error sequence starting with $[E_0, E_1] = [R'_0, R'_1] = [26, 4]$ to yield

$$\boldsymbol{E} = [26, 4, 1, 23]$$

and subtract from $\boldsymbol{R'} = [26, 4, 13, 15]$ to yield $\boldsymbol{C} = [0, 0, 12, 21]$ and hence recover the information digits $\boldsymbol{U} = [12, 21]$.

[10%]

This was a popular question that was very similar in style to previous years' question. The aim of this question is to test whether students have understood the concepts of linear codes over Galois fields and the encoder and decoder methods learned for Reed Solomon codes. Students' performance on this question has improved year on year as a result of improvements in the delivery of this material during the year and in the lecture notes. This year decoding for erasure channels was included in the question and the majority of students succeeded in doing this without fault.

Question 3

(a) Figure 1 illustrates the function E_s. The function is convex, increasing and $E_s(0) = 0$. From the lectures we know that the ensemble average error probability over all randomly generated codes is

$$\bar{p}_e \leq e^{-n(\rho R - E_s(\rho))}$$

for any parameter $\rho \in [0, 1]$ that can be optimised. Therefore, the ensemble average error probability vanishes exponentially with n as long as $\rho R > E_s(\rho)$. This means that there exists at least a code that meets this performance. Since $E_s(\rho)$ is convex, increasing and $E_s(0) = 0$, the smallest slope ρR that allows for a positive difference has to be such greater than $E'_s(0)$, where $E'_s(\rho)$ is the first derivative of $E_s(\rho)$. Since $E'_s(0) = H(V)$ we have that for $R > H(V)$ there exists a code of rate $R = \frac{1}{n} \log M$ whose error probability vanishes exponentially with n.

[20%]

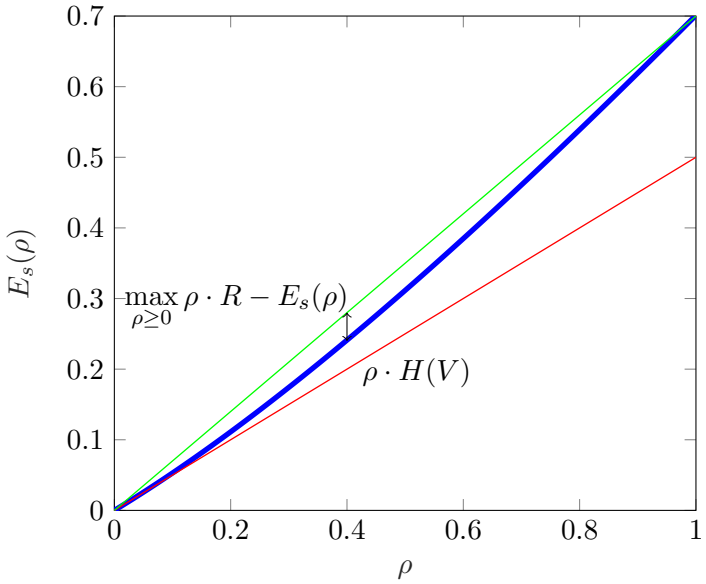

Figure 1: Function $E_s(\rho)$ for a BMS with $P_V(0) = 0.11$. The entropy of the source is $H(V) = 0.5$.

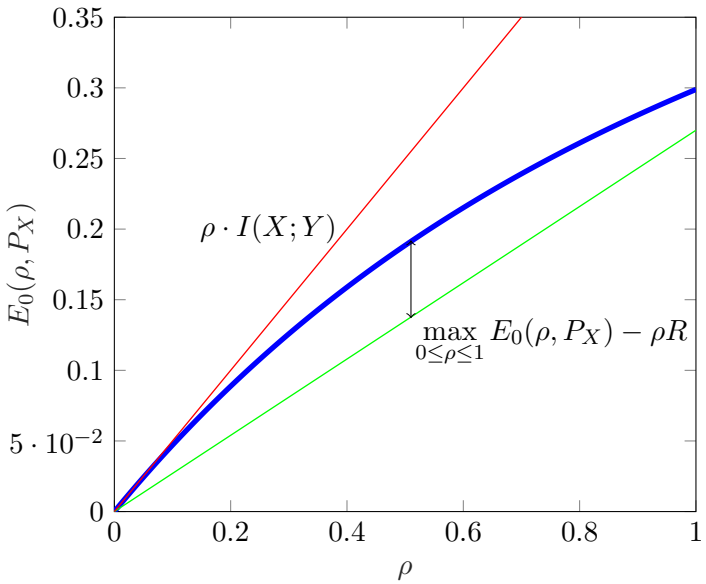

Figure 2: $E_0(\rho, P_X)$ function for a BSC with $\delta = 0.11$ with an equiprobable input distribution. The mutual information of the channel is $I(X;Y) = 0.5$.

(b) Figure 2 illustrates the function E_0. The function is concave, increasing and $E_0(0, P_X) = 0$.

From the lectures we know that the ensemble average error probability over all randomly generated codes is

$$\bar{p}_e \leq e^{-n(E_0(\rho,P_X)-\rho R)}$$

for any parameter $\rho \in [0, 1]$ that can be optimised. Therefore, the ensemble average error probability vanishes exponentially with n as long as $\rho R < E_0(\rho, P_X)$. This means that there exists at least a code that meets this performance. Since $E_0(\rho, P_X)$ is convex, increasing and $E_0(0, P_X) = 0$, the largest slope ρR that allows for a positive difference has to be such smaller than $E_0'(0, P_X)$, where $E_0'(\rho, P_X)$ is the first derivative of $E_0(\rho, P_X)$. Since $E_0'(0, P_X) = I(X;Y)$ we have that for $R < I(X;Y)$ there exists a code of rate $R = \frac{1}{n}\log M$ whose error probability vanishes exponentially with n.

[20%]

(c) (i)

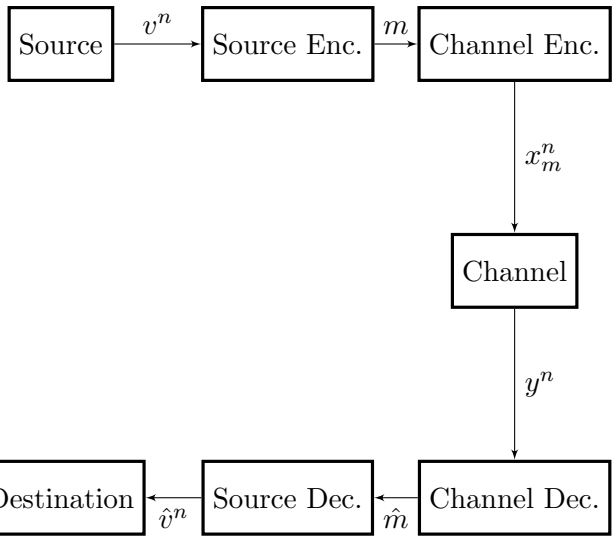

Figure 3: Block diagram of a separate source-channel coding system.

[10%]

(ii) We write the error probability of this system as

$$\begin{aligned} p_e &= \mathbb{P}[\hat{V}^n \neq V^n] && (1)\\ &= \mathbb{P}[\hat{V}^n \neq V^n \cap \hat{m} = m] + \mathbb{P}[\hat{V}^n \neq V^n \cap \hat{m} \neq m] && (2)\\ &= \mathbb{P}[\hat{V}^n \neq V^n | \hat{m} = m] \cdot \mathbb{P}[\hat{m} = m] + \mathbb{P}[\hat{V}^n \neq V^n | \hat{m} \neq m] \cdot \mathbb{P}[\hat{m} \neq m] && (3)\\ &\leq \mathbb{P}[\hat{V}^n \neq V^n | \hat{m} = m] + \mathbb{P}[\hat{m} \neq m] && (4) \end{aligned}$$

where we have upper bounded $\mathbb{P}[\hat{m} = m]$ and $\mathbb{P}[\hat{V}^n \neq V^n | \hat{m} \neq m]$ by one.

[20%]

(iii) From (4) we see that the first term corresponds to the probability of a source coding error, while the second corresponds to the probability of a channel coding error. Therefore, we bound each term by their corresponding error exponent, as done in the lectures.

$$\bar{p}_e \leq e^{-n(\rho_s R - E_s(\rho_s))} + e^{-n(E_0(\rho_c,P_X)-\rho_c R)}. \quad (5)$$

[15%]

(iv) If $R > H(V)$ we know that the first term in (5) vanishes. Likewise, if $R < I(X;Y)$, the second term in (5)vanishes. Thus any rate $H(V) < R < I(X;Y)$ makes the error probability vanish.

[15%]

Generally well answered question. Parts (a) and (b) were almost straight from the notes and answered correctly, although some candidates did not fully answer the question. Many candidates answered correctly Part (c)(i). Part (c)(ii) was mixed. Given Part (c)(ii), Parts (c)(iii) and (c)(iv) were answered generally well.

Question 4

(a) (i) Consider the random variable $p_e(\mathsf{C}_n)$ and directly apply Markov's inequality to its tail

$$\mathbb{P}\big[p_e(\mathsf{C}_n)^s \geq a^s\big] \leq \frac{\mathbb{E}[p_e(\mathsf{C}_n)^s]}{a^s} \tag{6}$$

where $a > 0, s > 0$. If we choose

$$a^s = \gamma_n \mathbb{E}[p_e(\mathsf{C}_n)^s]$$

we obtain the result.

[15%]

(ii) The derivation in question (i) implies that with probability smaller than $\frac{1}{\gamma_n}$ we will have

$$p_e(\mathcal{C}_n) \geq \gamma_n^{\frac{1}{s}}\, \mathbb{E}[p_e(\mathsf{C}_n)^s]^{\frac{1}{s}}. \tag{7}$$

which is equivalent to saying that with probability higher than $1 - \frac{1}{\gamma_n}$

$$p_e(\mathcal{C}_n) < \gamma_n^{\frac{1}{s}}\, \mathbb{E}[p_e(\mathsf{C}_n)^s]^{\frac{1}{s}}. \tag{8}$$

[10%]

(iii) If the sequence γ_n is such that

$$\lim_{n\to\infty} \gamma_n = \infty$$

with probability approaching one we have that

$$p_e(\mathcal{C}_n) < \gamma_n^{\frac{1}{s}}\, \mathbb{E}[p_e(\mathsf{C}_n)^s]^{\frac{1}{s}}. \tag{9}$$

or equivalently

$$-\frac{1}{n}\log p_e(\mathcal{C}_n) > -\frac{1}{n}\log \gamma_n^{\frac{1}{s}} - \frac{1}{n}\log \mathbb{E}[p_e(\mathsf{C}_n)^s]^{\frac{1}{s}}. \tag{10}$$

Thus, we see that if

$$\lim_{n\to\infty} -\frac{1}{n}\log \gamma_n^{\frac{1}{s}} = 0$$

we obtain the desired result. [15%]

(iv) For $s = 1$ we have that with high probability

$$-\frac{1}{n}\log p_e(\mathcal{C}_n) > -\frac{1}{n}\log \mathbb{E}[p_e(\mathsf{C}_n)]. \tag{11}$$

which means that with high probability the error exponent of a randomly generated code $\mathcal{C}_n$ will be higher than the exponent of the ensemble average error probability. [10%]

(b) (i) The Stein exponent is the maximum exponent (over all tests) of the pairwise error probability in hypothesis testing where the observations are i.i.d. and one of the pairwise error probabilities is bounded by a constant. Specifically, assuming that $\epsilon_0(P_Y, T) \leq \alpha$, it is

$$\lim_{n\to\infty} -\frac{1}{n}\log \epsilon_1(Q_Y, T) = D(P\|Q).$$

It provides an indication as to how difficult is to tell distributions P, Q apart from i.i.d. observations. The more different P and Q are, the higher $D(P\|Q)$ and the lower the error probabillity. [15%]

(ii) The Stein exponent is the relative entropy

$$D(P\|Q) = \sum_y P(y) \log \frac{P(y)}{Q(y)} \tag{12}$$

$$= \sum_y P(y) \log \frac{2^{-y}}{q \cdot p^{y-1}} \tag{13}$$

$$= \sum_y P(y) \log \frac{\frac{p}{q}}{(2p)^y} \tag{14}$$

$$= \log \frac{p}{q} \sum_y P(y) - \log(2p) \sum_y P(y) y \tag{15}$$

$$= \log \frac{p}{q} - \log(2p) \sum_y y 2^{-y} \tag{16}$$

$$= \log \frac{p}{q} - 2 \log(2p) \tag{17}$$

$$= -\log(4pq) \tag{18}$$

where $\sum_y y2^{-y} = 2$. [20%]

(iii) The Stein exponent is the relative entropy

$$D(P\|Q) = \sum_y P(y) \log \frac{P(y)}{Q(y)} \tag{19}$$

$$= \sum_y P(y) \log \frac{\alpha e^{-y}}{\frac{1}{A}} \tag{20}$$

$$= \sum_y P(y) \log(\alpha A) - \sum_y P(y) \cdot y \tag{21}$$

$$= \log \alpha + \log A - \mathbb{E}_P[Y]. \tag{22}$$

[15%]

Part (a)(i) was answered correctly by many candidates; a common mistake was not using Markov's inequality correctly. Most candidates answered Part (a)(ii) correctly. Part (a)(iii) proved difficult although several candidates answered correctly. Part (a)(iv) was mixed. Part (b)(i) was generally answered correctly, although some candidates did not answer it well despite being almost straight from the notes. Parts (b)(ii) and b(iii) involved calculations and were generally answered well.