4F5 Advanced Information Theory and Coding 2023 Crib

Question 1

(a) If the second operand is divisible by 9, its digits must sum to a multiple of 9, hence $i_2 = 0$ or 9. In Problem 1.3.3, we showed that the "casting out nines" method is a simple consequence of the property of remainders and implies that the sum modulo 9 of digits of the operands of a sum must be equal to the sum modulo 9 of the digits of the result. The sum is 1 on the left and $i_2 + 2$ on the right, hence $R_9(i_2 + 2) = 1$ implying that $i_2 = 8$. We can verify using a calculator that $i_1 = 9$. [10%]

(b) $R_{144}(100n) = 12$ implies that there exists a q such that $100n - 144q = 12$. Let us compute the greatest common divisor of 100 and 144 using the extended Euclid algorithm

n_1	n_2	a	b	a'	b'
144	100	1	0	0	1
44	100	1	-1	0	1
44	12	1	-1	-2	3
8	12	7	-10	-2	3
8	4	7	-10	-9	13
0	4	25	-36	-9	13

implying that $\gcd(100, 144) = 4 = -9 \times 144 + 13 \times 100$. We can multiply this equation by 3 to yield $12 = -27 \times 144 + 39 \times 100$ which yields a solution $n = 39$, i.e., $R_{144}(100 \times 39) = 12$ as required. Another solution could have been found by stopping the extended algorithm after the 3rd step when we found that $12 = -2 \times 144 + 3 \times 100$, giving the solution $n = 3$, i.e., $R_{144}(100 \times 3) = 12$. [10%]

(c) We note that $143 = 11 \times 13$ and hence use the Chinese Remainder Theorem to determine the residues of 27 to be $(5, 1)$ with respect to the moduli $(11, 13)$. The inverse of 5 in multiplication mod 11 is easily determined by inspection to be 9, and the inverse of 1 in multiplication mod 13 is obviously 1, so the residues of the inverse of 27 are $(9, 1)$. We can now either use the method described in the notes to obtain a number from its residues, or proceed by inspection to find an n such that $R_{11}(13n + 1) = 9$, which one readily finds to be $n = 4$, so the inverse of 27 is 53. [10%]

(d) The multiplicative group of GF(8) has order 7 which is a prime number. Hence, irrespective of which irreducible polynomial of degree 3 is used to define the group (indeed, we are not given that polynomial!) the order of all non-neutral elements including X is 7, the order of the group. Hence, $X^5 \cdot X^4 = X^{R_7(5+4)} = X^2$. [10%]

(e) By row manipulation, we bring the generator matrix into systematic form

$$\boldsymbol{G}_{\text{sys}} = \begin{bmatrix}\boldsymbol{I}_2 & \boldsymbol{P}\end{bmatrix} = \begin{bmatrix}1 & 0 & 4\\ 0 & 1 & 2\end{bmatrix}$$

and obtain the corresponding systematic parity-check matrix

$$\boldsymbol{H}_{\text{sys}} = \begin{bmatrix}-\boldsymbol{P}^T & \boldsymbol{I}_1\end{bmatrix} = \begin{bmatrix}1 & 3 & 1\end{bmatrix}$$

[10%]

(f) (a) We precompute the exponential powers of 2

2^1	2^2	2^4	2^8	2^{16}
2	4	16	20	46

and obtain $y_A = 2^{x_A} = 2^{27} = 2^{16+8+2+1} = 46 \times 20 \times 4 \times 2 = 44.$ [15%]

(b) We precompute the exponential powers of y_B

43^1	43^2	43^4	43^8	43^{16}
43	20	46	51	5

and obtain the common secret $y_B^{x_A} = 43^{27} = 43^{16+8+2+1} = 5 \times 51 \times 20 \times 43 = 56$. The binary key is 56 in binary, i.e., 111000. [15%]

(c) The new method generates the public keys $y_A = 2^{x'_A}$ and $y_B = 2^{x'_B}$ and obtains a common secret because

$$s = y_B^{x'_A} = (2^{x'_B})^{x'_A} = 2^{x'_A x'_B} = (2^{x'_A})^{x'_B} = y_A^{x'_B}$$

where s stands for the commen secret. However, note that we can express s in terms of x_A and x_B to obtain

$$s = 2^{x'_A x'_B} = 2^{2^{x_A}2^{x_B}} = 2^{2^{x_A+x_B}}$$

where the power of 2 in the exponent is taken in $\mathbb{Z}_{58}$ because the order of the multiplicative group of $\mathbb{Z}_{59}$ is 58, whereas the sum in the "double-exponent" is in $\mathbb{Z}_{\varphi(58)}$ where $\varphi(58) = (2-1)(29-1) = 28$ is the Euler function of 58. This is because exponentation in $\mathbb{Z}_{58}$ is over the multiplicative subgroup of invertible elements which has order $\varphi(58) = 28$. Hence, using double exponentiation reduces the search space for a brute force attack from 58 to 28 since there are only 28 possible values for the exponent of 2 that gives the common key, thereby weakening the method rather than strengthening it as claimed by Eve. [20%]

Question 2

(a) A binary single parity check bit achieves this. The easiest way to implement this is for the data on the last hard disk to be the bit-wise XOR of the other 4. If any single hard disk becomes unavailable, its content can be recovered by taking the bit-wise XOR of the remaining 4. [10%]

(b) Only a Maximum Distance Separable (MDS) code can achieve this, such as a Reed-Solomon (RS) code. Any RS code of block length 5 can be used. For example, one could use an RS code over GF(256) which would have the advantage of operating directly on bytes. One would need to pick a primitive 5th root of unity rather than an element of the maximum order 255 to get a block length of 5. Since 5 divides 255, there is a good chance of there being an element of order 5. In such a setup, the 3 first hard disks could contain information and the remaining 2 hard disks would contain the byte-wise two parity symbols of GF(256) obtained via a systematic encoder matrix from the corresponding information bytes in the first 3 hard disks. Alternatively, one could use an RS code of length 5 over GF(16), and operate on nibbles (half-bytes, 4 bits, as seen in the IA Microprocessor lab) instead of bytes. [10%]

(c) We compute powers of X modulo $\pi(X)$ to obtain $X^0 = 1, X^1 = X, X^2, X^3, X^4 = 1+X+X^2+X^3, X^5 = X(1+X+X^2+X^4) = 1$ and conclude that the order of X in the multiplicative group is 5. Hence the code length is $N = 5$. [10%]

(d) A parity-check matrix of an RS code consists of the first 2 rows of the DFT matrix

$$\boldsymbol{H} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & X & X^2 & X^3 & 1+X+X^2+X^3 \end{bmatrix}$$

[10%]

(e) We begin by transforming the parity-check matrix into systematic form

$$\begin{aligned}
\boldsymbol{H} &= \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & X & X^2 & X^3 & 1+X+X^2+X^3 \end{bmatrix} \begin{matrix} L_1 \\ \to L_2 + X^3 L_1 \end{matrix} \\
\boldsymbol{H'} &= \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1+X^3 & X+X^3 & X^2+X^3 & 0 & 1+X+X^2 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ \alpha^8 & \alpha^{14} & \alpha^{10} & 0 & \alpha^7 \end{bmatrix} \begin{matrix} L_1 \\ \to \alpha^{-7} L_2 \end{matrix} \\
\boldsymbol{H''} &= \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ \alpha & \alpha^7 & \alpha^3 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1+X & 1+X+X^2 & 1+X+X^2+X^3 & 0 & 1 \end{bmatrix} \begin{matrix} \to L_1 + L_2 \\ L_2 \end{matrix} \\
\boldsymbol{H}_{\text{sys}} &= \begin{bmatrix} X & X+X^2 & X+X^2+X^3 & 1 & 0 \\ 1+X & 1+X+X^2 & 1+X+X^2+X^3 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} \alpha^{12} & \alpha^{13} & \alpha^4 & 1 & 0 \\ \alpha & \alpha^7 & \alpha^3 & 0 & 1 \end{bmatrix}
\end{aligned}$$

We now obtain the systematic encoder matrix from $\boldsymbol{H}_{\text{sys}} = \begin{bmatrix} -\boldsymbol{P}^T & \boldsymbol{I}_2 \end{bmatrix}$ to

$$\begin{aligned}
\boldsymbol{G}_{\text{sys}} = \begin{bmatrix} \boldsymbol{I}_3 & \boldsymbol{P} \end{bmatrix} &= \begin{bmatrix} 1 & 0 & 0 & \alpha^{12} & \alpha \\ 0 & 1 & 0 & \alpha^{13} & \alpha^7 \\ 0 & 0 & 1 & \alpha^4 & \alpha^3 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 & X & 1+X \\ 0 & 1 & 0 & X+X^2 & 1+X+X^2 \\ 0 & 0 & 1 & X+X^2+X^3 & 1+X+X^2+X^3 \end{bmatrix}
\end{aligned}$$

[20%]

(f) The erasures are in position 1,2 in the codeword. We begin by replacing the erasures with zeros and consider the received word

$$\boldsymbol{r} = [1+X+X^3, 0, 0, X, 0]$$

and compute its DFT

$$\boldsymbol{R} = \boldsymbol{rF} = [1+X+X^3, 0, 0, X, 0] \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & X & X^2 & X^3 & X^4 \\ 1 & X^2 & X^4 & X & X^3 \\ 1 & X^3 & X & X^4 & X^2 \\ 1 & X^4 & X^3 & X^2 & X \end{bmatrix}$$

$$= [\alpha^{11}, 0, 0, \alpha^{12}, 0] \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & \alpha^{12} & \alpha^9 & \alpha^6 & \alpha^3 \\ 1 & \alpha^9 & \alpha^3 & \alpha^{12} & \alpha^6 \\ 1 & \alpha^6 & \alpha^{12} & \alpha^3 & \alpha^9 \\ 1 & \alpha^3 & \alpha^6 & \alpha^9 & \alpha^{12} \end{bmatrix}$$

$$= [\alpha^{11}+\alpha^{12}, \alpha^{11}+\alpha^3, \alpha^{11}+\alpha^9, \alpha^{11}+1, \alpha^{11}+\alpha^6]$$

$$= [1+X^3, X^2, 1+X+X^2+X^3, X+X^3, 1+X]$$

$$= [\alpha^8, \alpha^9, \alpha^3, \alpha^{14}, \alpha].$$

The first two symbols of the received vector in the frequency domain are non-zero, and since the codeword is zero in these positions we now know the first two symbols of the error vector in the frequency domain, $[E_0, E_1] = [\alpha^8, \alpha^9]$. The D transform of the recurrence relation is determined by the locations 1 and 2 of the erasures (or potential errors if those code digits weren't equal to zero), i.e.,

$$C(D) = (1-\beta D)(1-\beta^2 D) = 1 + (X+X^2)D + X^3 D^2 = 1 + \alpha^{13} D + \alpha^6 D^2,$$

resulting in the recurrence relation

$$E_k + \alpha^{13} E_{k-1} + \alpha^6 E_{k-2} = 0,$$

or, equivalently,

$$E_k = \alpha^{13} E_{k-1} + \alpha^6 E_{k-2}.$$

We apply the recurrence relation to obtain

$$\begin{cases} E_2 &= \alpha^{13} E_1 + \alpha^6 E_0 = \alpha^7 + \alpha^{14} = 1 + X^2 + X^3 = \alpha^5 \\ E_3 &= \alpha^{13} E_2 + \alpha^6 E_1 = \alpha^3 + \alpha^0 = X + X^2 + X^3 = \alpha^4 \\ E_4 &= \alpha^{13} E_3 + \alpha^6 E_2 = \alpha^2 + \alpha^{11} = X + X^2 + X^3 = \alpha^4 \end{cases}$$

We subtract (or, equivalently, add) the error vector from the received vector to obtain the codeword in the frequency domain

$$\boldsymbol{C} = [0, 0, X, X^2, 1+X^2+X^3] = [0, 0, \alpha^{12}, \alpha^9, \alpha^5].$$

We now take the inverse DFT to obtain the codeword

$$\boldsymbol{c} = \boldsymbol{CF^{-1}} = [0, 0, \alpha^{12}, \alpha^9, \alpha^5] \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & \alpha^3 & \alpha^6 & \alpha^9 & \alpha^{12} \\ 1 & \alpha^6 & \alpha^{12} & \alpha^3 & \alpha^9 \\ 1 & \alpha^9 & \alpha^3 & \alpha^{12} & \alpha^6 \\ 1 & \alpha^{12} & \alpha^9 & \alpha^6 & \alpha^3 \end{bmatrix}$$

$$= [\alpha^{12}+\alpha^9+\alpha^5, \alpha^3+\alpha^6+\alpha^2, \alpha^9+\alpha^{12}+\alpha^{14}, \alpha^0+\alpha^6+\alpha^{11}, \alpha^6+\alpha^0+\alpha^8]$$

$$= [1+X+X^3, X, X^2+X^3 X, 0]$$

and we finally read the 12 bit key from the systematic part of the codeword, $k = [1101, 0100, 0011]$.

[40%]

Question 3

(a) (i) By Markov's inequality we have that

$$\mathbb{P}\big[p_{e,m}(\mathcal{C}_n)^s \geq 2 \cdot \mathbb{E}[p_{e,m}(\mathcal{C}_n)^s]\big] \leq \frac{1}{2}. \tag{1}$$

[15%]

(ii) The quantity

$$\mathbb{E}\Big[\sum_{m=1}^{M'} Z_m\Big].$$

is the expected number of codewords in $\mathcal{C}_n$ that satisfy the property $p_{e,m}(\mathcal{C}_n) < 2^{1/s} \cdot \mathbb{E}[p_{e,m}(\mathcal{C}_n)^s]^{1/s}$.

[10%]

(iii) We have that

$$\begin{aligned}
\mathbb{E}\Big[\sum_{m=1}^{M} Z_m\Big] &= \mathbb{E}\Big[\sum_{m=1}^{M} \mathbb{1}\Big\{p_{e,m}(\mathcal{C}_n) < 2^{1/s} \cdot \mathbb{E}[p_{e,m}(\mathcal{C}_n)^s]^{1/s}\Big\}\Big] &(2)\\
&= \sum_{m=1}^{M} \mathbb{E}\Big[\mathbb{1}\Big\{p_{e,m}(\mathcal{C}_n) < 2^{1/s} \cdot \mathbb{E}[p_{e,m}(\mathcal{C}_n)^s]^{1/s}\Big\}\Big] &(3)\\
&= \sum_{m=1}^{M} \mathbb{P}\Big[p_{e,m}(\mathcal{C}_n) < 2^{1/s} \cdot \mathbb{E}[p_{e,m}(\mathcal{C}_n)^s]^{1/s}\Big] &(4)\\
&\geq \frac{M}{2} &(5)
\end{aligned}$$

where (5) follows from (1). [20%]

(b) (i) The rate of $\mathcal{C}_n$ is $R = \frac{1}{n}\log_2 M$ while that of $\mathcal{C}'_n$ is

$$R' = \frac{1}{n}\log_2(2M-1) \geq R + \frac{1}{n} \tag{6}$$

[10%]

(ii) If the expected number of codewords that satisfy the property $p_{e,m'}(\mathcal{C}'_n) < 2^{1/s} \cdot \mathbb{E}[p_{e,m'}(\mathcal{C}'_n)^s]^{1/s}$ is at least $\frac{M'}{2} = M - \frac{1}{2}$ (from (5)) it means that this number should be at least M (since $\frac{1}{2}$ is not an integer). If the aforementioned expectation is at least M it means that there exists at least a codebook $\mathcal{C}'_n$ for which are at least M codewords satisfy the required property. Removing the codewords for which the property is not satisfied yields the resulting code. [15%]

(iii) We have that with independent codewords generated with distribution Q_{X^n}

$$\begin{aligned}
p_{e,m'}(\mathcal{C}'_n)^s &\leq \Bigg(\sum_{\bar{m}\neq m'}\sum_{y^n}\sqrt{W^n(y^n|x^n(\bar{m}))W^n(y^n|x^n(m'))}\Bigg)^s &(7)\\
&\leq \sum_{\bar{m}\neq m'}\Bigg(\sum_{y^n}\sqrt{W^n(y^n|x^n(\bar{m}))W^n(y^n|x^n(m'))}\Bigg)^s. &(8)
\end{aligned}$$

Then

$$\mathbb{E}[p_{e,m'}(\mathcal{C}'_n)^s] = \sum_{\bar{m}\neq m'} \sum_{x_1^n \dots x_{M'}^n} Q_{X^n}(x_1^n)\cdots Q_{X^n}(x_{M'}^n)\left(\sum_{y^n}\sqrt{W^n(y^n|x^n(\bar{m}))W^n(y^n|x^n(m'))}\right)^s \tag{9}$$

$$= \sum_{\bar{m}\neq m'} \sum_{x_{m'}^n, x_{\bar{m}}^n} Q_{X^n}(x_{m'}^n)Q_{X^n}(x_{\bar{m}}^n)\left(\sum_{y^n}\sqrt{W^n(y^n|x^n(\bar{m}))W^n(y^n|x^n(m'))}\right)^s \tag{10}$$

$$= (M'-1)\sum_{x^n,\bar{x}^n} Q_{X^n}(x^n)Q_{X^n}(\bar{x}^n)\left(\sum_{y^n}\sqrt{W^n(y^n|x^n)W^n(y^n|\bar{x}^n)}\right)^s \tag{11}$$

where (10) follows since the term $\sum_{y^n}\sqrt{W^n(y^n|x^n(\bar{m}))W^n(y^n|x^n(m'))}$ only depends on m' and $\bar{m}$. Eq. (11) follows since each term in the sum over $\bar{m}\neq m'$ is the same. [20%]

(iv) Putting together parts (b)(ii) and (iii) one gets the result. [10%]

Question 4

(a) (i) Figure 1 illustrates the function E_s. The function is convex, increasing and $E_s(0) = 0$. From the

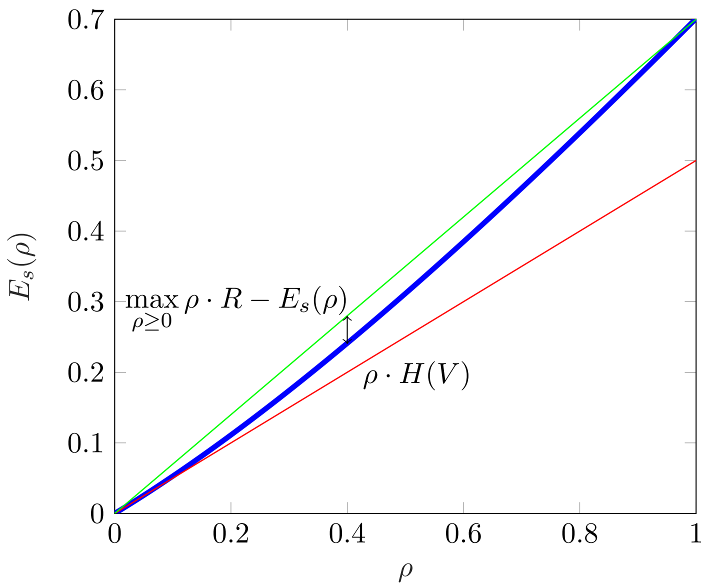

Figure 1: Function $E_s(\rho)$ for a BMS with $P_V(0) = 0.11$. The entropy of the source is $H(V) = 0.5$.

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

$$\bar{p}_e \leq e^{-n(\rho R - E_s(\rho))} \tag{12}$$

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.

[15%]

(ii) Figure 2 shows an example of such a case. The two functions have the same derivative at zero since $H(V) = H(Z)$ and then diverge.

[15%]

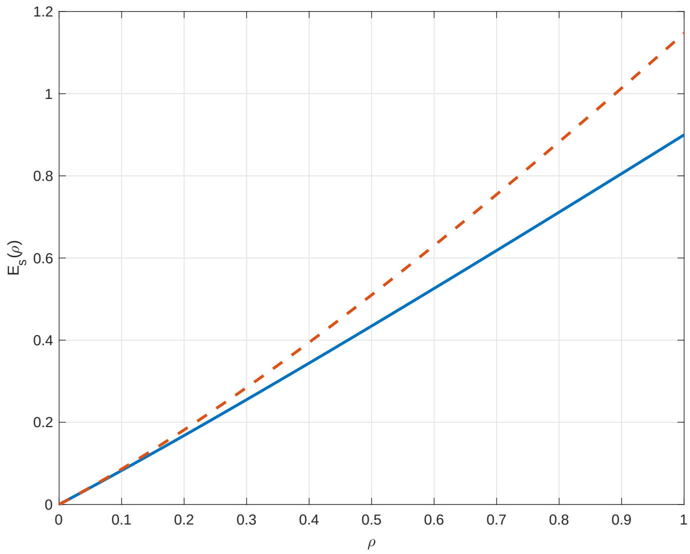

Figure 2: Example of two functions $E_s(\rho)$ for memoryless sources with $H(V) = H(Z)$.

(iii) Since the two sources are defined over the same alphabet $\mathcal{V}$, this is equivalent to encoding a single discrete memoryless source X whose alphabet is $\mathcal{X} = \mathcal{V} \times \mathcal{V} = \mathcal{V}^2$ and probability distribution is $P_X(x) = P_V(v)P_Z(z)$ when $x = (v, z)$. Therefore,

$$E_s^X(\rho) = \log\left(\sum_{(v,z)\in\mathcal{V}^2}\Big(P_V(v)P_Z(z)\Big)^{\frac{1}{1+\rho}}\right)^{1+\rho} \tag{13}$$

$$= \log\left(\sum_{v\in\mathcal{V}} P_V(v)^{\frac{1}{1+\rho}} \sum_{z\in\mathcal{V}} P_Z(z)^{\frac{1}{1+\rho}}\right)^{1+\rho} \tag{14}$$

$$= E_s^V(\rho) + E_s^Z(\rho). \tag{15}$$

[20%]

(b) (i) This is a M-ary hypothesis testing problem where, upon processing observation $y \in \mathcal{Y}$, the test $P_{\hat{V}|Y}$ outputs which of the M known distributions $P_{Y|V=v}, v = 1, \dots, M$ generated the observation. The joint distributions $P_{V=v,Y} = P_V \times P_{Y|V=v}, v = 1, \dots, M$ induce known priors $P_V(v), v = 1, \dots, M$. For a given test, the error probability can be written as

$$p_e(\mathcal{V}, P_{\hat{V}|Y}) = \mathbb{P}[\hat{V} \neq V] \tag{16}$$

$$= \sum_{v,y} P_{V,Y}(v,y)\Big(1 - P_{\hat{V}|Y}(v|y)\Big). \tag{17}$$

[20%]

(ii) We have that

$$\mathbb{P}[(V,Y) \in \mathcal{S}(\gamma)] = \mathbb{P}[(V,Y) \in \mathcal{S}(\gamma) \cap \text{error}] + \mathbb{P}[(V,Y) \in \mathcal{S}(\gamma) \cap \text{no error}] \tag{18}$$

$$\leq \mathbb{P}[\text{error}] + \mathbb{P}[(V,Y) \in \mathcal{S}(\gamma) \cap \text{no error}] \tag{19}$$

$$= \mathbb{P}[\text{error}] + \sum_{\substack{(v,y)\in\mathcal{S}(\gamma) \\ \text{no error}}} P_{V,Y}(v,y) \tag{20}$$

$$\leq \mathbb{P}[\text{error}] + \sum_{\substack{(v,y)\in\mathcal{S}(\gamma) \\ \text{no error}}} \gamma P_Y(y) \cdot 1 \tag{21}$$

$$\leq \mathbb{P}[\text{error}] + \gamma \tag{22}$$

[30%]