Problem 1

Ind	Codeword	Dict	Text	Ind	Codeword	Dict	Text
128	73	I␣	I	147	115	s␣	s
129	32	␣w	␣	148	140	␣whe	␣wh
130	119	wi	w	149	101	en	e
131	105	is	i	150	110	n␣	n
132	115	sh	s	151	144	␣I␣w	␣I␣
133	104	h␣	h	152	130	wis	wi
134	32	␣I	␣	153	132	she	sh
135	128	I␣w	I␣	154	101	ed	e
136	119	we	w	155	100	d␣	d
137	101	er	e	156	151	␣I␣we	␣I␣w
138	114	re	r	157	137	ere	er
139	101	e␣	e	158	139	e␣w	e␣
140	129	␣wh	␣w	159	119	wh	w
141	104	ha	h	160	141	hat	ha
142	97	at	a	161	143	t␣I	t␣
143	116	t␣	t	162	128	I␣a	I␣
144	134	␣I␣	␣I	163	97	am	a
145	129	␣wa	␣w	164	109	m.	m
146	97	as	a	165	46	.	.

Text: I␣wish␣I␣were␣what␣I␣was␣when␣I␣wished␣I␣were␣what␣I␣am.

Problem 2

Both X and Y are uniformly distributed with probability functions

$$p_X(k) = \tfrac{1}{5}, \quad k = 1, \dots, 5 \text{ and } p_Y(k) = \tfrac{1}{8}, \quad k = 1, \dots, 8$$

Then $p_{Z_a}(k)$ is described by the convolution of $p_X(k)$ and $p_Y(k)$. Since the uniform distribution is symmetric $-Y$ has the probability function

$$p_{-Y}(k) = \tfrac{1}{8}, \quad k = -8, \dots, -1$$

and $P_{Z_b}(k)$ is described by the convolution of $p_X(k)$ and $p_{-Y}(k)$. Finally $p_{Z_c}(k)$ is obtained by folding the negativ axis on to the positiv for $p_{Z_b}(k)$. These distributions are shown in Figure 2.1.

Alternatively, the distributions can be derived by first considering tables of the functions, as shown below, and then counting the number of occurances. This will of cource give the same probabilities as shown in Figure 2.1.

Z_a	1	2	3	4	5	6	7	8
1	2	3	4	5	6	7	8	9
2	3	4	5	6	7	8	9	10
3	4	5	6	7	8	9	10	11
4	5	6	7	8	9	10	11	12
5	6	7	8	9	10	11	12	13

Z_b	1	2	3	4	5	6	7	8
1	0	−1	−2	−3	−4	−5	−6	−7
2	1	0	−1	−2	−3	−4	−5	−6
3	2	1	0	−1	−2	−3	−4	−5
4	3	2	1	0	−1	−2	−3	−4
5	4	3	2	1	0	−1	−2	−3

Z_c	1	2	3	4	5	6	7	8
1	0	1	2	3	4	5	6	7
2	1	0	1	2	3	4	5	6
3	2	1	0	1	2	3	4	5
4	3	2	1	0	1	2	3	4
5	4	3	2	1	0	1	2	3

The entropies are

$$H(Z_a) = 3.4232 \text{ bit}$$
$$H(Z_b) = 3.4232 \text{ bit}$$
$$H(Z_c) = 2.7681 \text{ bit}$$

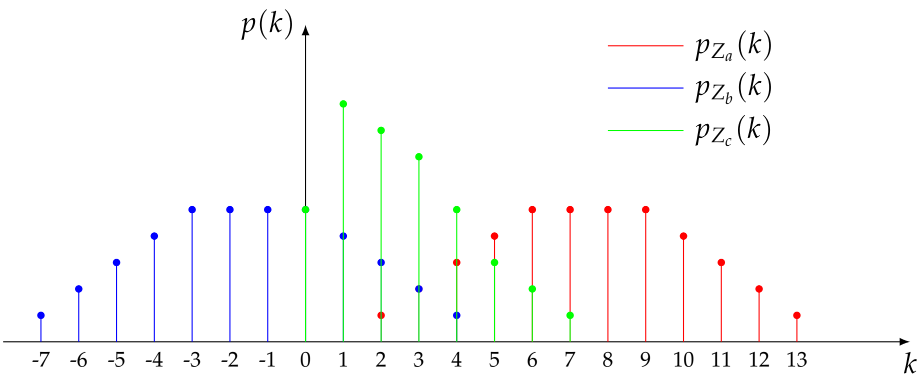


Figure 2.1: Probability functions for Z_a, Z_b and Z_c.

Problem 3

The state transition matrix is given by

$$P = \begin{pmatrix} 0 & 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \\ 1/5 & 0 & 1/5 & 1/5 & 1/5 & 1/5 \\ 1/5 & 1/5 & 0 & 1/5 & 1/5 & 1/5 \\ 1/5 & 1/5 & 1/5 & 0 & 1/5 & 1/5 \\ 1/5 & 1/5 & 1/5 & 1/5 & 0 & 1/5 \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 & 0 \end{pmatrix}$$

(a) Since $(\frac{1}{6}\ \frac{1}{6}\ \frac{1}{6}\ \frac{1}{6}\ \frac{1}{6}\ \frac{1}{6})P = (\frac{1}{6}\ \frac{1}{6}\ \frac{1}{6}\ \frac{1}{6}\ \frac{1}{6}\ \frac{1}{6})$ the steady state distribution is given by

$$\pi = (\tfrac{1}{6}\ \tfrac{1}{6}\ \tfrac{1}{6}\ \tfrac{1}{6}\ \tfrac{1}{6}\ \tfrac{1}{6})$$

That gives that without any knowledge about the previous symbols, the entropy is

$$H(\pi) = \log 6 = 1 + \log 3 = 2.58 \text{ bit}$$

(b) At the steady state solution $H(X_1|X_0 = j) = H(\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}) = \log 5$. Then,

$$H_\infty(X) = \sum_x \pi_x H(X_1|X_0 = x) = \log 5 = 2.32 \text{ bit}$$

(c) Starting at state $X_0 = j$ gives the same distribution, up to rearangments, for all j. So fixing $X_0 = 1$ gives the probability for the state at time n, i.e. X_n, conditioned on starting in $X_0 = 1$ is $P(X_n|X_0 = 1) = (1\,0\,0\,0\,0\,0)P^n$, and the corresponding entropy is $H\big((1\,0\,0\,0\,0\,0)P^n\big)$. Plotted as a function of n (using e.g. MATLAB) gives Figure 3.1. It is clearly seen that the distribution quickly approaches the asymptotic distribution (which equals the staeady state distribution).

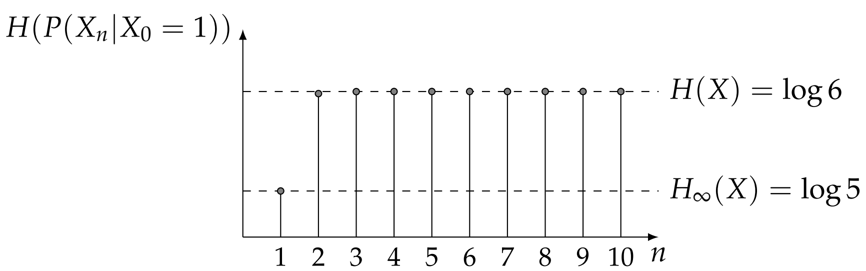
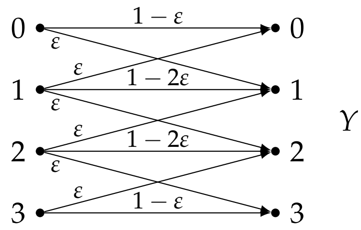


Figure 3.1: $H(P(X_n|X_0 = 1))$ as a function of n.

Problem 4

(a)

$P(Y\|X)$		$Y=0$	$Y=1$	$Y=2$	$Y=3$
X	0	$1-\varepsilon$	ε	0	0
	1	ε	$1-2\varepsilon$	ε	0
	2	0	ε	$1-2\varepsilon$	ε
	3	0	0	ε	$1-\varepsilon$



(b) With $p(x)=\frac{1}{4}$ the joint probability is $p(x,y)=\frac{1}{4}p(y|x)$ and, consequently,

$$p(y)=\sum_x p(x,y)=\begin{cases}\frac{1}{4}(1-\varepsilon)+\frac{1}{4}\varepsilon=\frac{1}{4}, & y=0,3\\ \frac{1}{4}(1-2\varepsilon)+2\frac{1}{4}\varepsilon=\frac{1}{4}, & y=1,2\end{cases}$$

That is, $H(Y)=\log 4=2$. Then,

$$\begin{aligned}H(Y|X)&=\sum_x H(Y|X{=}x)P(X{=}x)=2\tfrac{1}{4}H(1-\varepsilon),\varepsilon)+2\tfrac{1}{4}H(1-2\varepsilon,\varepsilon,\varepsilon)\\&=\tfrac{1}{2}h(\varepsilon)+\tfrac{1}{2}\Big(h(2\varepsilon)+2\varepsilon\underbrace{h(\tfrac{\varepsilon}{2\varepsilon})}_{=1}\Big)=\tfrac{1}{2}h(\varepsilon)+\tfrac{1}{2}h(2\varepsilon)+\varepsilon\end{aligned}$$

where in the third equality it is used that $H(\alpha,\beta,\gamma)=h(\alpha)+(1-\alpha)h(\frac{\beta}{1-\alpha})$, see Problem 3.14. The mutual information then becomes

$$I(X;Y)=H(Y)-H(Y|X)=2-\tfrac{1}{2}h(\varepsilon)-\tfrac{1}{2}h(2\varepsilon)-\varepsilon$$

In Figure 4.1 the mutual information is plotted as a function of ε.

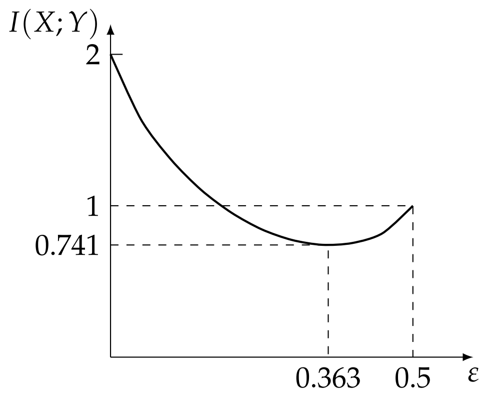


Figure 4.1: $I(X;Y)$ as a function of ε.

Problem 5

Th parameters used in the problem are $P=1000\,\text{mW}$ and $W_\Delta=1\,\text{MHz}$. Since $\frac{N_{0,i}}{|G_i|^2}$ is given in dBm/Hz it has to be converted to linear scale as

$$\frac{N_{0,i}}{|G_i|^2}=10^{\frac{[N_{0,i}/|G_i|^2]_{dB}}{10}}\quad[\text{mW/Hz}]$$

The specification that a PAM system should work at $P_e=10^{-6}$ means the SNR gap is set to $\Gamma=9\text{ dB}=7.94$.

(a) The estimated bit rate for sub-channel i is

$$R_{b,i}=W_\Delta\log\Big(1+\frac{|G_i|^2P_i}{W_\Delta N_{0,i}\Gamma}\Big)=W_\Delta\log\Big(1+\frac{P_i}{\frac{N_{0,i}}{|G_i|^2}W_\Delta\Gamma}\Big)$$

where P_i is the power used in channel i. Since the sub-channels are supposed to be independent the total bit rate is $R_b = \sum_i R_{b,i}$, which should be maximised when $\sum_i P_i = P$. Assign an optimisation function

$$J = \sum_{i=1}^{10} W_\Delta \log\Big(1 + \frac{P_i}{\frac{N_{0,i}}{|G_i|^2} W_\Delta \Gamma}\Big) + \lambda\Big(\sum_{i=1}^{10} P_i - P\Big)$$

By setting the derivative equal zero, $\frac{\partial}{\partial P_j} J = 0$, gives

$$P_j + \frac{N_{0,j}}{|G_j|^2} W_\Delta \Gamma = -\frac{W_\Delta}{\lambda \ln 2} = W_\Delta B \quad \Rightarrow \quad P_j = W_\Delta\Big(B - \frac{N_{0,j}}{|G_j|^2}\Gamma\Big)$$

With Khun-Tucker optimisation this gives

$$\begin{cases} P_j = W_\Delta\Big(B - \frac{N_{0,j}}{|G_j|^2}\Gamma\Big)^+ \\ \sum_j P_j = P \end{cases}$$

which is the water-filling procedure.

(b) As a first attempt, distribute the power on all sub-channels,

$$\sum_j P_j = \sum_j W_\Delta\Big(B - \frac{N_{0,j}}{|G_j|^2}\Gamma\Big) = W_\Delta\Big(10B - \Gamma\sum_j \frac{N_{0,j}}{|G_j|^2}\Big) = P$$

Rewriting gives an expression for B as

$$B = \frac{P}{10 W_\Delta} + \frac{\Gamma}{10}\sum_j \frac{N_{0,j}}{|G_j|^2} = 2.85 \cdot 10^{-4}\ \text{mW/Hz} = -35.45\ \text{dBm/Hz}$$

With this the first itaration P_j can be derived as

$$P_j = (245,\ -113,\ 159,\ 185,\ -216,\ 235,\ 34,\ 34,\ 235,\ 185) \quad \text{mW}$$

Here it is seen that sub-channels 2 and 5 have negative power and have to set to zero. This will affect the nioise levels on the remaining sub-channels, and the power distribution must be restarted. With eight sub-channels left the corresponding calculations give

$$B = \frac{P}{8 W_\Delta} + \frac{\Gamma}{8}\sum_{j \neq 2,5} \frac{N_{0,j}}{|G_j|^2} = 2.44 \cdot 10^{-4}\ \text{mW/Hz} = -36.13\ \text{dBm/Hz}$$

and the power per sub-channel

$$P_j = (204,\ 0,\ 118,\ 144,\ 0,\ 212,\ -7,\ -7,\ 194,\ 144) \quad \text{mW}$$

Again two sub-channels have negative powers, 7 and 8, and hve to be disconnected.
The third itaration gives the water filling level

$$B = \frac{P}{6 W_\Delta} + \frac{\Gamma}{6}\sum_{j \neq 2,5,7,8} \frac{N_{0,j}}{|G_j|^2} = 2.41 \cdot 10^{-4}\ \text{mW/Hz} = -36.18\ \text{dBm/Hz}$$

which gives

$$P_j = (201,\ 0,\ 115,\ 141,\ 0,\ 210,\ 0,\ 0,\ 191,\ 141) \quad \text{mW}$$

Since all used sub-channels have positive powers, we can continue to derive the corresponding bit rates,

$$R_{b,j} = (2599,\ 0,\ 938,\ 1270,\ 0,\ 2931,\ 0,\ 0,\ 2267,\ 1270) \quad \text{kbps}$$

and the total bit rate is $\sum_j R_{b,j} = 11.28$ Mbps This can be compared with the case when the power is distributed equally in all the sub-channels, $P_j = 100$ mW, which gives a total of 9.85 Mbps.

(c) With the coding gain $\gamma_c = 3$ dB the bit rate for the jth sub-chanel can be expressed as

$$R_{b,j} = W_\Delta \log\Big(1 + \frac{P_j \gamma_c}{\frac{N_{0,j}}{|G_j|^2} W_\Delta \Gamma}\Big) = W_\Delta \log\Big(1 + \frac{P_j}{\frac{N_{0,j}}{|G_j|^2} W_\Delta \Gamma_{eff}}\Big)$$

where $\Gamma_{eff} = \frac{\Gamma}{\gamma_c}$, or equivalently in dB scale $\Gamma_{eff} = \Gamma - \gamma_c = 6$ dB. That means the same optimisation as before can be used, but with $\Gamma = 6$ dB. In the first itration

$$B = 1.93 \cdot 10^{-4}\ \text{mW/Hz} = -37.15\ \text{dBm/Hz}$$
$$P_j = (173,\ -7,\ 130,\ 143,\ -59,\ 177,\ 67,\ 67,\ 168,\ 143)$$

which means that sub-channels 2 and 5 should not be used. Then, the second itaration gives

$$B = 1.85 \cdot 10^{-4}\ \text{mW/Hz} = -37.34\ \text{dBm/Hz}$$
$$P_j = (165,\ 0,\ 121,\ 134,\ 0,\ 169,\ 59,\ 59,\ 159,\ 134)$$

which gives

$$R_{b,j} = (3209,\ 0,\ 1548,\ 1880,\ 0,\ 3541,\ 551,\ 551,\ 2877,\ 1880) \quad [\text{kbps}]$$
$$R_b = 16.04\ \text{Mbps}$$

Problem 6

(a) Consider the difference between $H(Q)$ and $H(P)$,

$$\begin{aligned}
H(Q) - H(P) &= -q_1 \log q_1 - q_2 \log q_2 - \sum_{i=3}^{k} q_i \log q_i + \sum_{i=1}^{k} p_i \log p_i \\
&= -(p1 - \delta) \log q_1 - (p_2 + \delta) \log q_2 - \sum_{i=3}^{k} q_i \log q_i + \sum_{i=1}^{k} p_i \log p_i \\
&= \underbrace{\delta \log \frac{q_1}{q_2}}_{\geqslant 0 \text{ since } q_1 \geqslant q_2} - p_1 \log q_1 - p_2 \log q_2 - \sum_{i=3}^{k} q_i \log q_i + \sum_{i=1}^{k} p_i \log p_i \\
&\geqslant -\sum_{i=1}^{k} p_i \log q_i + \sum_{i=1}^{k} p_i \log p_i = \sum_{i=1}^{k} p_i \log \frac{p_i}{q_i} = D(p||q) \geqslant 0
\end{aligned}$$

(b) Use that $t \log t$ is a convex function to get

$$\begin{aligned}
H(Q) &= -\sum_i q_i \log q_i = -\sum_i \underbrace{\sum_j a_{ij} p_j \log \sum_j a_{ij} p_j}_{\leqslant \sum_j a_{ij} p_j \log p_j} \\
&\geqslant -\sum_i \sum_j a_{ij} p_j \log p_j = -\sum_j \underbrace{\sum_i a_{ij}}_{=1} p_j \log p_j = -\sum_j p_j \log p_j = H(P)
\end{aligned}$$

where the inequality follows from Jensen's inequality with a_{ij} as probabilities over j.

(c) The equation system in the hint can be written as

$$\begin{pmatrix} p_1 - \delta \\ p_2 + \delta \end{pmatrix} = \begin{pmatrix} p_1 & p_2 \\ p_2 & p_1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad \Rightarrow \quad \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} p_1 & p_2 \\ p_2 & p_1 \end{pmatrix}^{-1} \begin{pmatrix} p_1 - \delta \\ p_2 + \delta \end{pmatrix} = \begin{pmatrix} \frac{p_1 - p_2 - \delta}{p_1 - p_2} \\ \frac{\delta}{p_1 - p_2} \end{pmatrix}$$

This gives the A matrix as

$$A = \begin{pmatrix} \alpha & \beta & & & \\ \beta & \alpha & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}$$