Problem 1

The channel probabilities are given as $P(Y|X)$. With $P(X=0)=p$ and $P(X=1)=1-p$, the probabilities for Y are given by

$$P(Y=y)=\begin{cases}\alpha p+\alpha(1-p)=\alpha, & y=0\\ \beta p+\beta(1-p)=\beta, & y=\Delta\\ \gamma p+\gamma(1-p)=\gamma, & y=1\end{cases} \tag{1.1}$$

Hence,

$$\begin{aligned}H(Y)&=H(\alpha,\beta,\gamma)\\ H(Y|X)&=\sum_x H(Y|X=x)P(X=x)=H(\alpha,\beta,\gamma)p+H(\alpha,\beta,\gamma)(1-p)=H(\alpha,\beta,\gamma)\end{aligned}$$

and the mutual information is $I(X;Y)=H(Y)-H(Y|X)=0$. Since that imply $I(X;Y)=H(X)-H(X|Y)=0$ it means $H(X)=H(X|Y)$.

An alternativ solution is to observe that (1.1) means $P(Y=y)=P(Y=y|X=x)$, which gives

$$P(X=x)P(Y=y)=P(X=x)P(Y=y|X=x)=P(X=x,Y=y)$$

i.e. X and Y are independent and hence, $H(X)=H(X|Y)$.

Problem 2

Let X be the outcome of the die and Y the number of heads tossed with the coin. Then $P(X=n)=1/6$ for $n=1,\ldots,6$ and $P(Y=k|X=n)=\binom{n}{k}2^{-n}$, $k=0,\ldots,n$. Thus the joint probability is $P(X=n,Y=k)=\frac{\binom{n}{k}}{6\cdot 2^n}$, which is shown in the matrix

$$P=\begin{pmatrix}1/12 & 1/12 & 0 & 0 & 0 & 0 & 0\\ 1/24 & 1/12 & 1/24 & 0 & 0 & 0 & 0\\ 1/48 & 1/16 & 1/16 & 1/48 & 0 & 0 & 0\\ 1/96 & 1/24 & 1/16 & 1/24 & 1/96 & 0 & 0\\ 1/192 & 5/192 & 5/96 & 5/96 & 5/192 & 1/192 & 0\\ 1/384 & 1/64 & 5/128 & 5/96 & 5/128 & 1/64 & 1/384\end{pmatrix}$$

The probability for the number of heads can be derived as

$$P(Y=k)=\sum_n P(X=n,Y=k)=\begin{pmatrix}21/128 & 5/16 & 33/128 & 1/6 & 29/384 & 1/48 & 1/384\end{pmatrix}$$

for $k=0,\ldots,6$. With these the entropies can be derived as $H(X)=\log 6\approx 2.5850$, $H(Y)\approx 2.3074$ and $H(X,Y)\approx 4.3972$. Hence, the mutual information is $I(X;Y)=H(X)+H(Y)-H(XY)\approx 0.4952$.

Problem 3

The bandwidth for each channel is $W_\Delta=2$ MHz and the channel parameters expressed in linear scale becomes

$$\begin{aligned}|H_i|^2&=\begin{pmatrix}0.1 & 0.05 & 0.025 & 0.0126 & 0.0063\end{pmatrix}\\ N_0&=10^{-15}\text{ mW/Hz}\\ P&=10^{-6}\text{ mW}\end{aligned}$$

(a) To maximise the capacity the power is distributed according to the water-filling argument

$$P_i = \left(B - \frac{N_0 W_\Delta}{|H_i|^2}\right)^+ \quad \text{where} \quad \sum_i P_i = P$$

Summing the powers gives

$$\sum_i P_i = 5B - \sum_i \frac{N_0 W_\Delta}{|H_i|^2} = P$$

which gives

$$B = \frac{1}{5}\left(P + \frac{N_0 W_\Delta}{|H_i|^2}\right) = 3.23 \cdot 10^{-7}$$

and

$$P_i = \begin{pmatrix} 0.3031 & 0.2832 & 0.2435 & 0.1642 & 0.0061 \end{pmatrix} \cdot 10^{-6}$$

The capacity is given by

$$\begin{aligned} C &= \sum_i W_\Delta \log\left(1 + \frac{P_i |H_i|^2}{N_0 W_\Delta}\right) \\ &= 8.0276 + 6.0344 + 4.0413 + 2.0481 + 0.055 = 20.2 \text{ Mbps} \end{aligned}$$

(b) For an M-PAM communication scheme we can estimate the total bit rate by

$$R = \sum_i R_i = \sum_i W_\Delta \log\left(1 + \frac{P_i |H_i|^2}{\Gamma N_0 W_\Delta}\right) \quad \text{where} \quad \sum_i P_i = P$$

and $\Gamma = 9 \text{ dB} = 10^{0.9} = 7.94$. To maximise the bit rate the following maximisation function is used

$$J = \sum_i W_\Delta \log\left(1 + \frac{P_i |H_i|^2}{\Gamma N_0 W_\Delta}\right) + \lambda\Big(\sum_i P_i - P\Big)$$

By setting the derivative equal to zero, the water-filling function can be derived as

$$P_i = \left(B - \frac{\Gamma N_0 W_\Delta}{|H_i|^2}\right)^+$$

After succesiv cancelations of sub-channels with negative power, i.e. sub-channel 4 and 5, the optimal power distribution is given by

$$P_i = \begin{pmatrix} 0.5439 & 0.3858 & 0.0703 & 0 & 0 \end{pmatrix}$$

and the corresponding bit rate

$$R = \sum_i W_\Delta \log\left(1 + \frac{P_i |H_i|^2}{\Gamma N_0 W_\Delta}\right) = 4.2905 + 2.2973 + 0.3042 = 6.89 \text{ Mbps}$$

Problem 4

The transition probability matrix is

$$P = \begin{pmatrix} 1/3 & 2/3 & 0 \\ 1/3 & 1/3 & 1/3 \\ 0 & 2/3 & 1/3 \end{pmatrix}$$

(a) Through the equation system $\boldsymbol{\pi} P = \boldsymbol{\pi}$ and $\sum_i \pi_i = 1$ the stationary distribution is given by $\boldsymbol{\pi} = \begin{pmatrix} 1/4 & 1/2 & 1/4 \end{pmatrix}$. The entropy rate is given by

$$H_\infty(X) = \pi_1 h\big(\tfrac{1}{3}\big) + \pi_2 \log(3) + \pi_3 h\big(\tfrac{1}{3}\big) \approx 1.2516 \text{ bit}$$

(b) The sequence $\boldsymbol{Y}$ is given by taking two steps at a time in the graph. That corresponds to the transition probability matrix

$$P_b = P^2 = \begin{pmatrix} 1/3 & 4/9 & 2/9 \\ 2/9 & 5/9 & 2/9 \\ 2/9 & 4/9 & 1/3 \end{pmatrix}$$

Deriving the stationary distribution, by observing that $\boldsymbol{\pi} P^2 = \boldsymbol{\pi} P$, will give the same equation as in a), i.e. $\boldsymbol{\pi} = \begin{pmatrix} 1/4 & 1/2 & 1/4 \end{pmatrix}$. The entropy rate is given by

$$H_\infty(X) = \pi_1 H\big(\tfrac{1}{3}, \tfrac{4}{9}, \tfrac{2}{9}\big) + \pi_2 H\big(\tfrac{2}{9}, \tfrac{5}{9}, \tfrac{2}{9}\big) + \pi_3 H\big(\tfrac{2}{9}, \tfrac{4}{9}, \tfrac{1}{3}\big) \approx 1.4830 \text{ bit}$$

Problem 5

(a) To form the new code replace x_k and x_l be replaced by their parent node z_m with probability $q_m = p_k + p_l$. The new code tree has the leaf probabilities $\tilde{p}_j,\ j = 1, \dots n-1$, which are equal to $p_i,\ i \neq k, l$ and q_m. Then

$$\begin{aligned}
H(X) &= -\sum_{i=1}^{n} p_i \log p_i = -\sum_{i \neq k,l} p_i \log p_i - p_k \log p_k - p_l \log p_l \\
&= -\sum_{j=1}^{n-1} \tilde{p}_j \log \tilde{p}_j + q_m \log q_m - p_k \log p_k - p_l \log p_l \\
&= H(\tilde{X}) + (p_k + p_l)\Big(\frac{p_k + p_l}{p_k + p_l} \log(p_k + p_l) - \frac{p_k}{p_k + p_l} \log p_k - \frac{p_l}{p_k + p_l} \log p_l\Big) \\
&= H(\tilde{X}) + (p_k + p_l)\Big(-\frac{p_k}{p_k + p_l} \log \frac{p_k}{p_k + p_l} - \frac{p_l}{p_k + p_l} \log \frac{p_l}{p_k + p_l}\Big) \\
&= H(\tilde{X}) + (p_k + p_l) h\Big(\frac{p_k}{p_k + p_l}\Big)
\end{aligned}$$

and, hence $\alpha = \frac{p_k}{p_k + p_l}$.

(b) From the structure of the formula in a), assume that $H(X) = \sum_{i=1}^{n-1} q_i h(\alpha_i)$, where α_i is the probability for choosing branch 0 from the inner node z_i. That is, If z_k and z_l are the children nodes of z_i, then $\alpha_i = \frac{q_k}{q_k + q_l}$.

- For $n = 2$ there is one inner node, the root with $q_1 = 1$ and two leaves with probabilities p_1 and p_2. Then $\alpha_1 = p_k$ and, hence, $H(X) = h(p_1) = q_1 h(\alpha_1)$. That means our assumption is true for this case.

- For $n > 2$, assume the formula is true for a tree with $n-1$ leaves. Then construct a code according to a),

$$\sum_{i=1}^{n-1} q_i h(\alpha_i) = \sum_{i\neq m} q_i h(\alpha_i) + q_m h(\alpha_m) = H(\tilde{X})(p_k + p_l)h\Big(\frac{p_k}{p_k+p_l}\Big) = H(X)$$

(c) From the path length lemma and the fact that the entropy lower bounds the codeword length, we get

$$L = \sum_{i=1}^{n-1} q_i \geqslant H(X) = \sum_{i=1}^{n-1} q_i h(\alpha_i)$$

with equality if and only if $h(\alpha_i) = 1$ for all i. This is equivalent to saying that $\alpha_i = \frac{1}{2}$, for all i. Since the leaf node x_i uniquely determines the path from the root to the leaf, $z_{i_0}, z_{i_1}, \dots, z_{i_{\ell-1}}$, where z_{i_0} is the root node, we get

$$\begin{aligned}
p(x_i) &= P(z_{i_0}, \dots, z_{i_{\ell-1}}, x_i) \\
&= P(z_{i_0}) \prod_{j=1}^{\ell-1} P(z_{i_j} | z_{i_0}, \dots, z_{i_{j-1}}) P(x_i | z_{i_0}, \dots, z_{i_{\ell-1}}) \\
&= P(z_{i_0}) \prod_{j=1}^{\ell-1} P(z_{i_j} | z_{i_{j-1}}) P(x_i | z_{i_{\ell-1}}) \\
&= 1 \prod_{j=1}^{\ell} \alpha_{i_j} = 1 \prod_{j=1}^{\ell} \frac{1}{2} = \frac{1}{2^\ell}
\end{aligned}$$

That is, to get $L = H(X)$ it is required that the leaf probabilities are powers of 2. (This fact was also seen from the optimal codeword length, where $\log p_i$ must be an integer to fulfil the bound.)