Exam (January 14, 2017) Solution

Last modified January 25, 2017

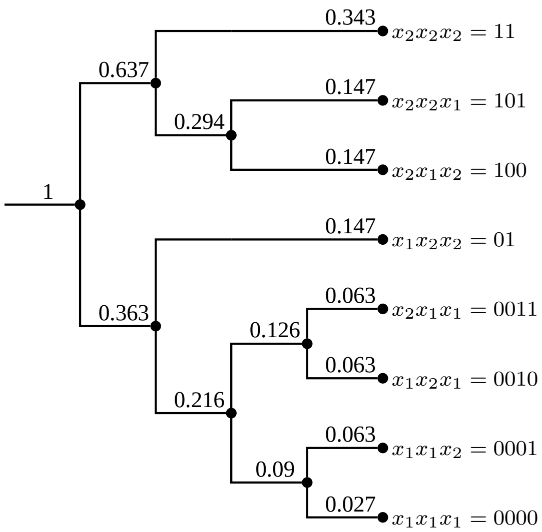
Problem 1 - Source Coding and Channel Capacity [15 points]

Part I

1. A possible Huffman code of the source, encoding blocks of three symbols, is as follows,

symbol	probability	codeword
$x_2x_2x_2$	0.343	11
$x_1x_2x_2$	0.147	01
$x_2x_1x_2$	0.147	100
$x_2x_2x_1$	0.147	101
$x_1x_1x_2$	0.063	0001
$x_1x_2x_1$	0.063	0010
$x_2x_1x_1$	0.063	0011
$x_1x_1x_1$	0.027	0000

where the corresponding tree is depicted below.



2. The source entropy is

$$\begin{aligned} H(X) &= -0.3\log_2(0.3) - 0.7\log_2(0.7) \\ &\approx 0.8813. \end{aligned}$$

The average codeword length is

$$\begin{aligned} \bar{L}_3 &= 2\cdot 0.343 + 3\cdot 0.147 + 3\cdot 0.147 + 2\cdot 0.147 + 4\cdot 0.063 + 4\cdot 0.063 + 4\cdot 0.063 + 4\cdot 0.027 \\ &= 2.726. \end{aligned}$$

Therefore, the efficiency is

$$\eta = \frac{3H(X)}{\bar{L}_3} \approx 96.99\%.$$

3. The codeword lengths are $\boldsymbol{l} = (2, 2, 3, 3, 4, 4, 4, 4)$. We get

$$\sum_{i=1}^{8} 2^{-l_i} = 1.$$

Therefore, Kraft's inequality is satisfied.

Part II

1. The joint distribution of X and Y is given by

$$\begin{aligned}
P_{X,Y}(x_1, y_1) &= \frac{1}{2}, \\
P_{X,Y}(x_1, y_2) &= 0, \\
P_{X,Y}(x_2, y_1) &= \frac{1}{2}\varepsilon, \\
P_{X,Y}(x_2, y_2) &= \frac{1}{2}(1-\varepsilon).
\end{aligned}$$

The distribution of Y can be obtained from $P_{X,Y}(x, y)$ through marginalization, which gives

$$\begin{aligned}
P_Y(y_1) &= \frac{1}{2}(1+\varepsilon), \\
P_Y(y_2) &= \frac{1}{2}(1-\varepsilon).
\end{aligned}$$

2. By definition, we get

$$\begin{aligned}
H(Y) &= -\sum_{y \in \mathcal{Y}} P_Y(y) \log_2(P_Y(y)) \\
&= -\frac{1}{2}(1+\varepsilon) \log_2\left(\frac{1}{2}(1+\varepsilon)\right) - \frac{1}{2}(1-\varepsilon) \log_2\left(\frac{1}{2}(1-\varepsilon)\right) \\
&= H_\mathrm{b}\left(\frac{1+\varepsilon}{2}\right), \\
H(Y|X) &= -\sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} P_{X,Y}(x, y) \log_2(P_{Y|X}(y|x)) \\
&= -\frac{1}{2}\log_2(1) - \frac{1}{2}\varepsilon \log_2(\varepsilon) - \frac{1}{2}(1-\varepsilon)\log_2(1-\varepsilon) \\
&= \frac{1}{2} H_\mathrm{b}(\varepsilon),
\end{aligned}$$

where

$$H_\mathrm{b}(\alpha) = -\alpha \log_2(\alpha) - (1-\alpha)\log_2(1-\alpha)$$

is the binary entropy function.

The mutual information between X and Y is given by

$$\begin{aligned}
I(Y; X) &= H(Y) - H(Y|X) \\
&= \begin{cases} H_\mathrm{b}(1/2) - H_\mathrm{b}(0)/2 = 1 \text{ bit}, & \varepsilon = 0 \\ H_\mathrm{b}(3/4) - H_\mathrm{b}(1/2)/2 \approx 0.3113 \text{ bits}, & \varepsilon = 1/2 \\ H_\mathrm{b}(1) - H_\mathrm{b}(1)/2 = 0 \text{ bits}, & \varepsilon = 1. \end{cases}
\end{aligned}$$

3. When $\varepsilon = 1$, $I(Y; X) = 0$, and thus X and Y are independent. For $\varepsilon = 0$ and $\varepsilon = 1/2$, X and Y are not independent since $I(Y; X) > 0$.

4. The capacity is not a function of the input distribution, and therefore not dependent on the probabilities of the input symbols, since it is obtained by maximizing $I(X; Y)$ over all possible input distributions.

Problem 2 - Signal Constellations and Detection [15 points]

1. The minimum distance is either the distance between two neighbors on the inner circle, or the distance between points on the inner and the outer circle. For symmetry reasons it suffices to look at s_1 and its neighbors. The minimum distance of Ω is thus either $|s_1 - s_7|$ or $|s_1 - s_2|$. Clearly, $|s_1 - s_7| = r_2 - r_1$. Furthermore,

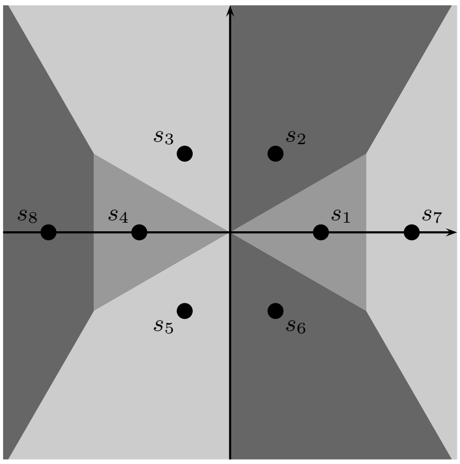
$$\begin{aligned}
|s_1 - s_2| &= |r_1 - r_1 e^{j2\pi/6}| \\
&= r_1 |1 - e^{j\pi/3}| \\
&= r_1 \sqrt{(1 - e^{j\pi/3})(1 - e^{j\pi/3})^*} \\
&= r_1 \sqrt{1 - 2\Re\{e^{j\pi/3}\} + 1} \\
&= r_1 \sqrt{2 - 2\cos(\pi/3)} \\
&= r_1.
\end{aligned}$$

   Therefore, the minimum distance of Ω is $\min(r_1, r_2 - r_1)$, which is maximized when

$$r_1 = r_2 - r_1 \quad \Leftrightarrow \quad r_1 = \frac{1}{2} r_2.$$

2. 6 points lie on a circle with radius r_1, and 2 points lie on a circle with radius $r_2 = 2r_1$. The average comes out to $E_s = 7r_1^2/4 = 1$, and therefore $r_1 = \sqrt{4/7}$ and $r_2 = 2\sqrt{4/7}$.

3. The maximum likelihood decision regions for Ω are depicted below.



4. In general, the nearest neighbor approximation is given by

$$P_s(e) \approx \bar{A}_{\min} Q\left(\sqrt{\frac{d_{E,\min}^2}{2N_0}}\right),$$

   where $\bar{A}_{\min}$ is the average number of nearest neighbors at $d_{E,\min}$.

   For Ω_1, we have $d_{E,\min} = r_1 = \sqrt{4/7}$ and $\bar{A}_{\min} = 2$. Therefore

$$P_s(e) \approx 2Q\left(\sqrt{\frac{2}{7N_0}}\right).$$

5. Both constellations have 8 points, therefore they have the same spectral efficiency. Moreover, the minimum distance of 8-PSK is $2\sin(\pi/8) = \sqrt{2-\sqrt{2}} \approx 0.7654$, which is larger than the minimum distance of Ω. Thus, 8-PSK will have lower $P_s(e)$ than Ω at high SNR.

6. The power efficiencies of Ω and Ω_{R} will only differ if the rotation changes the distances between constellation points. Since each point in Ω is rotated by the same phase ϕ, the Euclidean distance between any two points in Ω_{R} becomes

$$\begin{aligned}
|s_i e^{j\phi} - s_j e^{j\phi}| &= |e^{j\phi}(s_i - s_j)| \\
&= |e^{j\phi} z| \\
&= |e^{j\phi}| z| e^{j\angle z}| \\
&= |z||e^{j(\phi+\angle z)}| \\
&= |z| \\
&= |s_i - s_j|
\end{aligned}$$

for $i = j = 1, 2, \ldots, 8$, where $\angle z$ denotes the angle of a complex number z. Since the rotation does not affect the distances between the constellation points, Ω and Ω_{R} have the same power efficiency.

7. The minimum angular difference between two points in Ω is $\pi/3$. Hence, if $\theta > \pi/6$ then the points are rotated over to the decision regions of their neighboring points. Moreover, if $\theta = \pi/6$ then points are rotated on the decision region boundary between two neighboring points. Therefore, the symbol error probability for Ω goes to zero as the SNR goes to infinity as long as $\theta < \pi/6$.

8. The maximum likelihood decision rule can be simplified as

$$\begin{aligned}
\hat{s}_{\mathrm{ML}} &= \operatorname*{argmax}_{s\in\mathcal{X}} p(r|s) \\
&= \operatorname*{argmax}_{s\in\mathcal{X}} \log p(r|s) \\
&= \operatorname*{argmax}_{s\in\mathcal{X}} -|r-s|^2 \\
&= \operatorname*{argmax}_{s\in\mathcal{X}} \left(-|r|^2 + 2\Re\{rs^*\} - |s|^2\right) \\
&= \operatorname*{argmax}_{s\in\mathcal{X}} \left(-|r|^2 + 2\Re\{rs\} - |s|^2\right) \\
&= \operatorname*{argmax}_{s\in\mathcal{X}} \left(2\Re\{rs\} - |s|^2\right). \qquad (1)
\end{aligned}$$

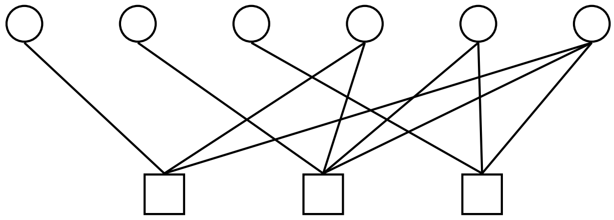
Since PAM constellations with more than 2 points do not have constant amplitude, the latter term in (1) can not be neglected in the maximization. Therefore, the maximum likelihood decision rule for s can *not* be written as

$$\hat{s}_{\mathrm{ML}} = \operatorname*{argmax}_{s\in\mathcal{X}} \Re\{rs\}.$$

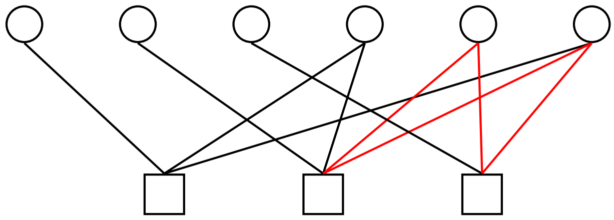
Problem 3 - Linear Block Codes and LDPC Codes [15 points]

Part I

1. The Tanner graph for $\mathbf{H}$ is depicted below. We can identify one cycle with length 4. Since this is the



shortest cycle, we know that the girth of $\mathbf{H}$ has to be 4. Below, this cycle is highlighted in the graph.



2. By considering the Tanner graph, we get

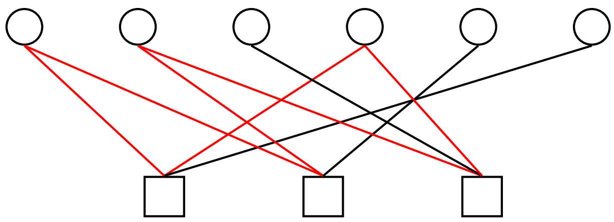
$$\Lambda(x) = \frac{1}{2}x + \frac{1}{3}x^2 + \frac{1}{6}x^3 \qquad\qquad P(x) = \frac{2}{3}x^3 + \frac{1}{3}x^4.$$

The given code is an irregular LDPC code since the VNs are of different degrees.

3. Yes, it can be done by linear operations on the matrix $\mathbf{H}$. By adding the first and the second row, and by adding the second and third row, we get

$$\tilde{\mathbf{H}} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \end{bmatrix}.$$

From the new Tanner graph, we can find that the shortest cycle (i.e., the girth) is of length 6. Therefore $\tilde{\mathbf{H}}$ has a larger girth than $\boldsymbol{H}$. The new Tanner graph with the shortest cycle highlighted is shown below.



Part II

1.

$$\mathbf{G}_s = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \qquad \mathbf{H}_s = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$$

2. From the size of the generator matrix, we can deduce $N = 7$, $K = 4$. Since the code is a linear code, we can conclude, that $d_{\min} = w_{\mathrm{H,min}} = 2$.
The code rate is $R = \frac{4}{7}$. The code can detect $w_{\mathrm{H}}(\boldsymbol{e}) < d_{\min} = 2$ errors and can correct all errors patterns with weight $w_{\mathrm{H}}(\boldsymbol{e}) \leq \lfloor \frac{d_{\min}-1}{2} \rfloor = 0$.

3. The syndrom table contains $2^3 - 1 = 7$ entries. We first compute the syndroms of all errors of weight 1. These are the columns of the parity check matrix, since for an $\mathbf{e}$ with exactly a single one, $\mathbf{H}_s\mathbf{e}^T$ is a column on $\mathbf{H}_s$. The error patterns for the remaining two sequences can be found by considering the equation

$$\mathbf{s} = \mathbf{H}_s\mathbf{e}^T$$

so that

$$s_1 = e_1 + e_5 \qquad s_2 = e_2 + e_6 \qquad s_3 = e_1 + e_2 + e_3 + e_4 + e_7.$$

Note that the syndrome table in Tab. 1 is not unique since multiple error patterns with the same weight map to the same syndrom.

syndrome	error pattern
101	1000000
011	0100000
001	0010000
100	0000100
010	0000010
110	1100000
111	0100100

Table 1: Syndrome table

4. For hard decision decoding, the AWGN channel reduces to a BSC. We first decide on a received bit according to the received soft information, i.e.,

$$\overline{r_i} = \begin{cases} 0 & r_i < 0 \\ 1 & r_i \geq 0 \end{cases}.$$

We therefore get

$$\bar{\mathbf{r}} = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

Next we compute the syndrom $\mathbf{s} = \mathbf{H}_s\bar{\mathbf{r}}^T = [101]$, which, from Tab. 1, gives an error pattern $\mathbf{e} = [1000000]$. The ML codeword then results in

$$\begin{aligned}\hat{\mathbf{x}} &= \bar{\mathbf{r}} + \mathbf{e} = [0101111] + [1000000] \\ &= [1101111]\,.\end{aligned}$$

Results for $\mathbf{G}_s^\star$ and $\mathbf{H}_s^\star$ (only final results):

2.
- $N = 7$
- $K = 4$
- $d_{\min} = 2$
- $R = \frac{4}{7}$
- Detect error patterns of weight $w_{\mathrm{H}}(\boldsymbol{e}) < 2$.
- Correct all errors patterns with weight $w_{\mathrm{H}}(\boldsymbol{e}) \leq 0$.

3.

syndrome	error pattern
001	1000000
111	0100000
011	0010000
101	0001000
100	0000100
010	0000010
110	0011000

Table 2: Syndrome table

4.

$$\bar{\mathbf{r}} = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 & 1 & 1 \end{bmatrix} \qquad \mathbf{s} = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix}$$

$$\mathbf{e} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \qquad \hat{\mathbf{x}} = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

Problem 4 - Viterbi Algorithm [15 points]

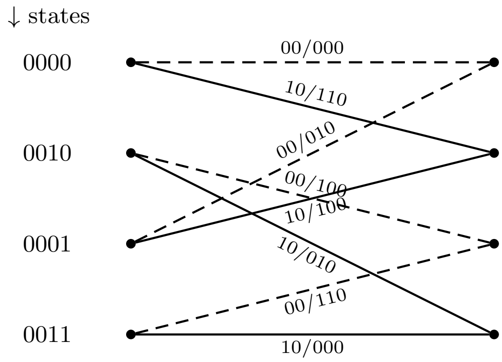
1.

$$\mathbf{G} = \begin{bmatrix} 1+D & 1+D^2 & 0 \\ 0 & 1+D+D^2 & D+D^2 \end{bmatrix}$$

2. Dividing the first row by $(1+D)$ and the second row by $(D+D^2)$ and then swapping column two and three results in

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & \frac{1+D^2}{1+D} \\ 0 & 1 & \frac{1+D+D^2}{D+D^2} \end{bmatrix}.$$

3. The lower two memory elements are always zero, therefore only four states are reachable.



4. From the Viterbi algorithm below, we can see that the received information sequence results to $\hat{\boldsymbol{u}} = (00100000)$.