Exam (January 19, 2019) Solution

Last modified January 22, 2019

Problem 1 - Channel Capacity [15 points]

Part I

1. Using the fact that $P_{X,Y}(x,y) = P_{Y|X}(y|x)P_X(x)$ and $\alpha = \beta = 0.5$, the joint distribution is given by

	$X=0$	$X=1$
$Y=0$	0	1/4
$Y=\varepsilon$	1/4	1/4
$Y=1$	1/4	0

Moreover, $P_Y(y) = \sum_{x\in\mathcal{X}} P_{X,Y}(x,y)$ so $P_Y(0) = 1/4$, $P_Y(\varepsilon) = 1/2$, and $P_Y(1) = 1/4$.

2. $\mathsf{H}(X) = 1$ bit, $\mathsf{H}(Y) = 3/2$ bits, $\mathsf{H}(Y|X) = 1$ bit, and $\mathsf{H}(X,Y) = \mathsf{H}(Y|X) + \mathsf{H}(X) = 2$ bits.

3. $\mathsf{I}(X;Y) = \mathsf{I}(Y;X) = \mathsf{H}(Y) - \mathsf{H}(Y|X) = 1/2$ bit. The channel capacity is the mutual information, maximized with respect to the input distribution, i.e., $C = \max_{P_X} \mathsf{I}(X;Y)$.

4. When the input distribution is fixed at $P_X(0) = P_Y(0) = 1/2$, the mutual information can be upper bounded as $\mathsf{I}(X;Y) = \mathsf{H}(X) - \mathsf{H}(X|Y) \leq \mathsf{H}(X)$. Letting $\alpha = 0$ as well as $\beta = 0$ or $\beta = 1$ yields a noise-free channel, in which case $\mathsf{H}(Y|X) = \mathsf{H}(X|Y) = 0$. This gives $\mathsf{I}(X;Y) = \mathsf{H}(X) - \mathsf{H}(X|Y) = \mathsf{H}(X) = 1$ bit.

Part II

1. A complex Gaussian distribution with zero mean and variance $\mathsf{E_s}$.

2. $\mathsf{E_s}/\mathsf{N_0} = 1000$, and therefore $C = \log_2(1+1000) \approx 9.97$ bits/symbol. Reliable communication is therefore not possible at $R = 10$ bits/symbols. The lowest $\mathsf{E_s}/\mathsf{N_0}$ required for this rate can be found as

$$C = \log_2\left(1 + \frac{\mathsf{E_s}}{\mathsf{N_0}}\right) = 10 \quad \Rightarrow \quad \frac{\mathsf{E_s}}{\mathsf{N_0}} = 2^{10} - 1 = 1023 \approx 30.1 \text{ dB}.$$

3. The minimum required $\mathsf{E_b}/\mathsf{N_0}$ required for reliable transmission at an information rate R is (see lecture notes for details)

$$\frac{\mathsf{E_b}}{\mathsf{N_0}} > \frac{2^R - 1}{R},$$

and

$$\lim_{R\to 0} \frac{2^R - 1}{R} = \ln 2.$$

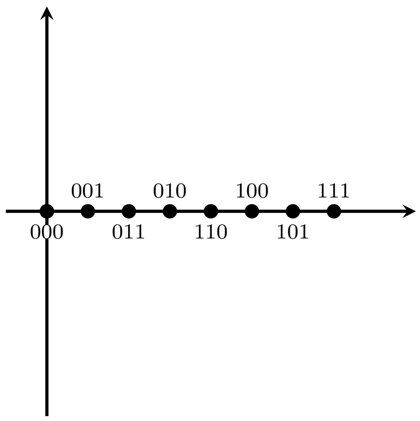
Therefore, $R \to 0$ at the minimum required $\mathsf{E_b}/\mathsf{N_0}$.

Problem 2 - Signal Constellations and Maximum Likelihood [15 points]

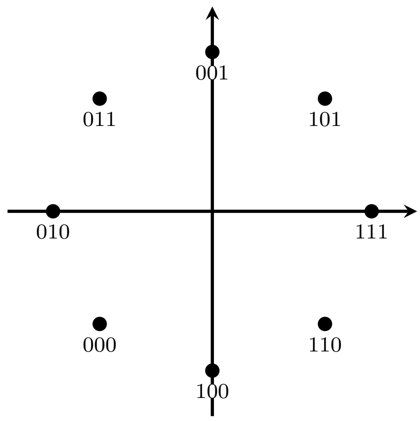
1. For Ω_1, the average energy per symbol is $\mathsf{E_s} = (a^2 + 4a^2 + 9a^2 + 16a^2 + 25a^2 + 36a^2 + 49a^2)/8 = 35a^2/2$, and therefore $a = \sqrt{2\mathsf{E_s}/35}$.

   For Ω_2, the average energy per symbol is $\mathsf{E_s} = (8b^2)/8 = b^2$, and therefore $b = \sqrt{\mathsf{E_s}}$.

2. An example of a Gray mapping for each constellation is depicted below.



(a) Ω_1



(b) Ω_2

Figure 1: Gray mapping examples for the constellations.

3. The symbol error probability can be upper bounded using the union bound, i.e., as the sum of pairwise error probabilities (see lecture notes for details). Hence,

$$
\begin{aligned}
P_{\mathsf{s}} &\leq \frac{1}{8} \cdot 8 \Bigg[2Q\left(\sqrt{\frac{4b^2 \sin^2(\pi/8)}{2\mathsf{N_0}}}\right) + 2Q\left(\sqrt{\frac{4b^2 \sin^2(\pi/4)}{2\mathsf{N_0}}}\right) \\
&\qquad + 2Q\left(\sqrt{\frac{4b^2 \sin^2(3\pi/8)}{2\mathsf{N_0}}}\right) + Q\left(\sqrt{\frac{4b^2 \sin^2(\pi/2)}{2\mathsf{N_0}}}\right)\Bigg] \\
&= 2Q\left(\sqrt{\frac{2\mathsf{E_s} \sin^2(\pi/8)}{\mathsf{N_0}}}\right) + 2Q\left(\sqrt{\frac{2\mathsf{E_s} \sin^2(\pi/4)}{\mathsf{N_0}}}\right) + 2Q\left(\sqrt{\frac{2\mathsf{E_s} \sin^2(3\pi/8)}{\mathsf{N_0}}}\right) + Q\left(\sqrt{\frac{2\mathsf{E_s} \sin^2(\pi/2)}{\mathsf{N_0}}}\right).
\end{aligned}
$$

   Note that for PSK formats, we can discard all terms except the one corresponding to the nearest neighbors and still maintain an upper bound. Hence, a tighter upper bound on P_{s} is

$$
P_{\mathsf{s}} \leq 2Q\left(\sqrt{\frac{2\mathsf{E_s} \sin^2(\pi/8)}{\mathsf{N_0}}}\right),
$$

   which also happens to be the nearest neighbor approximation of P_{s}. Moreover, since $Q(\,\cdot\,)$ decreases with increasing arguments, $\lim_{\mathsf{N_0} \to 0} P_{\mathsf{s}} = 0$.

4. The average symbol energy of Ω_1 is given by $\mathbb{E}[|S|^2]$, i.e., the second moment of s. In general, $\mathbb{E}[|S|^2] = \mathrm{Var}(S) + |\mathbb{E}[S]|^2$, where $\mathrm{Var}(S)$ is the variance of s and $|\mathbb{E}[S]|^2$ is the expectation of s, i.e., the mean value of the constellation. A shift applied to the constellation points does not affect $\mathrm{Var}(S)$ but it does change $|\mathbb{E}[S]|^2$. Therefore, to minimize the average symbol energy, the constellation points must be shifted by $-\mathbb{E}[S]$, which yields $\mathbb{E}[|S|^2] = \mathrm{Var}(S) + |\mathbb{E}[S]|^2 = \mathrm{Var}(S) + 0 = \mathrm{Var}(S)$.

5. The maximum-likelihood estimator of θ is obtained as

$$\begin{aligned}\hat{\theta}_{\mathrm{ML}} &= \underset{\theta\in[0,2\pi)}{\operatorname{argmax}}\, p(r|s,\theta)\\ &= \underset{\theta\in[0,2\pi)}{\operatorname{argmin}}\, |r|^2 - 2\Re\{rs^* e^{-j\theta}\} + |s|^2\\ &= \underset{\theta\in[0,2\pi)}{\operatorname{argmax}}\, \Re\{rs^* e^{-j\theta}\}\\ &= \underset{\theta\in[0,2\pi)}{\operatorname{argmax}}\, |r||s|\Re\{e^{j\angle(rs^*)} e^{-j\theta}\}\\ &= \underset{\theta\in[0,2\pi)}{\operatorname{argmax}}\, \cos(\angle(rs^*) - \theta).\end{aligned}$$

The cosine is maximized when its argument is zero, and therefore, $\hat{\theta}_{\mathrm{ML}} = \angle(rs^*)$.

6. The first strategy is maximum-likelihood detection assuming an AWGN channel. Hence, when θ is uniformly distributed in $[0, 2\pi)$, P_{s} will not tend to zero when $N_0 \to 0$ because the received signals will lie uniformly on circles of different radii.

   The second strategy exploits the fact that information is only encoded in the amplitude of the signal, and finds the constellation point that is closest (in terms of Euclidean distance) to the amplitude of the received signal. When $\mathsf{N}_0 \to 0$ there will be no noise in the amplitude and $P_{\mathsf{s}} \to 0$. Therefore, the second strategy will perform better when $\mathsf{N}_0 \to 0$.

7. Denote the set of constellation points for 8PAM with $\{-7c, -5c, -3c, -c, +c, +3c, +5c, +7c\}$. Then, the average symbol energy is

$$\mathsf{E}_{\mathsf{s}} = 3\mathsf{E}_{\mathsf{b}} = \frac{1}{8}\left(2c^2 + 2\cdot 9c^2 + 2\cdot 25c^2 + 2\cdot 49c^2\right) = 21c^2,$$

   and hence, $c = \sqrt{3\mathsf{E}_{\mathsf{b}}/21} = \sqrt{\mathsf{E}_{\mathsf{b}}/7}$. If an outermost constellation points is transmitted, the error probability is $Q(\sqrt{d^2/(2\mathsf{N}_0)})$, where $d = 2c$. Otherwise, the error probability is $2Q(\sqrt{d^2/(2\mathsf{N}_0)})$. Therefore, the average symbol error probability is

$$P_{\mathsf{s}}^{\mathsf{8PAM}} = \frac{1}{8}\left(2\cdot Q\left(\sqrt{\frac{d^2}{2\mathsf{N}_0}}\right) + 6\cdot 2Q\left(\sqrt{\frac{d^2}{2\mathsf{N}_0}}\right)\right) = \frac{7}{4}Q\left(\sqrt{\frac{2c^2}{\mathsf{N}_0}}\right) = \frac{7}{4}Q\left(\sqrt{\frac{2\mathsf{E}_{\mathsf{b}}}{7\mathsf{N}_0}}\right) = AQ\left(\sqrt{B\frac{\mathsf{E}_{\mathsf{b}}}{\mathsf{N}_0}}\right),$$

   i.e., $A = 7/4$ and $B = 2/7$.

   In order to determine the exact $P_{\mathsf{s}}^{64\mathsf{QAM}}$, 64QAM can be regarded as two independent 8PAM in the in-phase and quadrature components of the signal. Since the complex AWGN affects each component of the signal independently, the exact symbol error probability of 64QAM can be derived as follows. Let e_{I} and e_{Q} denote the events “symbol error occurred in in-phase component" and “symbol error occurred in quadrature component", respectively, and note that $\mathsf{e}_{\mathsf{I}} = \mathsf{e}_{\mathsf{Q}} = P_{\mathsf{s}}^{\mathsf{8PAM}}$. Then,

$$\begin{aligned}P_{\mathsf{s}}^{64\mathsf{QAM}} &= \Pr(\mathsf{e}_{\mathsf{I}} \cup \mathsf{e}_{\mathsf{Q}})\\ &= \Pr(\mathsf{e}_{\mathsf{I}}) + \Pr(\mathsf{e}_{\mathsf{Q}}) - \Pr(\mathsf{e}_{\mathsf{I}} \cap \mathsf{e}_{\mathsf{Q}})\\ &= \Pr(\mathsf{e}_{\mathsf{I}}) + \Pr(\mathsf{e}_{\mathsf{Q}}) - \Pr(\mathsf{e}_{\mathsf{I}})\Pr(\mathsf{e}_{\mathsf{Q}})\\ &= 2P_{\mathsf{s}}^{\mathsf{8PAM}} - \left(P_{\mathsf{s}}^{\mathsf{8PAM}}\right)^2\\ &= 2AQ\left(\sqrt{B\frac{\mathsf{E}_{\mathsf{b}}}{\mathsf{N}_0}}\right) - A^2Q^2\left(\sqrt{B\frac{\mathsf{E}_{\mathsf{b}}}{\mathsf{N}_0}}\right)\\ &= \frac{7}{2}Q\left(\sqrt{\frac{6\mathsf{E}_{\mathsf{b}}}{21\mathsf{N}_0}}\right) - \frac{49}{16}Q^2\left(\sqrt{\frac{6\mathsf{E}_{\mathsf{b}}}{21\mathsf{N}_0}}\right).\end{aligned}$$

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

Part I

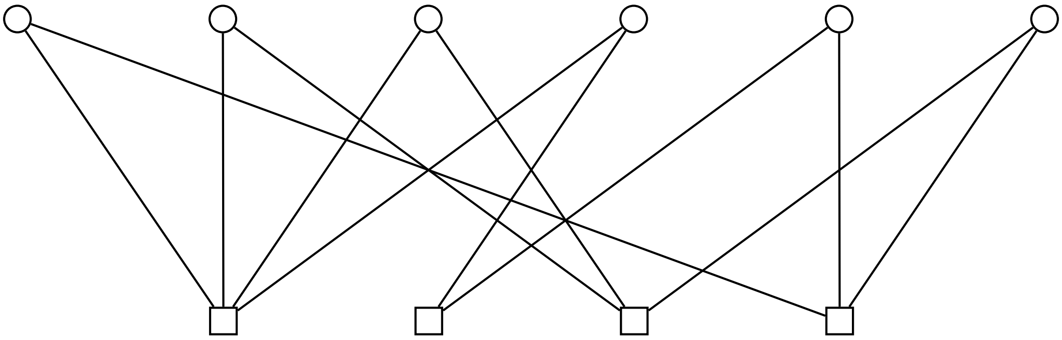
1. 

Figure 2: Tanner graph for $\mathcal{C}_1$.

2. The girth is the length of the shortest cycle in the Tanner graph, which is 4 in this case. It is highlighted in Fig. 3.

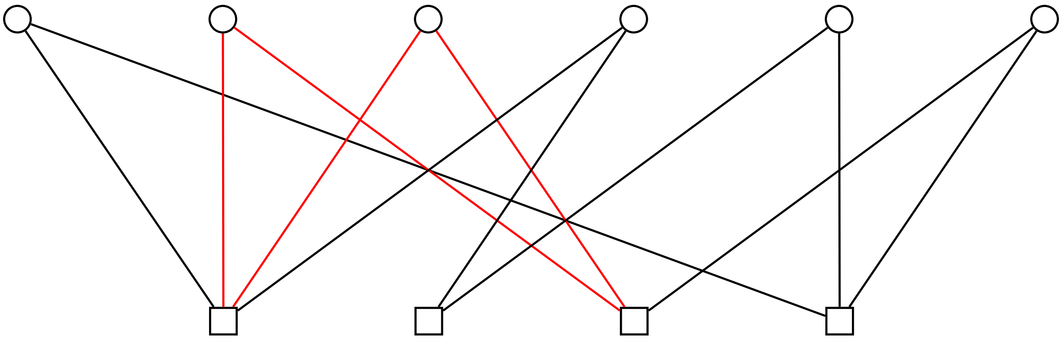


Figure 3: Tanner graph with highlighted girth.

3. By considering the Tanner graph, we get

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

Alternatively, we can get the same results from the parity check matrix $\boldsymbol{G}_1$ by considering the weight of the rows and the columns. The given code is an irregular LDPC code since the CNs are of different degrees.

Part II

1. If we can find a generator matrix that creates all the codewords in $\mathcal{C}_2$. We can see that the generator matrix must consist of 2 rows and pick the thrid and the fourth codeword as the generator matrix,

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

We now need to check for each possible information sequence if it creates a valid codeword,

$$\begin{aligned}[00]\,\boldsymbol{G}_{\mathsf{s}} &= [000000]\\ [01]\,\boldsymbol{G}_{\mathsf{s}} &= [011111]\\ [10]\,\boldsymbol{G}_{\mathsf{s}} &= [101011]\\ [11]\,\boldsymbol{G}_{\mathsf{s}} &= [110100].\end{aligned}$$

We note that all of those results are codewords and hence $\mathcal{C}_2$ is a linear block code.

By the choice of our generator matrix, $\boldsymbol{G}$ is already in systematic form, i.e., $\boldsymbol{G}_{\mathsf{s}} = \boldsymbol{G}$. The parity check matrix then results to

$$\boldsymbol{H}_{\mathsf{s}} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

2. The error detection and error correction capabilities over the BSC depend on the minimum Hamming distance of the code, which is $d_{\mathsf{min}} = 3$. Hence, $\mathcal{C}_2$ can detect error patterns with up to 2 errors and correct all error patterns with up to 1 error.

3. It is important to realize that minimizing the Hamming distance does *not* correspond to ML decoding here. Notationwise, we introduce a lower index i to refer to a particular use of the channel. The ML codeword is given by

$$\begin{aligned} c_{\mathrm{ML}} &= \arg\max_{c\in\mathcal{C}} p(\hat{\boldsymbol{c}}|c) \\ &= \arg\max_{c\in\mathcal{C}} \prod_{i=1}^{3} p(\hat{c}_i^{(1)}, \hat{c}_i^{(2)}|c_i^{(1)}, c_i^{(2)}) \\ &= \arg\max_{c\in\mathcal{C}} \sum_{i=1}^{3} \log p(\hat{c}_i^{(1)}, \hat{c}_i^{(2)}|c_i^{(1)}, c_i^{(2)}) \\ &= \arg\max_{c\in\mathcal{C}} \sum_{i=1}^{3} \lambda_i, \end{aligned}$$

where the second equality follows from the fact that the channel is memoryless (from QPSK symbol to QPSK symbol), and for the second equality we just applied the log function, which does not change the argument of the maximization. For the last equality, we have introduced λ_i as our decoding metric:

$$\lambda_i \triangleq \log p(\hat{c}_i^{(1)}, \hat{c}_i^{(2)}|c_i^{(1)}, c_i^{(2)}).$$

We now create a transition table for $p(\hat{c}_i^{(1)}, \hat{c}_i^{(2)}|c_i^{(1)}, c_i^{(2)})$,

	00	01	10	11
00	$(1-\mathsf{Q})^2$	Q^2	$\mathsf{Q}-\mathsf{Q}^2$	$\mathsf{Q}-\mathsf{Q}^2$
01	Q^2	$(1-\mathsf{Q})^2$	$\mathsf{Q}-\mathsf{Q}^2$	$\mathsf{Q}-\mathsf{Q}^2$
10	$\mathsf{Q}-\mathsf{Q}^2$	$\mathsf{Q}-\mathsf{Q}^2$	$(1-\mathsf{Q})^2$	Q^2
11	$\mathsf{Q}-\mathsf{Q}^2$	$\mathsf{Q}-\mathsf{Q}^2$	Q^2	$(1-\mathsf{Q})^2$

where Q refers to the q-function$\mathsf{Q}(\mathsf{E_s}/N_0)$. We then normalize the whole table by $(1-\mathsf{Q})^2$, take the logarithm and remove any common factor

	00	01	10	11
00	0	2	1	1
01	2	0	1	1
10	1	1	0	2
11	1	1	2	0

For each codeword in $\mathcal{C}_2$ we can then calculate the cumulative metric,

$$\begin{aligned}
\boldsymbol{c} = (000000) &: \sum_{i=1}^{3} \lambda_i = 1+1+1 = 3 \\
\boldsymbol{c} = (110100) &: \sum_{i=1}^{3} \lambda_i = 2+1+1 = 4 \\
\boldsymbol{c} = (101011) &: \sum_{i=1}^{3} \lambda_i = 0+0+2 = 2 \\
\boldsymbol{c} = (011111) &: \sum_{i=1}^{3} \lambda_i = 1+2+2 = 5.
\end{aligned}$$

We note that the minimum cumulative metric is achieved in the case of $\boldsymbol{c} = (101011)$. Hence, the ML codeword is $\boldsymbol{c} = (101011)$.

Problem 4 - Convolutional Codes and the Viterbi Algorithm [15 points]

Part I

1.

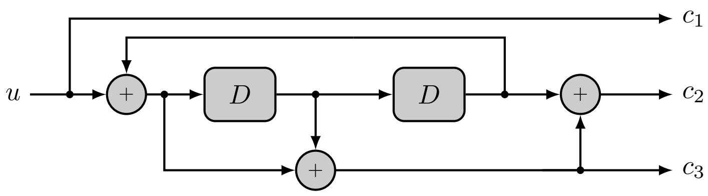


Figure 4: Encoder $\mathcal{E}_{\mathsf{RSC}}$

2. Multiplying $\boldsymbol{G}(\mathsf{D})$ by $(1+\mathsf{D}^2)$ results in

$$\boldsymbol{G}_{\mathsf{s}}(\mathsf{D}) = \begin{bmatrix} 1+\mathsf{D}^2 & 1+\mathsf{D}+\mathsf{D}^2 & 1+\mathsf{D} \end{bmatrix}.$$

The corresponding block diagram is depicted in Fig. 5.

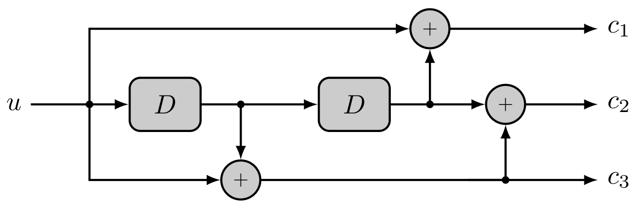


Figure 5: Encoder $\mathcal{E}_1$

3.

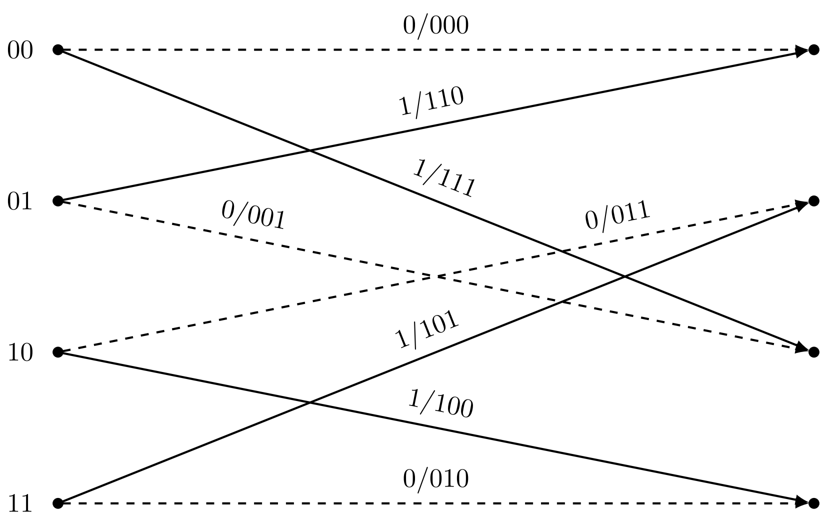


Figure 6: Full Trellis of $\mathcal{E}_{\mathsf{RSC}}$. Dashed lines correspond to a 0 bit and soild lines to a 1 bit.

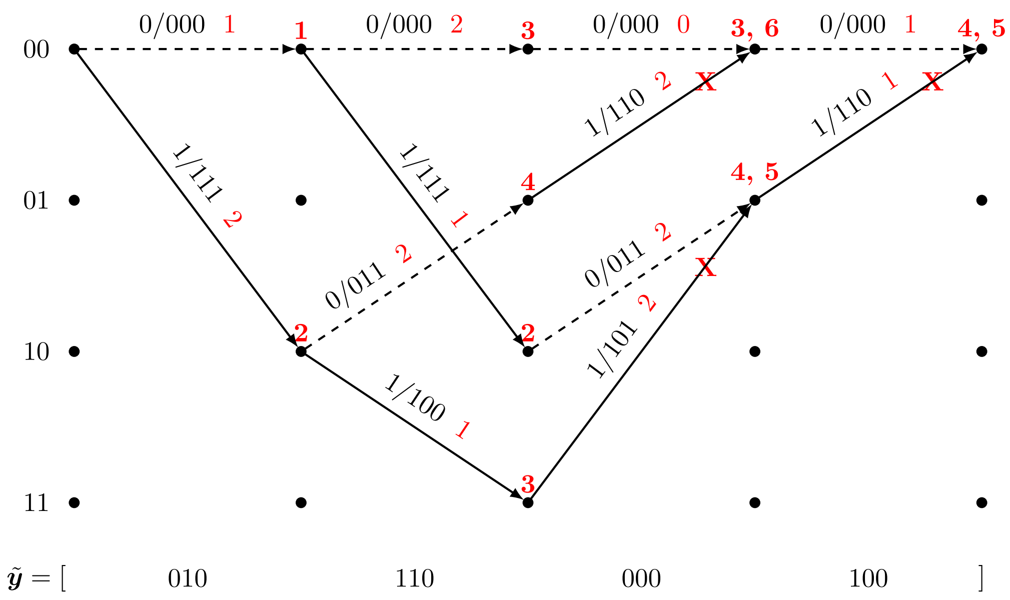


Figure 7: Viterbi $\mathcal{E}_{\mathsf{RSC}}$. Dashed lines correspond to a 0 bit and soild lines to a 1 bit.

4. We can see in 7 that the path with the lowest cumulative metric corredsonds to the codeword 000 000 000 000. The corresponding information sequences is 0000.

Part II

1. Since we have five columns in $\boldsymbol{H}$, we need five sections in the trellis.

2. The parity check equation needs to sum to zero. Hence, the trellis must end in the zero state and therefore be zero-terminated.

3.

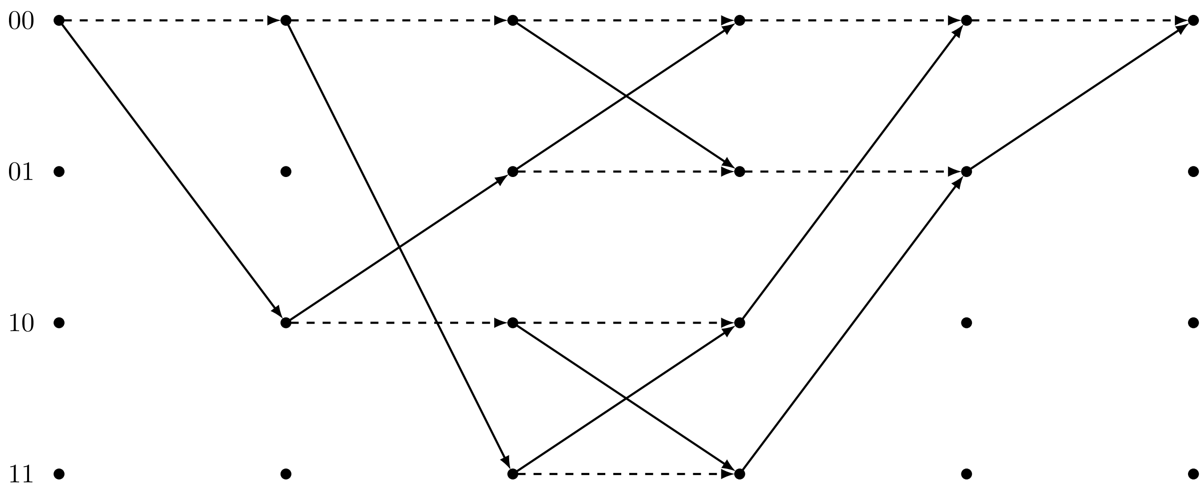


Figure 8: Trellis of the block code. Dashed lines correspond to a 0 bit and soild lines to a 1 bit.