Solution Sheet 2

Last modified November 18, 2019

Problem 1

1. We know that for reliable transmission the minimum required $\mathsf{E_b/N_0}$ is given by $\mathsf{E_b/N_0} > (2^r - 1)/r$. Substituting the different spectral efficiencies we obtain:

   For $r = 0.5$, $\mathsf{E_b/N_0} > .828 = -0.82$ dB

   For $r = 1$, $\mathsf{E_b/N_0} > 1 = 0$ dB

   For $r = 2$, $\mathsf{E_b/N_0} > 1.5 = 1.76$ dB

2. We know that for reliable transmission the minimum required $\mathsf{E_s/N_0}$ is given by $\mathsf{E_s/N_0} > 2^r - 1$. Substituting the different spectral efficiencies we obtain:

   For $r = 0.5$, $\mathsf{E_s/N_0} > .414 = -3.83$ dB

   For $r = 1$, $\mathsf{E_s/N_0} > 1 = 0$ dB

   For $r = 2$, $\mathsf{E_s/N_0} > 3 = 4.77$ dB

3. We know that for reliable transmission the minimum required $\mathsf{E_b/N_0}$ is given by $\mathsf{E_b/N_0} > (2^r - 1)/r$.

   For $R_\mathsf{b} = 60$ kbits/s, the spectral efficiency is $r_1 = R_\mathsf{b}/W = 12$ and the required $\mathsf{E_b/N_0} > 341.25 = 25.33$ dB, while with the given limited power of 0.01 W, $\mathsf{E_b/N_0} = 0.1667 = -7.78$ dB. Reliable communication is not possible.

   For $R_\mathsf{b} = 50$ kbits/s, the spectral efficiency is $r_1 = R_\mathsf{b}/W = 10$ and the required $\mathsf{E_b/N_0} > 102.3 = 20.09$ dB, while with the given limited power of 0.01 W, $\mathsf{E_b/N_0} = 0.2 = -6.98$ dB. Reliable communication is not possible.

4. Considering no bandwidth constraints, meaning $W \to \infty$, the capacity approaches a limit at

   $\mathsf{C_{AWGN-C}} = \frac{P}{\mathsf{N_0}} \log_2(e) = \frac{0.001}{10^{-8}} \log_2(e) = 14.43 \log_2(e) = 144.43$ kbits/s.

   Hence, reliable communication is not possible for $R_\mathsf{b} = 150$ kbits/s. The maximum possible rate is 144.43 kbits/s when the spectral efficiency $r \to 0$.

5. There are two main disadvantages when using coding: delay and complexity. In principle, the delay comes from the fact that in the channel coding theorem, it is necessary to let the block length N go to ∞. The complexity comes from the fact that we need to decode.

6. When $P \to \infty$, capacity goes to infinity

   $\lim_{P\to\infty} \mathsf{C_{AWGN-C}} = \infty$.

   When $W \to \infty$, capacity approaches a limit

   $\lim_{W\to\infty} \mathsf{C_{AWGN-C}} = \frac{P}{\mathsf{N_0}} \log_2(e)$.

Problem 2

In order to find the number of (approximately) disjoint noise spheres M, we simply divide the volume of the entire output sphere by the volume of one noise sphere and obtain

$$M = \frac{\frac{(\pi(N\mathsf{E_s}+N\mathsf{N_0}/2)^{N/2}}{(N/2)!}}{\frac{(\pi N\mathsf{N_0}/2)^{N/2}}{(N/2)!}} = \left(\frac{N\mathsf{E_s} + N\mathsf{N_0}/2}{N\mathsf{N_0}/2}\right)^{N/2} = \left(1 + \frac{\mathsf{E_s}}{\mathsf{N_0}/2}\right)^{N/2} = \left(1 + \frac{\mathsf{E_s}}{\sigma^2}\right)^{N/2}$$

This is then also the maximum number of input vectors (or signals) that can be reliably distinguished at a receiver. If we take the logarithm to base 2 of this number, then we get the total number of bits that can be transmitted (in N uses of the channel):

$$\log_2 M = \frac{N}{2}\log_2\left(1+\frac{\mathsf{E_s}}{\sigma^2}\right).$$

The last step involves normalizing the above number by N, to obtain the maximum number of bits *per channel use*:

$$\frac{\log_2 M}{N} = \frac{1}{2}\log_2\left(1+\frac{\mathsf{E_s}}{\sigma^2}\right),$$

which is equal to the capacity of the channel, $\mathsf{C_{AWGN-D}}$, as defined in the lecture.

Problem 3

Part I

We use the notation

$$\mathcal{N}_x(\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

to denote a Gaussian distribution with argument x. Now, the random variable X is our parameter of interest, which has a certain prior $P_X(0)=a$, $P_X(1)=1-a$. Y is the observation, where the likelihood $p_{Y|X}(y|x)$ is directly given in the problem description, that is, $p_{Y|X}(y|0)=\mathcal{N}_y(-1,1)$ and $p_{Y|X}(y|1)=\mathcal{N}_y(1,4)$.

1. The optimal detector (maximum a posteriori) is given by

$$\hat{x} = \underset{x\in\{0,1\}}{\operatorname{argmax}}\, p_{Y|X}(y|x)P_X(x).$$

   Because the maximization is only over two possibilities, one can rewrite the detector as

$$p_{Y|X}(y|0)P_X(0) \underset{\hat{x}=0}{\overset{\hat{x}=1}{\lessgtr}} p_{Y|X}(y|1)P_X(1).$$

   As a first step, we plug in the likelihood and the prior to obtain

$$\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{(y+1)^2}{2}\right)\cdot a \underset{\hat{x}=0}{\overset{\hat{x}=1}{\lessgtr}} \frac{1}{2\sqrt{2\pi}}\exp\left(-\frac{(y-1)^2}{2\cdot 4}\right)\cdot(1-a).$$

   Rearranging and taking the (natural) logarithm eventually leads to

$$3y^2+10y \underset{\hat{x}=1}{\overset{\hat{x}=0}{\lessgtr}} 8\log\left(2\cdot\frac{a}{1-a}\right)-3,$$

   which has the desired form.

2. The only thing left to do here is to insert the prior into the above rule, obtaining

$$3y^2+10y \underset{\hat{x}=1}{\overset{\hat{x}=0}{\lessgtr}} -3.$$

Part II

Formally, the ML estimator is always written in terms of the likelihood as

$$\hat{\alpha}_{\mathsf{ML}} = \underset{\alpha}{\operatorname{argmax}}\, p(\boldsymbol{r}|\boldsymbol{s},\alpha).$$

The likelihood in this case can be written as a product of many Gaussian distributions (because the noise is i.i.d.) as

$$p(\boldsymbol{r}|\boldsymbol{s},\alpha) = \prod_i \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(r_i-\alpha s_i)^2}{2\sigma^2}\right).$$

Notice that the maximization of the likelihood in terms of α can alternatively be done over the log-likelihood, which is proportional to

$$\log p(\boldsymbol{r}|\boldsymbol{s},\alpha) \propto -\sum_i (r_i-\alpha s_i)^2 = f(\alpha).$$

Maximization of $f(\alpha)$ is achieved by setting the derivate $f'(\alpha)$ equal to zero which leads to

$$2\sum_i s_i(r_i-\alpha s_i) = 0.$$

Rearranging the above equality for α gives the final answer as

$$\hat{\alpha}_{\mathsf{ML}} = \frac{\sum_i s_i r_i}{\sum_i s_i^2}.$$

Problem 4

The ML estimator is given by

$$\hat{\phi}_{\mathsf{ML}} = \underset{\phi}{\operatorname{argmax}}\, p(\boldsymbol{y}|\boldsymbol{x},\phi),$$

where the likelihood function in this case is

$$p(\boldsymbol{y}|\boldsymbol{x},\phi) = \left(\frac{1}{2\pi\sigma^2}\right)^{N/2} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=0}^{N-1}\left|y_i - x_i e^{\jmath\phi}\right|^2\right).$$

Alternatively, we can maximize over the log-likelihood function, which is proportional to

$$\begin{aligned}
\log p(\boldsymbol{y}|\boldsymbol{x},\phi) &\propto -\sum_{i=0}^{N-1} |y_i - x_i e^{\jmath\phi}|^2 \\
&= -\sum_{i=0}^{N-1}\left(|y_i|^2 - 2\Re(y_i x_i^* e^{-\jmath\phi}) + |x_i|^2\right) \\
&\propto \sum_{i=0}^{N-1} \Re(y_i x_i^* e^{-\jmath\phi}) \\
&= \sum_{i=0}^{N-1} \frac{1}{2}(y_i x_i^* e^{-\jmath\phi} + y_i^* x_i e^{\jmath\phi}) \\
&\propto z e^{-\jmath\phi} + z^* e^{\jmath\phi} \\
&= \Re(z e^{-\jmath\phi}),
\end{aligned}$$

where we defined $z = \sum_{i=0}^{N-1} y_i x_i^*$. Maximization of the last equation with respect to z is obtained by setting $\hat{\phi}_{\mathrm{ML}} = \arg z$. This result can be obtained intuitively: Imagine a complex number z in the two-dimensional complex plane. Then, for which phase rotation do you maximize the real part? Alternatively, the same result can be obtained by differentiating the second last equation with respect to ϕ and setting the derivate to zero as

$$-\jmath z e^{-\jmath\phi} + \jmath z^* e^{\jmath\phi} = 0.$$

After rearranging, we obtain $e^{2\jmath\phi} = z/z^*$, which gives exactly the same answer. In summary, the ML phase offset estimator is

$$\hat{\phi}_{\mathsf{ML}} = \arg\left(\sum_{i=0}^{N-1} y_i x_i^*\right).$$

Problem 5

From $P_X(x) = \sum_y P_{X,Y}(x,y)$ (and similarly for $P_Y(y)$), we have $P_X(0) = 2/3$, $P_X(1) = 1/3$, $P_Y(0) = 1/3$, and $P_Y(1) = 2/3$.

1. $\mathsf{H}(X) = 0.9183$

   $\mathsf{H}(Y) = 0.9183$

2. $\mathsf{H}(X|Y) = P_Y(0)\mathsf{H}(X|Y=0) + P_Y(1)\mathsf{H}(X|Y=1) = 0 + 2/3 \cdot 1 = 2/3$

   $\mathsf{H}(Y|X) = P_X(0)\mathsf{H}(Y|X=0) + P_X(1)\mathsf{H}(Y|X=1) = 2/3 \cdot 1 + 0 = 2/3$

3. $\mathsf{H}(X,Y) = \mathsf{H}(X) + \mathsf{H}(Y|X) = 1.585$

4. $\mathsf{H}(Y) - \mathsf{H}(Y|X) = \mathsf{I}(X;Y) = 0.2516$

Problem 6

1. Applying the Huffman algorithm gives us the following table

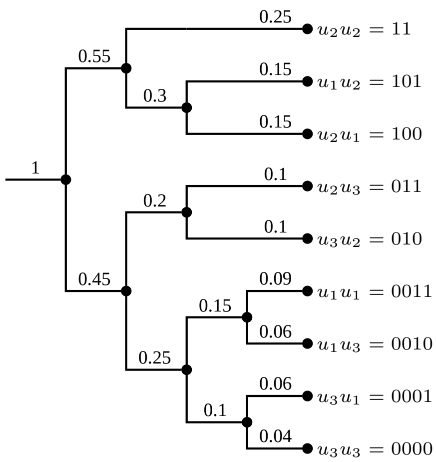
Code	Symbol	Probability			
0	1	1/3	1/3	2/3	1
11	2	1/3	1/3	1/3	
101	3	1/4	1/3		
100	4	1/12			

   which gives codeword lengths of 1,2,3,3 for the different codewords.

2. Both set of lengths 1,2,3,3 and 2,2,2,2 satisfy the Kraft inequality, and they both achieve the same expected length (2 bits) for the above distribution. Therefore they are both optimal.

Problem 7

1. The Huffman coding algorithm can yield the following tree for two consecutive symbols of the source.



2. The source entropy is given by $\mathsf{H}(U) = -\sum_{i=1}^{3} P_U(u_i)\log_2(P_U(u_i)) = 1.49$ bits.

   The expected codeword length is given by $\bar{L}_2 = \sum_{i=1}^{3} P_U(u_i)c_i = 3$ bits.

   Therefore, the code efficiency is $\eta = 2\mathsf{H}(U)/\bar{L}_2 \approx 0.99$.

3. The code efficiency can be increased by coding over multiple symbols. However, this also increases the complexity and the decoding delay.

Problem 8

1. $Z = X + Y$. Hence $P(Z = z|X = x) = P(Y = z - x|X = x)$.

$$
\begin{aligned}
\mathsf{H}(Z|X) &= \sum_x P(x)\mathsf{H}(Z|X = x) \\
&= -\sum_x P(x) \sum_z P(Z = z|X = x)\log P(Z = z|X = x) \\
&= -\sum_x P(x) \sum_z P(Y = z - x|X = x)\log P(Y = z - x|X = x) \\
&= -\sum_x P(x) \sum_y P(Y = y|X = x)\log P(Y = y|X = x) \\
&= \sum_x P(x)\mathsf{H}(Y|X = x) \\
&= \mathsf{H}(Y|X).
\end{aligned}
$$

   If X and Y are independent, then $\mathsf{H}(Y|X) = \mathsf{H}(Y)$. Since $\mathsf{I}(X;Z) \geq 0$, we have $\mathsf{H}(Z) \geq \mathsf{H}(Z|X) = \mathsf{H}(Y|X) = \mathsf{H}(Y)$. Similarly we can show that $\mathsf{H}(Z) \geq \mathsf{H}(X)$.

2. Consider the following joint distribution for X and Y. Let

$$
X = -Y = \begin{cases} 1 & \text{with probability } 1/2 \\ 0 & \text{with probability } 1/2. \end{cases}
$$

   Then $\mathsf{H}(X) = \mathsf{H}(Y) = 1$, but $Z = 0$ with probability 1 and hence $\mathsf{H}(Z) = 0$.