École Polytechnique Fédérale de Lausanne
School of Computer and Communication Sciences

Handout 19 Advanced Digital Communications
Solutions to Midterm Exam Nov. 9, 2016

Solution 1.

(a)

(i) The MAP rule is simply

$$\hat{X}(y) = \begin{cases} +1, & y \geq 0, \\ -1, & y \leq 0. \end{cases}$$

(ii) The MAP rule is

$$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y+1)^2}{2\sigma^2}} (1-p) \underset{\hat{X}=+1}{\overset{\hat{X}=-1}{\gtrless}} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-1)^2}{2\sigma^2}} p$$

which simplifies to

$$\hat{X}(y) = \begin{cases} +1, & y \geq \frac{\sigma^2}{2} \ln \frac{1-p}{p}, \\ -1, & y \leq \frac{\sigma^2}{2} \ln \frac{1-p}{p}. \end{cases}$$

(iii) Let $\tau := \frac{\sigma^2}{2} \ln \frac{1-p}{p}$. Then

$$\begin{aligned} \Pr\{\text{error}|X=+1\} &= \Pr\{+1+Z \leq \tau\} \\ &= \Pr\{Z \leq \tau - 1\} \\ &= \Pr\{Z \geq 1 - \tau\} \\ &= Q\Big(\frac{1-\tau}{\sigma}\Big) = Q\Big(\frac{1}{\sigma} - \frac{\sigma}{2} \ln \frac{1-p}{p}\Big). \end{aligned}$$

Similarly,

$$\begin{aligned} \Pr\{\text{error}|X=-1\} &= \Pr\{-1+Z \geq \tau\} \\ &= \Pr\{Z \geq 1 + \tau\} \\ &= Q\Big(\frac{1+\tau}{\sigma}\Big) = Q\Big(\frac{1}{\sigma} + \frac{\sigma}{2} \ln \frac{1-p}{p}\Big). \end{aligned}$$

Consequently,

$$\Pr\{\text{error}\} = pQ\Big(\frac{1}{\sigma} - \frac{\sigma}{2} \ln \frac{1-p}{p}\Big) + (1-p)Q\Big(\frac{1}{\sigma} + \frac{\sigma}{2} \ln \frac{1-p}{p}\Big).$$

(iv) Let $P_{\mathrm{e}}(p)$ denote the probability of error of the MAP rule (that we computed in (ii)). We know from Homework 3 that $P_{\mathrm{e}}(p)$ is concave in p. In this case it is also symmetric around $p = \frac{1}{2}$ (i.e., $P_{\mathrm{e}}(p) = P_{\mathrm{e}}(1-p)$). Thus it is maximized at $p^\star = \frac{1}{2}$. (A more formal proof: Suppose P_{e} is maximized at some other point $\tilde{p} < \frac{1}{2}$. Then $P_{\mathrm{e}}(\tilde{p}) = P_{\mathrm{e}}(1-\tilde{p})$ but for any value of $p \in [\tilde{p}, 1-\tilde{p}]$, $P_{\mathrm{e}}(p) \geq P_{\mathrm{e}}(\tilde{p})$ because of concavity — the function lies above the cord. Thus, $\tilde{p}$ cannot be the maximizer.)

(v) Going along the same lines as in (iii), we get

$$P_{\mathrm{e}}(p,q) := qQ\Big(\frac{1}{\sigma} - \frac{\sigma}{2}\ln\frac{1-p}{p}\Big) + (1-q)Q\Big(\frac{1}{\sigma} + \frac{\sigma}{2}\ln\frac{1-p}{p}\Big).$$

(vi) For any pair of real numbers α and β, $\max_{q\in[0:1]}\{\alpha q+\beta(1-q)\} = \max\{\alpha,\beta\}$ (and the maximum is attained with $q=1$ if $\alpha \geq \beta$ and $q=0$ otherwise). Therefore,

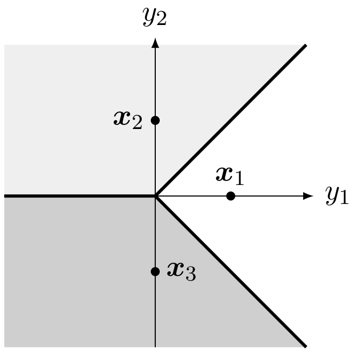
$$\begin{aligned}\max_{q\in[0:1]} P_{\mathrm{e}}(p,q) &= \max\Big\{Q\Big(\frac{1}{\sigma} - \frac{\sigma}{2}\ln\frac{1-p}{p}\Big), Q\Big(\frac{1}{\sigma} + \frac{\sigma}{2}\ln\frac{1-p}{p}\Big)\Big\} \\ &= \begin{cases} Q\Big(\frac{1}{\sigma} - \frac{\sigma}{2}\ln\frac{1-p}{p}\Big), & \text{if } p \leq \frac{1}{2}, \\ Q\Big(\frac{1}{\sigma} + \frac{\sigma}{2}\ln\frac{1-p}{p}\Big), & \text{if } p \geq \frac{1}{2}. \end{cases}\end{aligned}$$

(vii) Since $\ln\frac{1-p}{p} = \ln(1/p-1)$ is decreasing in p and $Q(\cdot)$ is decreasing in its argument, from the results of (vi) it is obvious that $P_{\mathrm{e}}(p,*)$ decreases from $Q(-\infty)=1$ to $Q(\frac{1}{\sigma})$ as p increases from 0 to $\frac{1}{2}$ and again increases to $Q(-\infty)=1$ as p further increases to 1. (Indeed, $P_{\mathrm{e}}(p,*)$ is symmetric around $p=\frac{1}{2}$.) Consequently, it is minimized at $p^\star = \frac{1}{2}$ with the minimum value of $Q(\frac{1}{\sigma})$.

This is, by construction, the optimal decoder since, for any other p, if the transmitter chooses the prior $\Pr\{X=+1\} = q \in \{0,1\}$ that maximizes $P_{\mathrm{e}}(p,q)$ we will have a probability of error $P_{\mathrm{e}}(p,q) = P_{\mathrm{e}}(p,*) \geq P_{\mathrm{e}}(p^\star,*)$.

(b)

(i) The decision regions are as follows:



(ii) The union bound gives $P_{\mathrm{e}} \leq (M-1)Q(d_{\min}/(2\sigma))$ where M is the number of codewords and $d_{\min}$ is the minimum distance. In this case $d_{\min} = \|\boldsymbol{x}_1 - \boldsymbol{x}_2\| = \sqrt{2}$ and $M=3$, thus,

$$P_{\mathrm{e}} \leq 2Q\Big(\frac{\sqrt{2}}{2\sigma}\Big).$$

Solution 2.

(a) The spectrum of V is

$$S_V(z) = -\frac{1}{2}z^{-1} + \frac{5}{4} - \frac{1}{2}z = \underbrace{(1-\frac{1}{2}z^{-1})}_{A(z)}\underbrace{(1-\frac{1}{2}z)}_{A^*(1/z^*)}.$$

The whitening filter that results in a causal effective channel is

$$D_W(z) = \frac{1}{A^*(1/z^*)} = \frac{1}{1-\frac{1}{2}z}.$$

The output of whitening filter will be

$$S[n] = d_W[n] * U[n] = \sum_k I[n-k]f[k] + W[n]$$

where $W[n]$ is a circularly symmetric zero-mean noise with variance 1 (since $S_W(z) = S_V(z)D_W(z)D_W^*(1/z^*) = 1$) and $F(z) = (1-\frac{1}{2}z^{-1})(1-\frac{1}{2}z)D_W(z) = (1-\frac{1}{2}z^{-1})$, hence $f[k] = \delta[k] - \frac{1}{2}\delta[k-1]$. We can write the channel model more explicitly as

$$S[n] = I[n] - \frac{1}{2}I[n-1] + W[n].$$

(b) Since the effective channel has 2 taps (hence a memory of length 1), the Viterbi algorithm requires $|\mathcal{A}|$ states.

(c) The zero-forcing equalizer is simply the inverse of the channel response, $D(z) = -\frac{1}{2}z^{-1} + \frac{5}{4} - \frac{1}{2}z$, that is

$$D_{\mathrm{ZF}}(z) = \frac{1}{-\frac{1}{2}z^{-1} + \frac{5}{4} - \frac{1}{2}z}.$$

Replacing $z = e^{j2\pi f}$ we get

$$D_{\mathrm{ZF}}(f) = \frac{1}{\frac{5}{4} - \cos(2\pi f)}.$$

(d) After filtering we will get

$$\hat{I}_{\mathrm{ZF}} = I[n] + \tilde{V}[n],$$

where $\tilde{V}[n]$ has spectrum

$$S_{\tilde{V}}(f) = S_V(f)|D_{\mathrm{ZF}}(f)|^2 = \frac{1}{D(f)}$$

(where the last equality follows since $S_V(f) = D(f)$ and $D_{\mathrm{ZF}}(f) = \frac{1}{D(f)}$.) In our case,

$$S_{\tilde{V}}(f) = \frac{1}{\frac{5}{4} - \cos(2\pi f)}.$$

We finally have

$$\mathbb{E}[\tilde{V}[n]^2] = R_{\tilde{V}}[0] = \int_{-\frac{1}{2}}^{\frac{1}{2}} S_{\tilde{V}}(f)\,\mathrm{d}f = \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{1}{\frac{5}{4} - \cos(2\pi f)}\,\mathrm{d}f = \frac{4}{3}.$$

(e) Since the information symbols are i.i.d., zero-mean, with unit variance, $S_I(z) = 1$ and the noise spectrum is $S_V(z) = D(z)$,

$$D_{\mathrm{LMMSE}}(z) = \frac{1}{D(z)+1} = \frac{1}{\frac{9}{4} - \frac{1}{2}z^{-1} - \frac{1}{2}z},$$

which shows

$$D_{\mathrm{LMMSE}}(f) = \frac{1}{\frac{9}{4} - \cos(2\pi f)}.$$

(f) In the Z-domain we have

$$\hat{I}_{\mathrm{LMMSE}}(z) = \frac{1}{1+D(z)}D(z)I(z) + \frac{1}{1+D(z)}V(z).$$

Consequently,

$$E(z) := \hat{I}_{\mathrm{LMMSE}}(z) - I(z) = \frac{-1}{1+D(z)}I(z) + V(z)\frac{1}{1+D(z)}.$$

Since $I[n]$ and $V[n]$ are independent zero-mean processes, $S_I(z) = 1$ and $S_V(z) = D(z)$,

$$S_E(z) = \frac{1}{(1+D(z))(1+D^*(1/z^*))} + \frac{D(z)}{(1+D(z))(1+D^*(1/z^*))} = \frac{1}{1+D^*(1/z^*)}.$$

Therefore,

$$S_E(f) = \frac{1}{1+D^*(f)} = \frac{1}{\frac{9}{4} - \cos(2\pi f)},$$

which implies

$$\mathbb{E}[|E[n]|^2] = R_E[0] = \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{1}{\frac{9}{4} - \cos(2\pi f)}\,\mathrm{d}f = \frac{4}{\sqrt{65}}.$$

(g) Under the assumption that the noise is Gaussian and with antipodal information symbols, the probability of error is smaller if the noise variance is smaller. The LMMSE, by definition, minimizes the effective noise variance at the input of the decision device, hence, yields the smallest probability of error (that no other linear equalizer can ever beat!).

Solution 3.

(a) It is equal to 3 symbols since we have 4 taps.

(b) We need $x[-3] = x[1]$, $x[-2] = x[2]$, and $x[-1] = x[3]$.

(c) In order to determine the channel gains we take the length-4 DFT of the vector $(\frac{3}{4}, \frac{1}{4}, \frac{3}{4}, \frac{1}{4})$ and we multiply the result by $\sqrt{N} = 2$ (this factor is due to the particular way we normalized the DFT; see the hint at the end of the problem). The 4 independent channels have gains 2, 0, 1, and 0 respectively. For each channel the noise is Gaussian of mean 0 and variance 1.

(d) To bring the channels into standard form we divide by the channel gain which changes the variance by the square of the channel gain. After the conversion all channel gains are 1 and the variances are $1/4$, ∞, 1, and ∞, respectively. Note that the second and fourth channels have infinite variance (they used to have zero gain). So these two channels are useless for information transmission.

(e) We use water-filling. Clearly, we will only use the first and the third channels. Their respective variances are $1/4$ and 1. If we pick the constant to be 2 and "fill up" the channels up to this constant then this means that we assign a power of $7/4$ to the first channel and a power of 1 to the second channel.

(f) If we plug in these noise powers and variance into the capacity formula we get a total capacity of $\log_2(1+7) + \log_2(1+1) = 4$. Hence we can transmit up to 4 bits over a length of $3+4$ channel inputs. The number of bits per channel use is therefore $4/7$.

(g) The most effective method would be to increase N to a much larger value. Currently we loose a lot of rate because the gap is of the same order as N.