MIDTERM SOLUTIONS

Problem 1

(a) Since the prior probability q is known at the decoder, we know from the class that probability of error is minimized by the maximum a posteriori (MAP) decision rule, i.e., for a given observation $y \in \{0,1\}$, we decide

$$\mathbb{P}[X=0|Y=y] \underset{\hat{X}=1}{\overset{\hat{X}=0}{\gtrless}} \mathbb{P}[X=1|Y=y].$$

If $Y=1$, then we have

$$\begin{aligned} \mathbb{P}[X=0|Y=1] &= 0 \\ \mathbb{P}[X=1|Y=1] &= 1, \end{aligned}$$

and one should always decide $\hat{X}=1$. If $Y=0$, we use Bayes' formula

$$\mathbb{P}[X=x|Y=y] = \frac{\mathbb{P}[Y=y|X=x]\mathbb{P}[X=x]}{\mathbb{P}[Y=y]}$$

and obtain

$$\begin{aligned} \mathbb{P}[X=0|Y=0] &= \frac{1\ q}{\mathbb{P}[Y=0]} \\ \mathbb{P}[X=1|Y=0] &= \frac{p(1-q)}{\mathbb{P}[Y=0]}. \end{aligned}$$

If $q > p(1-q)$ and $Y=0$, then we always decide $\hat{X}=0$. If $q < p(1-q)$ and $Y=0$, we always decide $\hat{X}=1$.

(b) Attention: in this scheme, the **same input symbol is reapeated** n times, i.e.,

$$\mathbf{X} = \begin{cases} (0,0,0,\dots,0) & \text{w.p.} \quad q \quad (\text{ if } X=0) \\ (1,1,1,\dots,1) & \text{w.p.} \quad 1-q \quad (\text{ if } X=1). \end{cases}$$

Again, we apply the MAP rule. Assume that a certain $\mathbf{y} = (y_1,\dots,y_n)$ is observed at the decoder. Note that as soon as $y_i = 1$ for at least one $i \in \{1,\dots,n\}$, we have

$$\begin{aligned} \mathbb{P}[\mathbf{X}=\mathbf{0}|\mathbf{Y}=\mathbf{y}] &= 0 \\ \mathbb{P}[\mathbf{X}=\mathbf{1}|\mathbf{Y}=\mathbf{y}] &= 1, \end{aligned}$$

because if $\mathbf{X}=\mathbf{0}$, we can impossibly have $y_i = 1$ for any i. Thus, the decision would be $\hat{X}=1$. The only ambiguous case is when $\mathbf{y}=\mathbf{0}=(0,\dots,0)$. In this case, we use Bayes'rule to compute

$$\begin{aligned} \mathbb{P}[X=0|\mathbf{Y}=\mathbf{0}] &= \frac{1^n\ q}{\mathbb{P}[\mathbf{Y}=\mathbf{0}]} \\ \mathbb{P}[X=1|\mathbf{Y}=\mathbf{0}] &= \frac{p^n(1-q)}{\mathbb{P}[\mathbf{Y}=\mathbf{0}]}. \end{aligned}$$

Therefore, if $\mathbf{Y} = \mathbf{0}$ and $p^n > \frac{q}{1-q}$, we decide $\hat{X} = 1$. If $\mathbf{Y} = \mathbf{0}$ and $p^n < \frac{q}{1-q}$, we rather decide $\hat{X} = 0$.

(c) This time, we ask for the maximum likelihood (ML) decision rule, i.e., for a given observation $\mathbf{y} \in \{0,1\}^n$, we decide

$$\mathbb{P}[\mathbf{Y} = \mathbf{y} | H = H_0] \underset{\hat{H}=H_1}{\overset{\hat{H}=H_0}{\gtrless}} \mathbb{P}[\mathbf{Y} = \mathbf{y} | H = H_1],$$

where H is a random variable that takes values H_0 or H_1 with equal probability. We write out the conditional distributions of $\mathbf{Y} = (Y_1, \ldots, Y_n)$:

$$\begin{aligned}\mathbb{P}[\mathbf{Y} = \mathbf{y} | H = H_0] &= \mathbb{P}[\mathbf{Y} = \mathbf{y} | \mathbf{X} = \mathbf{1}] \\ &= p^{n - N_1(\mathbf{y})}(1-p)^{N_1(\mathbf{y})},\end{aligned}$$

where $N_1(\mathbf{y})$ is the number of ones in the vector $\mathbf{y}$.

$$\begin{aligned}\mathbb{P}[\mathbf{Y} = \mathbf{y} | H = H_1] &= \mathbb{P}[\mathbf{Y} = \mathbf{y} | \mathbf{X} = \mathbf{S}] \\ &= \prod_{i=1}^{n} \mathbb{P}[Y_i = y_i | \mathbf{X} = \mathbf{S}] \\ &= \prod_{i=1}^{n} \mathbb{P}[Y_i = y_i | X_i = S_i] \\ &= \prod_{i=1}^{n} \big(\mathbb{P}[Y_i = y_i | X_i = 1]\mathbb{P}[S_i = 1] + \mathbb{P}[Y_i = y_i | X_i = 0]\mathbb{P}[S_i = 0]\big) \\ &= \prod_{i=1}^{n} \left\{ \begin{array}{ll} \text{if } y_i = 0: & p\frac{1}{2} + 1\ \frac{1}{2} \\ \text{if } y_i = 1: & (1-p)\frac{1}{2} + 0\ \frac{1}{2} \end{array} \right\} \\ &= \left(\frac{1}{2}(1+p)\right)^{n - N_1(\mathbf{y})} \left(\frac{1}{2}(1-p)\right)^{N_1(\mathbf{y})}.\end{aligned}$$

We take the log of both sides of the ML rule:

$$\begin{aligned} p^{n - N_1(\mathbf{y})}(1-p)^{N_1(\mathbf{y})} &\underset{\hat{H}=H_1}{\overset{\hat{H}=H_0}{\gtrless}} \left(\frac{1}{2}(1+p)\right)^{n - N_1(\mathbf{y})} \left(\frac{1}{2}(1-p)\right)^{N_1(\mathbf{y})} \\ N_1(\mathbf{y}) \log(1-p) + (n - N_1(\mathbf{y})) \log p &\underset{\hat{H}=H_1}{\overset{\hat{H}=H_0}{\gtrless}} (n - N_1(\mathbf{y})) \log\big(\frac{1}{2}(1+p)\big) \\ &\qquad + N_1(\mathbf{y}) \log\big(\frac{1}{2}(1-p)\big) \\ N_1(\mathbf{y})(\log(1+p) - \log p) &\underset{\hat{H}=H_1}{\overset{\hat{H}=H_0}{\gtrless}} n\big(\log\frac{1}{2} + \log(1+p) - \log p\big) \\ N_1(\mathbf{y}) &\underset{\hat{H}=H_1}{\overset{\hat{H}=H_0}{\gtrless}} \frac{n \log\frac{1}{2} + \log(1+p) - \log(1+p)}{\log(1+p) - \log p}. \end{aligned}$$

Problem 2

(a) We can write the usual binary MAP decoding rule as

$$\mathrm{P}(Y|X=+1)\mathbb{P}[X=+1] \underset{x\,=\,+1}{\overset{x\,=\,-1}{\lessgtr}} \mathrm{P}(Y|X=-1)\mathbb{P}[X=-1].$$

We also know the conditional density function of Y according to the channel model:

$$\mathrm{P}(y|X=x) = \mathrm{P}(Z=y-x|X=x) = \mathrm{P}(Z=y-x) = p_Z(y-x) = \frac{1}{2}e^{-|y-x|}$$

So we have the decision criterion as

$$\frac{1}{2}e^{-|y-1|}\cdot(1-q) \underset{x\,=\,+1}{\overset{x\,=\,-1}{\lessgtr}} \frac{1}{2}e^{-|y+1|}\cdot q$$

Assuming $q \neq 0, 1$ (for the extreme cases the criterion is obtained obviously), we have

$$e^{|y+1|-|y-1|} \underset{x\,=\,+1}{\overset{x\,=\,-1}{\lessgtr}} \frac{q}{1-q}$$

$$|y+1|-|y-1| \underset{x\,=\,+1}{\overset{x\,=\,-1}{\lessgtr}} \ln(\frac{q}{1-q})$$

In the following we consider three cases for q which is constant for the given channel and give the decision criterion for each of them.

- ▶ $q < \frac{1}{1+e^2}$: $\ln(\frac{q}{1-q}) < -2$
  - $y > 1 \Rightarrow |y+1|-|y-1| = 2 > 0 > \ln(\frac{q}{1-q}) \Rightarrow \hat{x} = 1$
  - $-1 \leq y \leq 1 \Rightarrow |y+1|-|y-1| = 2y > -2 > \ln(\frac{q}{1-q}) \Rightarrow \hat{x} = 1$
  - $y < -1 \Rightarrow |y+1|-|y-1| = -2 > \ln(\frac{q}{1-q}) \Rightarrow \hat{x} = 1$
- ▶ $\frac{1}{1+e^2} \leq q \leq \frac{e^2}{1+e^2}$: $-2 \leq \ln(\frac{q}{1-q}) \leq 2$
  - $y > 1 \Rightarrow |y+1|-|y-1| = 2 \geq \ln(\frac{q}{1-q}) \Rightarrow \hat{x} = 1$
  - $-1 \leq y \leq 1 \Rightarrow |y+1|-|y-1| = 2y \Rightarrow \begin{cases} \hat{x} = 1 & \text{if } y \geq \frac{1}{2}\ln(\frac{q}{1-q}) \\ \hat{x} = -1 & \text{if } y \leq \frac{1}{2}\ln(\frac{q}{1-q}) \end{cases}$
  - $y < -1 \Rightarrow |y+1|-|y-1| = -2 \leq \ln(\frac{q}{1-q}) \Rightarrow \hat{x} = -1$
- ▶ $q > \frac{e^2}{1+e^2}$: $\ln(\frac{q}{1-q}) > 2$
  - $y > 1 \Rightarrow |y+1|-|y-1| = 2 < \ln(\frac{q}{1-q}) \Rightarrow \hat{x} = -1$
  - $-1 \leq y \leq 1 \Rightarrow |y+1|-|y-1| = 2y < 2 < \ln(\frac{q}{1-q}) \Rightarrow \hat{x} = -1$

- $y < -1 \Rightarrow |y+1| - |y-1| = -2 < 0 < \ln(\frac{q}{1-q}) \Rightarrow \hat{x} = -1$

(b) Let H be the optimal decision rule for the channel when q is unknown.

$$P_{e,H} = q\mathbb{P}[e|X=-1] + (1-q)\mathbb{P}[e|X=+1]$$

where $\mathbb{P}[e|X=-1]$ and $\mathbb{P}[e|X=-1]$ are indepenet of q and depend on the characterization of the channel of the statistics of the noise. So the average error probability is a linear function of q and thus takes its maximum in one of its end points.

$$\min_H \max_q P_{e,H} = \max\{\mathbb{P}[e|X=-1], \mathbb{P}[e|X=+1]\}$$

Therefore the best decoder is obtained when $\mathbb{P}[e|X=-1] = \mathbb{P}[e|X=+1]$. So we have

$$\begin{aligned} \mathbb{P}[e|X=-1] = \int_\theta^\infty p(y|x=-1)dy &= \int_{-\infty}^\theta p(y|x=+1)dy = \mathbb{P}[e|X=+1] \\ \int_\theta^\infty \frac{1}{2}e^{-|y+1|}dy &= \int_{-\infty}^\theta \frac{1}{2}e^{-|y-1|}dy \end{aligned} \tag{1}$$

In order to solve the above equation, we can consider three cases:

- $\theta < -1$:

$$\begin{aligned} \int_\theta^{-1} e^{y+1}dy + \int_{-1}^\infty e^{-(y+1)}dy &= \int_{-\infty}^\theta e^{y-1}dy \\ (1-e^{\theta+1}) - (0-1) &= (e^{\theta-1} - 0) \\ e^{\theta-1} - e^{\theta+1} &= 2 \end{aligned}$$

  The last equality is impossible because both of the terms in the LHS are positive and less than 1.

- $\theta > 1$:

$$\begin{aligned} \int_\theta^\infty e^{-(y+1)}dy &= \int_{-\infty}^1 e^{y-1}dy + \int_1^\theta e^{-(y-1)}dy \\ -(0-e^{-(\theta+1)}) &= (1-0) - (e^{-(\theta-1)} - 1) \\ e^{-(\theta+1)} - e^{-(\theta-1)} &= 2 \end{aligned}$$

  Here also because of the same reason, we have a contradiction in the last equality.

- $-1 \le \theta \le 1$: In this case we have

$$\begin{aligned} \int_\theta^\infty e^{-(y+1)}dy &= \int_{-\infty}^\theta e^{y-1}dy \\ e^{-\theta-1} &= e^{\theta-1} \end{aligned}$$

  and so $\theta = 0$ is the only root of (1). (You could also show this by make an argument on the symmerty of the tails of the error function around zero.)

(c) We can write the MAP decoder for this case as

$$\mathrm{P}(Y_1, Y_2|X = +1)\mathbb{P}[X = +1] \underset{x=+1}{\overset{x=-1}{\lessgtr}} \mathrm{P}(Y_1, Y_2|X = -1)\mathbb{P}[X = -1].$$

where $\mathbb{P}[X = -1] = \mathbb{P}[X = 1] = \frac{1}{2}$. We have

$$\begin{aligned} p(y_1, y_2|x) &= p(Z_1 = y_1 - x, Z_2 = y_2 - x|x) \\ &\overset{(a)}{=} p(Z_1 = y_1 - x)p(Z_1 = y_1 - x) \\ &= \frac{1}{4}e^{-|y_1-x|-|y_2-x|} \end{aligned}$$

where (a) follows from the fact that Z_1 and Z_2 are independent from X and each other. So the dicision rule will be

$$e^{-|y_1-1|-|y_2-1|} \underset{x=+1}{\overset{x=-1}{\lessgtr}} e^{-|y_1+1|-|y_2+1|}.$$

Now, consider the following cases and write the dicision rule for each.

- $y_1 > 1, y_2 > 1$:

$$\Longrightarrow \quad e^{-y_1-y_2+2} \underset{x=+1}{\overset{x=-1}{\lessgtr}} e^{-y_1-y_2-2} \quad \Longrightarrow \quad e^2 \underset{x=+1}{\overset{x=-1}{\lessgtr}} e^{-2} \quad \Longrightarrow \quad \hat{x} = 1$$

- $y_1 > 1, -1 \leq y_2 \leq 1$:

$$\Longrightarrow \quad e^{-y_1+y_2} \underset{x=+1}{\overset{x=-1}{\lessgtr}} e^{-y_1-y_2-2} \quad \Longrightarrow \quad e^{-2y_2} \underset{x=+1}{\overset{x=-1}{\lessgtr}} e^{-2} \quad \Longrightarrow \quad \hat{x} = 1$$

- $y_1 > 1, y_2 < -11$:

$$\Longrightarrow \quad e^{-y_1+y_2} \underset{x=+1}{\overset{x=-1}{\lessgtr}} e^{-y_1+y_2} \quad \Longrightarrow \quad 1 \underset{x=+1}{\overset{x=-1}{\lessgtr}} 1 \quad \Longrightarrow \quad \text{either choice}$$

- $-1 \leq y_1 \leq 1, y_2 > 1$:

$$\Longrightarrow \quad e^{y_1-y_2} \underset{x=+1}{\overset{x=-1}{\lessgtr}} e^{-y_1-y_2-2} \quad \Longrightarrow \quad e^{-2y_1} \underset{x=+1}{\overset{x=-1}{\lessgtr}} e^{-2} \quad \Longrightarrow \quad \hat{x} = 1$$

- $-1 \leq y_1 \leq 1, -1 \leq y_2 \leq 1$:

$$\Longrightarrow \quad e^{y_1+y_2-2} \underset{x=+1}{\overset{x=-1}{\lessgtr}} e^{-y_1-y_2-2} \quad \Longrightarrow \quad e^{y_1+y_2} \underset{x=+1}{\overset{x=-1}{\lessgtr}} 1 \quad \Longrightarrow \quad \begin{cases} \hat{x} = 1 & \text{if } y_1 + y_2 \geq 0 \\ \hat{x} = -1 & \text{if } y_1 + y_2 < 0 \end{cases}$$

- $-1 \leq y_1 \leq 1, y_2 < -1$:

$$\Longrightarrow \quad e^{y_1+y_2-2} \underset{x=+1}{\overset{x=-1}{\lessgtr}} e^{-y_1+y_2} \quad \Longrightarrow \quad e^{2y_1} \underset{x=+1}{\overset{x=-1}{\lessgtr}} e^{2} \quad \Longrightarrow \quad \hat{x} = -1$$

- $y_1 < -1, y_2 > 1$:

$$\Longrightarrow \quad e^{y_1-y_2} \underset{x=+1}{\overset{x=-1}{\lessgtr}} e^{y_1-y_2} \quad \Longrightarrow \quad 1 \underset{x=+1}{\overset{x=-1}{\lessgtr}} 1 \quad \Longrightarrow \quad \text{either choice}$$

- $y_1 < -1, -1 \leq y_2 \leq 1$:

$$\Longrightarrow \quad e^{y_1+y_2-2} \underset{x=+1}{\overset{x=-1}{\lessgtr}} e^{y_1-y_2} \quad \Longrightarrow \quad e^{2y_2} \underset{x=+1}{\overset{x=-1}{\lessgtr}} e^{2} \quad \Longrightarrow \quad \hat{x} = -1$$

- $y_1 < -1, y_2 < -1$:

$$\Longrightarrow \quad e^{y_1+y_2-2} \underset{x=+1}{\overset{x=-1}{\lessgtr}} e^{y_1+y_2+2} \quad \Longrightarrow \quad e^{-2} \underset{x=+1}{\overset{x=-1}{\lessgtr}} e^{2} \quad \Longrightarrow \quad \hat{x} = -1$$

Finally, we will have the following decision regions on the $(y_1 - y_2)$-plane.

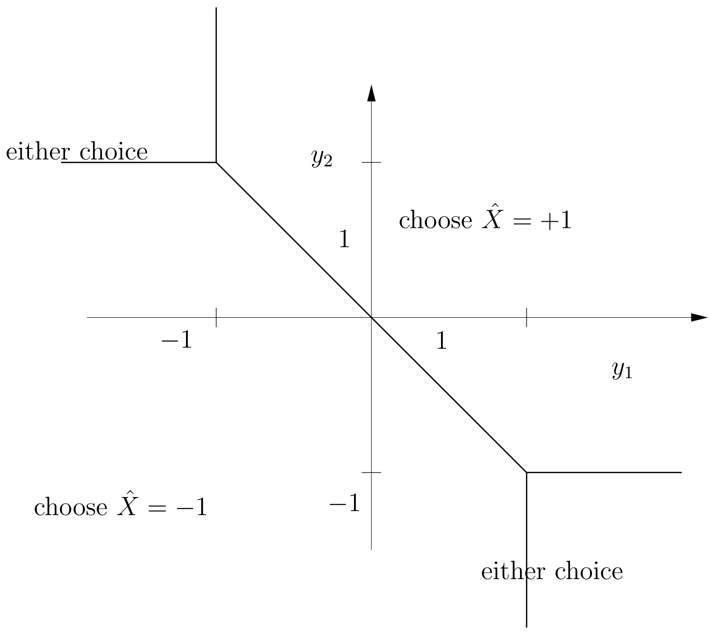

Figure 1: Decision regions.

Problem 3

(a) Note that $x_I(t)$ and $x_Q(t)$ are colored Gaussian processes. The baseband representation is given by

$$x_{bb}(t) = x_I(t) + jx_Q(t)$$

and is very useful for the following reason. If we pass $x_{bb}(t)$ through a linear filter $h(t)$, then the resulting process will be

$$x_{bb}(t) * h(t) = x_I(t) * h(t) + jx_Q(t) * h(t),$$

i.e., the filter will apply independently to the in-phase and to the quadrature-phase component and the resulting process will still be baseband. In addition, we know that $x_I(t)$ and $x_Q(t)$ have the same auto-correlation function $\phi_{X_I}(\tau) = \phi_{X_Q}(\tau) = e^{-2|\tau|}$. We wish to find a process $y_{bb}(t)$ that is baseband and whose in-phase and quadrature-phase components are **white** zero-mean random processes. To find such a process, it suffices to apply a whitening filter to $x_{bb}(t)$, i.e., to find a filter $h(t)$ that is a whitening filter for both $x_I(t)$ and $x_Q(t)$. Note that since we have continuous-time processes, the frequency domain is the Fourier-domain. The condition

$$\mathbb{E}[y_I(t)y_I(t-\tau)] = \frac{1}{2}\delta(\tau)$$

(meaning that $y_I(t)$ should be white) corresponds to

$$S_{y_I}(\omega) = \frac{1}{2}$$

in the frequency-domain. To whiten $x_I(t)$, we first compute its power-spectral density $S_{x_I}(\omega)$:

$$\begin{aligned} S_{x_I}(\omega) &= \mathcal{F}\mathbb{E}[x_I(t)x_I(t-\tau)] \\ &= \mathcal{F}e^{-2|\tau|} \\ &= \frac{1}{(1+\frac{1}{2}j\omega)(1-\frac{1}{2}j\omega)}, \end{aligned}$$

where the last equality comes from the hint on the first page of the exam (note that $\omega = 2\pi f$). If we filter $x_I(t)$ by $h(t)$ to obtain $y_I(t)$, then the output will be

$$y_I(t) = x_I(t) * h(t)$$

and the power spectral densities behave like

$$S_{y_I}(\omega) = |H(\omega)|^2 S_{x_I}(\omega),$$

where $H(\omega)$ is the Fourier transform of $h(t)$. To achieve the desired power spectral density of $y_I(t)$, we set

$$\begin{aligned} |H(\omega)|^2 &= \frac{S_{y_I}(\omega)}{S_{x_I}(\omega)} \\ &= \frac{\frac{1}{2}}{\frac{1}{(1+\frac{1}{2}j\omega)(1-\frac{1}{2}j\omega)}} \\ &= \frac{(1+\frac{1}{2}j\omega)(1-\frac{1}{2}j\omega)}{2} \\ &= H(\omega)H(\omega)^*. \end{aligned}$$

Note that the complex conjugate of $\frac{1}{\sqrt{2}}(1+\frac{1}{2}j\omega)$ is $\frac{1}{\sqrt{2}}(1-\frac{1}{2}j\omega)$. Therefore, we identify the whitening filter

$$H(\omega) = \frac{1}{\sqrt{2}}(1+\frac{1}{2}j\omega).$$

The same whitening filter also works for the quadrature-phase part $x_Q(t)$.

(b) The procedure is as follows. We follow part (a) to get from $x(t)$ to its baseband equivalent $x_{bb}(t)$ and then to the white baseband process $y_{bb}(t)$. Now, we color the in-phase and quadrature-phase parts of $y_{bb}(t)$ to get a baseband process $z_{bb}(t)$. The coloring procedure is the exact opposite of the whitening procedure, i.e., $z_{bb}(t) = y_{bb}(t) * g(t)$, where $g(t)$ is a coloring filter that should be such that

$$\begin{aligned}
|G(\omega)|^2 S_{y_I}(\omega) &= S_{z_I}(\omega) \\
&= \mathcal{F}\mathbb{E}[z_I(t)z_I(t-\tau)] \\
&= \mathcal{F}e^{-3|\tau|} \\
&= \frac{\frac{2}{3}}{(1+\frac{1}{3}j\omega)(1-\frac{1}{3}j\omega)},
\end{aligned}$$

where the last equality again comes from the hint on the first page. Hence,

$$\begin{aligned}
|G(\omega)|^2 &= G(\omega)G(\omega)^* \\
&= \frac{S_{z_I}(\omega)}{S_{y_I}(\omega)} \\
&= \frac{2}{3(1+\frac{1}{3}j\omega)(1-\frac{1}{3}j\omega)\frac{1}{2}},
\end{aligned}$$

and we find

$$G(\omega) = \frac{2}{\sqrt{3}(1+\frac{1}{3}j\omega)}.$$

By filtering $y_{bb}(t)$ with $g(t)$, both the in-phase and the quadrature-phase component of $y_{bb}(t)$ get colored. Now, from the colored baseband process $z_{bb}(t)$, we can construct a passband process

$$z(t) = z_I(t)\cos(\tilde{\omega}_c t) - z_Q(t)\sin(\tilde{\omega}_c t).$$

Problem 4

(a) We have to find $\{a_n\}_{n=1}^{\infty}$ such that $\mathbb{E}[|x_k - \hat{x}_k|^2]$ is minimized. Using the orthogonality principle, the error of estimition should be orthogonal to our observation, *i.e.,*

$$\begin{aligned}
\mathbb{E}[(x_k - \hat{x}_k)y_{k-m}^*] &= 0 \quad m = 1, 2, 3, \ldots \\
\mathbb{E}\left[\left(x_k - \sum_{n=1}^{\infty} a_n y_{k-n}\right) y_{k-m}^*\right] &= 0 \\
\mathbb{E}[x_k y_{k-m}^*] - \sum_{n=1}^{\infty} a_n \mathbb{E}[y_{k-n} y_{k-m}^*] &= 0 \\
r_{xy}(m) &= \sum_{n=1}^{\infty} a_n r_y(m-n) \qquad (2)
\end{aligned}$$

According to the channel model, we can relate all the auto-correlations to $S_x(D)$ and $S_z(D)$:

$$\begin{aligned}
r_y(\ell) &= \mathbb{E}[y_k y_{k-\ell}^*] \\
&= \mathbb{E}[(x_k + z_k)(x_{k-\ell}^* + z_{k-\ell}^*)] \\
&= \mathbb{E}[x_k x_{k-\ell}^*] + \mathbb{E}[z_k z_{k-\ell}^*] \\
&= r_x(\ell) + \delta(\ell) \quad \text{(note that } z_k \text{ and } z_{k-\ell}^* \text{ are independent unless for } \ell = 0) \qquad (3)
\end{aligned}$$

$$\begin{aligned}
r_{xy}(\ell) &= \mathbb{E}[x_k y_{k-\ell}^*] \\
&= \mathbb{E}[x_k(x_{k-\ell}^* + z_{k-\ell}^*)] \\
&= \mathbb{E}[x_k x_{k-\ell}^*] = r_x(\ell) \qquad (4)
\end{aligned}$$

Replacing (3) and (4) in (2), we have

$$\begin{aligned}
r_x(m) &= \sum_{n=1}^{\infty} a_n \left[r_x(m-n) + \delta(m-n) \right] \quad m = 1, 2, 3, \ldots \\
&= \sum_{n=1}^{\infty} a_n r_x(m-n) + a_m \\
g_m &= \sum_{n=1}^{\infty} a_n' r_x(m-n) + a_m' = 0 \quad m = 1, 2, 3, \ldots
\end{aligned}$$

where $a_0' = 1$ and $a_n' = -a_n$ for $n > 0$. Note that $g_m = 0$ for $m = 1, 2, 3, \ldots$, thus it is an anti-causal sequence. In the other hand, we know that by definition a_n' is a causal sequence. Going to the $\mathcal{D}$ domain we have

$$\begin{aligned}
G(D) &= A'(D)S_x(D) + A'(D) \\
&= A'(D)(S_x(D) + 1) \\
&= A'(D)\Gamma L(D) L^*(D^{-*})
\end{aligned}$$

where $\Gamma L(D) L^*(D^{-*})$ is the spectral factorization of $S_x(D) + 1$. So we have

$$\frac{G(D)}{\Gamma L^*(D^{-*})} = A'(D)L(D)$$

where the LHS and RHS are respectively anti-causal and causal sequences and thus both should be constant. Accoding to the monicity of $A'(D)$ and $L(D)$, we have

$$\begin{aligned}
A'(D)L(D) &= 1 \\
\Rightarrow A'(D) &= \frac{1}{L(D)} \\
\Rightarrow \{a_n\}_{n=1}^{\infty} &= \{-a_n'\}_{n=1}^{\infty} = -\mathcal{D}^{-1}\left\{ \frac{1}{L(D)} \right\}
\end{aligned}$$

(b)

$$\begin{aligned}
r_x(\ell) &= e^{-2|\ell|} \\
\Rightarrow S_x(D) &= \mathcal{D}\{r_x(\ell)\} \\
&= \sum_{\ell=0}^{\infty} r_x(\ell)D^\ell + \sum_{\ell=-\infty}^{-1} r_x(\ell)D^\ell \\
&= \sum_{\ell=0}^{\infty} e^{-2\ell}D^\ell + \sum_{\ell=-\infty}^{-1} e^{2\ell}D^\ell \\
&= \frac{1}{1-\frac{D}{e^2}} + \frac{\frac{1}{e^2 D}}{1-\frac{1}{e^2 D}} \\
&= \frac{1-\frac{1}{e^4}}{\left(1-\frac{D}{e^2}\right)\left(1-\frac{1}{e^2 D}\right)}
\end{aligned}$$

and

$$\begin{aligned}
S_x(D)+1 &= \frac{1-\frac{1}{e^4}}{\left(1-\frac{D}{e^2}\right)\left(1-\frac{1}{e^2 D}\right)} + 1 \\
&= \frac{-\frac{D}{e^2}+2(1+\frac{1}{e^4})-\frac{1}{e^2 D}}{\left(1-\frac{D}{e^2}\right)\left(1-\frac{1}{e^2 D}\right)} \\
&= \Gamma \frac{(1-\beta D)(1-\beta D^{-1})}{\left(1-\frac{1}{e^2}D\right)\left(1-\frac{1}{e^2}D^{-1}\right)}
\end{aligned}$$

where $\Gamma = 1+\frac{1}{e^4}+\sqrt{1+\frac{1}{e^4}+\frac{1}{e^8}} \simeq 2.0276$ and $\beta = e^2+\frac{1}{e^2}-\sqrt{e^4+1+\frac{1}{e^4}} \simeq 0.0667$. So, we have

$$L(D) = \frac{1-\beta D}{1-\frac{1}{e^2}D}$$

and

$$\begin{aligned}
A'(D) &= \frac{1}{L(D)} = \frac{1-\frac{1}{e^2}D}{1-\beta D} \\
&= (1-\frac{1}{e^2}D)(1+\beta D+\beta^2 D^2+\beta^3 D^3+\cdots) \\
&= 1+(\beta-\frac{1}{e^2})D+\beta(\beta-\frac{1}{e^2})D^2+\cdots
\end{aligned}$$

in general for $n \geq 1$ we have

$$a'_n = \beta^{n-1}(\beta - e^{-2}).$$

Finally, we have

$$a_n = \beta^{n-1}(e^{-2}-\beta) \qquad n = 1,2,3,\ldots.$$

Problem 5

(a) Bob receives

$$y_B(t) = \sum_n x[n]\phi(t-nT) + z_B(t),$$

and therefore, projecting $y_B(t)$ onto $\phi(t-kT)$ is a sufficient statistics for detecting $x[k]$. This projection (matched filter) yields

$$\begin{aligned} y_B[k] &= y_B(t) * \phi(-t)^*\big|_{t=kT} \\ &= \int_0^T y_B(t)\phi(t)^* dt \\ &= x[k] + \int_0^T z_B(t)\phi(t)^* dt. \end{aligned}$$

Thus, $z_B[k]$ is equal to the last integral. To find its power spectral density, we first compute its auto-correlation function.

$$\begin{aligned} r_{z_B}[m] &= \mathbb{E}\left[z_B[k] z_B[k-m]^*\right] \\ &= \mathbb{E}\left[\int_{t=0}^T z_B(t)\phi(t-kT)^* dt \int_{s=0}^T z_B(s)^*\phi\big(s-(k-m)T\big)ds\right] \\ &= \int_{t=0}^T \int_{s=0}^T \underbrace{\mathbb{E}\left[z_B(t) z_B(s)^*\right]}_{N_0\delta(t-s)} \phi(t-kT)^*\phi\big(s-(k-m)T\big)dsdt \\ &= \int_{t=0}^T N_0\phi(t-kT)^*\phi\big(t-(k-m)T\big)dt \\ &= N_0\delta_m, \end{aligned}$$

where the last equality follows from the fact that $\{\phi(t-nT)\}_{n\in\mathbb{Z}}$ is an orthonormal basis. The power spectral density is the D-transform of $r_{z_B}[k]$, which is

$$S_{z_B}(D) = N_0.$$

(b) Carol receives

$$\begin{aligned} y_C(t) &= \sum_n x[n]\phi(t-nT) + \sum_n y_B[n]\phi\big(t-(n+1)T\big) + z_C(t) \\ &= \sum_n x[n]\Big(\phi(t-nT) + \phi\big(t-(n+1)T\big)\Big) + \sum_n z_B[n]\phi\big(t-(n+1)T\big) + z_C(t). \end{aligned}$$

Note that the expression $\phi(t-nT)+\phi\big(t-(n+1)T\big)$ is exactly as in an inter-symbol interference (ISI) channel, since

$$\begin{aligned} \phi(t) + \phi(t-T) &= \phi(t) * h(t) \\ &\triangleq p(t), \end{aligned}$$

where $h(t) = \delta(t) + \delta(t-T)$. We therefore proceed as for an ISI channel. Define

$$\begin{aligned} \tilde{\phi}(t) &= \frac{p(t)}{||p(t)||} \\ &= \frac{1}{\sqrt{2}}\big(\phi(t) + \phi(t-T)\big), \end{aligned}$$

because $||p(t)|| = \langle p(t), p(t), \rangle^{\frac{1}{2}} = \sqrt{2}$. We know that in an ISI channel, we should use $\tilde{\phi}(t)$ as a basis function for the matched filter. If we just used $\phi(t)$, we would **not** end up with **a sufficient statistics**. Projecting $y_C(t)$ onto the kth basis function yields

$$
\begin{aligned}
y_C[k] &= \int_0^T y_C(t)\tilde{\phi}(t-kT)^* dt \\
&= \frac{1}{\sqrt{2}}\sum_n x[n]\int_0^T \big(\phi(t-nT)+\phi\big(t-(n+1)T\big)\big)\big(\phi(t-kT)^*+\phi\big(t-(k+1)T\big)^*\big)dt \\
&\quad + \frac{1}{\sqrt{2}}\sum_n z_B[n]\int_0^T \phi\big(t-(n+1)T\big)\big(\phi(t-kT)^*+\phi\big(t-(k+1)T\big)^*\big)dt \\
&\quad + \frac{1}{\sqrt{2}}\int_0^T z_C(t)\big(\phi(t-kT)^*+\phi\big(t-(k+1)T\big)^*\big)dt \\
&= \frac{2}{\sqrt{2}}x[k]+\frac{1}{\sqrt{2}}x[k-1]+\frac{1}{\sqrt{2}}x[k+1] \\
&\quad + \frac{1}{\sqrt{2}}\big(z_B[k-1]+z_B[k]\big) \\
&\quad + z_C[k] \\
&= \sqrt{2}q[k]*x[k]+\tilde{z}_B[k]+z_C[k],
\end{aligned}
$$

where we defined $z_C[k]=\frac{1}{\sqrt{2}}\int_0^T z_C(t)\big(\phi(t-kT)^*+\phi\big(t-(k+1)T\big)^*\big)dt$, and where

$$
q[k]=\int_0^T \tilde{\phi}(t)\tilde{\phi}(t-kT)=\delta_k+\frac{1}{2}\delta_{k-1}+\frac{1}{2}\delta_{k+1}.
$$

One can easily verify that $z_C[k]$ has autocorrelation-function

$$
\begin{aligned}
r_{z_C}[m] &= \mathbb{E}\left[z_C[k]z_C[k-m]^*\right] \\
&= \frac{1}{2}N_0(2\delta_m+\delta_{m-1}+\delta_{m+1}) \\
&= N_0 q[m].
\end{aligned}
$$

The first noise term is defined as $\tilde{z}_B[k]=\frac{1}{\sqrt{2}}\big(z_B[k-1]+z_B[k]\big)$. Remember that $z_B[k]$ has autocorrelation function $r_{z_B}[m]=N_0\delta_m$. One can again easily verify that $r_{\tilde{z}_B}[m]=N_0q[m]$. Hence

$$
S_{z_C}(D)=S_{\tilde{z}_B}(D)=N_0Q(D),
$$

i.e., the noise is no longer white.

(c) It suffices to find the D-transform $W(D)$ of the optimal filter $w[k]$. We write the orthogonality principle in the D-notation:

$$
\mathbb{E}\left[\big(X(D)-W(D)Y_C(D)\big)Y_C^*(D^{-*})\right]=S_{xy_C}(D)-W(D)S_{y_C}(D)=0.
$$

Hence,

$$
\begin{aligned}
W(D) &= \frac{S_{xy_C}(D)}{S_{y_C}(D)} \\
&= \frac{\sqrt{2}Q(D)\mathcal{E}_x}{2Q^2(D)\mathcal{E}_x+N_0Q(D)+N_0Q(D)}.
\end{aligned}
$$

Problem 6

(a) From

$$Q(D) + \frac{1}{SNR_{MFB}} = \gamma_0 G(D) G^*(D^{-*}) \tag{5}$$

we know that the constant terms (the coefficient of D^0) are the same in both sides. We also know $q_0 = 1$, and so the constant term in LHS of (5) is $1 + 1/SNR_{MFB}$. $G(D)$ is causal and can be written as $G(D) = g_0 + g_1 D + g_2 D^2 + g_3 D^3 + \cdots = \sum_{n=0}^{\infty} g_n D^n$. Thus, we also can write $G^*(D^{-*})$ as

$$G^*(D^{-*}) = \left(\sum_{n=0}^{\infty} g_n (D^{-*})^n \right)^* = \sum_{n=0}^{\infty} g_n^* D^{-n} = \sum_{-\infty}^{0} g_n' D^n$$

where $g_n' = g_{-n}^*$. Therefore the constant term in $\gamma_0 G(D) G^*(D^{-*}) \triangleq A(D) = \sum_{n=-\infty}^{\infty} a_n D^n$ is

$$a_0 = \gamma_0 \sum_{n=0}^{\infty} g_n g_{-n}' = \gamma_0 \sum_{n=0}^{\infty} g_n g_n^* = \gamma_0 \sum_{n=0}^{\infty} |g_n|^2 = \gamma_0 \parallel g \parallel^2 .$$

So,

$$1 + \frac{1}{SNR_{MFB}} = \gamma_0 \parallel g \parallel^2$$

(b) We know that $G(D)$ is *monic*, and so $g_0 = 1$. Thus,

$$\parallel g \parallel^2 = \sum_{n=0}^{\infty} |g_n|^2 = |g_0|^2 + \sum_{n=1}^{\infty} |g_n|^2 \geq 1 \tag{6}$$

where the inequality follows from the fact that all the terms in the summation are non-negative. In order to obtain the equality condition,

(i) assume that $\parallel g \parallel^2 = 1$. From (6), we see that all the terms in the summation should be zero, *i.e.*, $g_n = 0$ for $n > 0$. and therefore $G(D) = 1$ and $G^*(D^{-*}) = 1^* = 1$. So, the RHS of (5) is a constant and the LHS is, *i.e.*, $Q(D)$ is a constant polynomial. From the fact $q_0 = 1$, there is no choice for $Q(D)$ unless $Q(D) = 1$.

(ii) assume that $Q(D) = 1$, *i.e.*, the LHS of (5) is a constant.and has no zero or pole and its spectral factorization is just some constant. Assuming $G(D)$ is monic, we have $G(D) = 1$, and therefore $\parallel g \parallel^2 = 1$.

Combining $\parallel g \parallel^2 \geq 1$ with the fact $1 + \frac{1}{SNR_{MFB}} = \gamma_0 \parallel g \parallel^2$, we have

$$\gamma_0 \leq 1 + \frac{1}{SNR_{MFB}}.$$

(c) $G(D)$ is monic and causal, so is $L(D) = \frac{1}{G(D)}$. $L^*(D^{-*}) = \frac{1}{G^*(D^{-*})}$ is also monic and anti-causal, because $G^*(D^{-*})$ is monic and anti-causal. So, the constant term in the inverse $\mathcal{D}$-transform of $L(D)L^*(D^{-*})$ is 1. Thus,

$$\begin{aligned}
\mathcal{D}^{-1}\left[\frac{1}{Q(D)+\frac{1}{SNR_{MFB}}}\right]\Bigg|_0 &= \mathcal{D}^{-1}\left[\beta L(D)L^*(D^{-*})\right]\Big|_0 \\
&= \beta\mathcal{D}^{-1}\left[L(D)L^*(D^{-*})\right]\Big|_0 \\
&= \frac{1}{\gamma_0}\sum_{n=0}^{\infty}|\ell_n|^2 \\
&= \frac{1}{\gamma_0}\left(1+\sum_{n=1}^{\infty}|\ell_n|^2\right) \\
&\geq \frac{1}{\gamma_0}
\end{aligned}$$

(d) We know that the signal energy, $\mathcal{E}_x$, is the same in MMSE-LE and MMSE-DFE. So we just have to comprise the noise energies in the equlizers.

$$\begin{aligned}
\sigma^2_{MMSE-DFE} &= \frac{N_0}{\| p \|^2}\frac{1}{\gamma_0} \\
&\leq \frac{N_0}{\| p \|^2}\mathcal{D}^{-1}\left[\frac{1}{Q(D)+\frac{1}{SNR_{MFB}}}\right]\Bigg|_0 \qquad (7)\\
&= \sigma^2_{MMSE-LE}
\end{aligned}$$

So,

$$SNR_{MMSE-LE} = \frac{\mathcal{E}_x}{\sigma^2_{MMSE-LE}} \leq \frac{\mathcal{E}_x}{\sigma^2_{MMSE-DFE}} = SNR_{MMSE-DFE}$$

and

$$SNR_{MMSE-LE,U} = SNR_{MMSE-LE} - 1 \leq SNR_{MMSE-LE} - 1 = SNR_{MMSE-LE}.$$

Note that the inequality in (7) follows from the inequality of part (c) with the quality condition $\ell_n = 0$ for $n > 0$. This condition is equivalent to $L(D) = 1$, which is also equivalent to $G(D) = 1$ and so $Q(D) = 1$.

(e) The first inequality is already proved in (d) and we just have to show

$$SNR_{MMSE-DFE,U} \leq SNR_{MFB}:$$

$$\begin{aligned}
SNR_{MMSE-DFE,U} &= SNR_{MMSE-DFE} - 1 = \frac{\mathcal{E}_x}{\sigma^2_{MMSE-DFE}} - 1 \\
&\overset{(*)}{=} \frac{\mathcal{E}_x}{\frac{N_0}{\|p\|^2}}\gamma_0 - 1 \\
&\overset{(**)}{\leq} SNR_{MFB}\cdot\left(1+\frac{1}{SNR_{MFB}}\right) - 1 \\
&= SNR_{MFB}
\end{aligned}$$

where $(*)$ and $(**)$ follow from (d) and (b), respectively.