ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
School of Computer and Communication Sciences

Handout 11 Advanced Digital Communications
Solutions to Homework 4 Oct. 24, 2016

Solution 1.

(a) An i.i.d. random process implies that for an arbitrary set $\{n_1, \ldots, n_K\}$,

$$\begin{aligned} p_{X[n_1]\ldots X[n_K]}(x[n_1], \ldots x[n_K]) &= p_{X[n_1]}(x[n_1]) \ldots p_{X[n_k]}(x[n_K]) \\ p_{X[n_i]} &= p_{X[n_j]}, \quad i, j = 1 \ldots, K. \end{aligned}$$

Then, we have for any integer M,

$$\begin{aligned} p_{X[n_1+M]\ldots X[n_K+M]}(x[n_1+M], \ldots, x[n_K+M]) &\overset{(*)}{=} \prod_{i=n_1}^{n_K} p_{X[i+M]}(x) \\ &\overset{(\dagger)}{=} \prod_{i=n_1}^{n_K} p_{X[i]}(x) \\ &\overset{(\#)}{=} p_{X[n_1]\ldots X[n_K]}(x[n_1], \ldots, x[n_K]), \end{aligned}$$

where $(*)$ and $(\#)$ follow by independence and $(\dagger)$ follows by identical distribution. Thus, every i.i.d. random process is stationary.

Moreover, it is clear that the mean $m_{X[n]} = \mathbb{E}[X[n]]$ is constant 0 for any n. As for the autocorrelation function, we have

$$\begin{aligned} R_X[n, m] &= \mathbb{E}[X[n]X[m]] \\ &= \begin{cases} \mathbb{E}[X[n]]\,\mathbb{E}[X[m]] = 0 & \text{if } m \neq n \\ \mathbb{E}[X^2[n]] = 1 & \text{if } m = n \end{cases} \\ &= \delta[m-n] \end{aligned}$$

where the *Kronecker delta* function $\delta[i]$ equals 1 if $i = 0$ and 0 otherwise. Therefore, the i.i.d. random process is wide-sense stationary and its power spectral density is given by

$$\begin{aligned} S_X(f) &= \sum_{k=-\infty}^{\infty} R_X[k] e^{-j2\pi fk} \\ &= \sum_{k=-\infty}^{\infty} \delta[k] e^{-j2\pi fk} \\ &= 1. \end{aligned}$$

(b) The mean of $Y[n]$ is

$$\mathbb{E}[Y[n]] = \sum_{k=0}^{\infty} (1/2)^k\, \mathbb{E}[X[n-k]] = \sum_{k=0}^{\infty} (1/2)^k \times 0 = 0.$$

The autocorrelation function of $Y[n]_n$ is

$$\begin{aligned} R_Y[n,m] &= \mathbb{E}[Y[n]Y[m]] \\ &= \mathbb{E}\left[\left(\sum_{k=0}^{\infty}(1/2)^k X[n-k]\right)\left(\sum_{l=0}^{\infty}(1/2)^l X[m-l]\right)\right] \\ &= \mathbb{E}\left[\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}(1/2)^{k+l}X[n-k]X[m-l]\right] \\ &= \sum_{k=0}^{\infty}\sum_{l=0}^{\infty}(1/2)^{k+l}R_X[m-n-l+k]. \end{aligned}$$

It can, hence, be seen from here that the process $Y[n]$ is wide-sense stationary as $R_Y[n,m]$ only depends on the difference $m-n$. Using the result from (a), we can simplify this expression as follows

$$R_Y[n,m] = \sum_{k=0}^{\infty}\sum_{l=0}^{\infty}(1/2)^{k+l}\delta[m-n-l+k].$$

We first consider the case $n \geq m$:

$$\begin{aligned} R_Y[n,m] &\overset{(*)}{=} \sum_{k=n-m}^{\infty}(1/2)^{k+(m-n+k)} \\ &= (1/2)^{m-n}\sum_{k=n-m}^{\infty}(1/2)^{2k} \\ &= (1/2)^{m-n}\frac{(1/2)^{2(n-m)}}{1-1/4} \\ &= \frac{4}{3}(1/2)^{n-m}, \end{aligned}$$

where $(*)$ follows from the restriction that $\delta(m-n-l+k)$ is non-zero only if $l = m-n+k$, and the fact that l is non-negative. In a similar manner we can conclude $R_Y[n,m] = \frac{4}{3}(1/2)^{m-n}$ if $m > n$. The expression can be more simply written as

$$R_Y[n,m] = \frac{4}{3}(1/2)^{|n-m|},$$

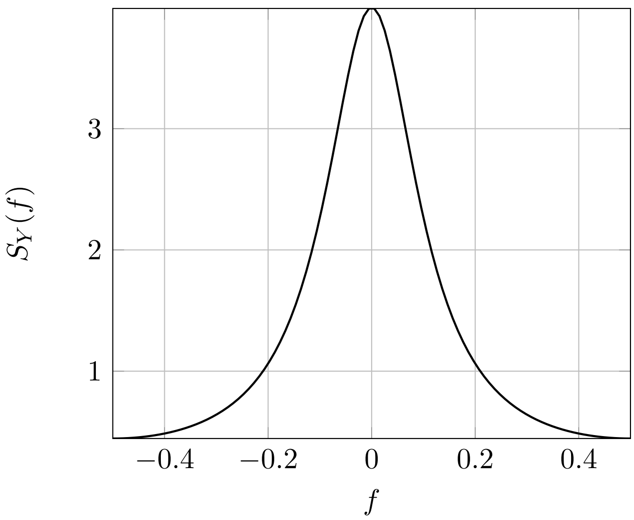
or equivalently,

$$R_Y[k] = \frac{4}{3}(1/2)^{|k|}.$$

The power spectral density of $Y[n]$ is

$$\begin{aligned}
S_Y(f) &= \sum_{k=-\infty}^{\infty} R_Y[k] e^{-j2\pi fk} \\
&= (4/3) \sum_{k=-\infty}^{\infty} (1/2)^{|k|} e^{-j2\pi fk} \\
&= 4/3 \left(-1 + \sum_{k=0}^{\infty} (1/2)^k e^{-j2\pi fk} + \sum_{k=-\infty}^{0} (1/2)^{-k} e^{-j2\pi fk} \right) \\
&= 4/3 \left(-1 + \sum_{k=0}^{\infty} (1/2)^k e^{-j2\pi fk} + \sum_{k=0}^{\infty} (1/2)^k e^{j2\pi fk} \right) \\
&= 4/3 \left(-1 + \frac{1}{1 - \frac{e^{-j2\pi f}}{2}} + \frac{1}{1 - \frac{e^{j2\pi f}}{2}} \right) \\
&= 4/3 \left(-1 + \frac{2 - (1/2)(e^{j2\pi f} + e^{-j2\pi f})}{1 + 1/4 - (1/2)(e^{j2\pi f} + e^{-j2\pi f})} \right) \\
&= 4/3 \left(-1 + \frac{2 - \cos(2\pi f)}{5/4 - \cos(2\pi f)} \right) \\
&= 4/3 \left(\frac{3/4}{5/4 - \cos(2\pi f)} \right) \\
&= \frac{1}{5/4 - \cos(2\pi f)}.
\end{aligned}$$

The power spectral density $S_Y(f)$ is plotted in the below figure. Originally, the power spectrum of the i.i.d. random process is flat. After passing through the low-pass filter, the low-frequency part is preserved, whereas the high-frequency part is attenuated.



(c) The mean of $Y[n]$ is

$$\mathbb{E}[Y[n]] = \mathbb{E}[X[n]] + \frac{1}{3}\,\mathbb{E}[X[n-1]] + \mathbb{E}[Z[n]] = 0.$$

The autocorrelation function of $Y[n]$ is

$$\begin{aligned}
R_Y[n,m] &= \mathbb{E}[Y[n]Y[m]] \\
&= \mathbb{E}\left[\left(X[n] + \frac{1}{3}X[n-1] + Z[n]\right)\left(X[m] + \frac{1}{3}X[m-1] + Z[m]\right)\right] \\
&= R_X[m-n] + \frac{1}{3}R_X[m-n-1] + \frac{1}{3}R_X[m-n+1] \\
&\quad + \frac{1}{9}R_X[m-n] + R_Z[m-n],
\end{aligned}$$

where we used the independence of $\{X[n]\}_n$ and $\{Z[n]\}_n$. Since $\mathbb{E}[Y[n]]$ is a constant and $R_Y[n,m]$ depends only on the difference $m-n$, the random process $\{Y[n]\}_n$ is also wide-sense stationary. Using the result from (a), we can conclude

$$R_Y[k] = \begin{cases} \frac{10}{9} + \sigma^2 & \text{if } k = 0, \\ \frac{1}{3} & \text{if } k = \pm 1, \\ 0 & \text{otherwise.} \end{cases}$$

The power spectral density of $\{Y[n]\}_n$ is

$$\begin{aligned}
S_Y(f) &= \sum_{k=-\infty}^{\infty} R_Y[k]e^{-j2\pi fk} \\
&= \frac{10}{9} + \sigma^2 + \frac{2}{3}\cos(2\pi f).
\end{aligned}$$

and

$$\begin{aligned}
S_Y(z) &= \sum_{k=-\infty}^{\infty} R_Y[k]z^{-k} \\
&= \frac{10}{9} + \sigma^2 + \frac{1}{3}(z + z^{-1}).
\end{aligned}$$

Solution 2.

(a) For the ease of notation, each entry in the table has been divided by $2\sqrt{\mathcal{E}}/3$.

b_0	b_1	b_2	$\tilde{I}[0]$	$\tilde{I}[1]$	$\tilde{I}[2]$	$\tilde{I}[3]$	$(\times 2\sqrt{\mathcal{E}}/3)$
0	0	0	-2	-2	-2	-2	
0	0	1	-2	-2	1	-1	
0	1	0	-2	1	-1	-2	
0	1	1	-2	1	2	-1	
1	0	0	1	-1	-2	-2	
1	0	1	1	-1	1	-1	
1	1	0	1	2	-1	-2	
1	1	1	1	2	2	-1	

(b) The minimum distance of the constellation is $2\sqrt{10\mathcal{E}}/3$ and the average error probability can be upper bounded as

$$\begin{aligned} P_e &< (M-1)Q\left(\frac{d_{min}}{2\sqrt{N_0/2}}\right) \\ &= (8-1)Q\left(\frac{2\sqrt{10\mathcal{E}/3}}{2\sqrt{N_0/2}}\right) \\ &= 7Q\left(\sqrt{\frac{20\mathcal{E}}{9N_0}}\right). \end{aligned}$$

(c) We know that the zero forcing filter is essentially the inverse of the channel response. Since the channel response (in z-domain) is

$$1+\frac{1}{3}z^{-1}$$

we find

$$D_{ZF}(z) = \frac{3}{3+z^{-1}}.$$

The power spectral density of the noise after filtering $S_{\tilde{V}}(z)$ is

$$\begin{aligned} S_{\tilde{V}}(z) &= \frac{N_0}{2}D_{ZF}(z)D_{ZF}(z^{-1}) \\ &= \frac{N_0}{2}\frac{9}{10+3(z+z^{-1})} \end{aligned}$$

The power of the noise is

$$\begin{aligned} \mathbb{E}[|\tilde{V}|^2] &= \int_{-1/2}^{1/2} S_{\tilde{V}}(e^{j2\pi f})df \\ &= \frac{N_0}{2}\int_{-1/2}^{1/2}\frac{9}{10+6\cos(2\pi f)}df \\ &= \frac{9}{16}N_0 \end{aligned}$$

(d) The error probability of estimation between $-\sqrt{\mathcal{E}}$ and $\sqrt{\mathcal{E}}$ under the above noise is

$$P_s = Q\left(\sqrt{\frac{16\mathcal{E}}{9N_0}}\right)$$

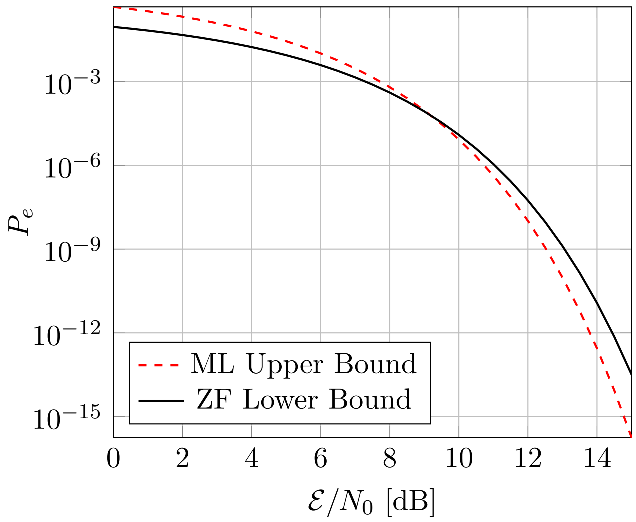
We cannot find the exact error probability of the estimation of all three symbols, so instead we will use our result as a lower bound:

$$P_{eZF} \geq P_s$$

In (b) we found that the union bound on the error probability of ML estimation is

$$P_{eML} \leq 7Q\left(\sqrt{\frac{20\mathcal{E}}{9N_0}}\right) =: P_{U,ML}$$

The plot below shows a comparison of these two bounds as a function of $\frac{\mathcal{E}}{N_0}$; we can see that once the $\frac{\mathcal{E}}{N_0}$ is larger than 9dB, our bounds indeed confirm that the ML is better than the ZF. Can we gain some more insight into how much better it is? To this end, let us consider the high-SNR case, i.e., when $\frac{\mathcal{E}}{N_0}$ becomes large. In this case, the factor 7 in the ML-bound no longer matters. Now, if we remove the factor 7 from the ML bound, we can gain further insight: let us denote the $\frac{\mathcal{E}}{N_0}$ in the ZF case by SNR_{ZF}, and in the ML case by SNR_{ML}. Then, to make the two error probability bounds the same, we have to pick $\mathrm{SNR}_{ZF} = \frac{20}{16}\mathrm{SNR}_{ML}$. That is, the ZF needs at least $10\log_{10}(20/16) \approx 0.9691$dB more transmit power than the ML. In reality, we should expect the ZF to perform strictly worse than the bound we have plotted, so the loss should be expected to be more than a dB.



SOLUTION 3.

(a) Let's use $'$ and $''$ notation to *split* any length-$2N$ vector to its first and second halves (each of length N). Namely, $\boldsymbol{x} = [\boldsymbol{x}', \boldsymbol{x}'']^T$ where $\boldsymbol{x}'$ and $\boldsymbol{x}''$ are length-N vectors. Then we can write the conditional pdf of the output $\boldsymbol{Y}$ given the input $\boldsymbol{x}$ as

$$f(\boldsymbol{y}|\boldsymbol{x}) = \frac{1}{(2\pi)^{N/2}\sigma_0^N} e^{-\frac{1}{2\sigma_0^2}\|\boldsymbol{y}'-\boldsymbol{x}'\|^2} \frac{1}{(2\pi)^{N/2}\sigma_1^N} e^{-\frac{1}{2\sigma_1^2}\|\boldsymbol{y}''-\boldsymbol{x}''\|^2}.$$

By definition, the ML estimation of the sent message is

$$\arg\max_m \frac{1}{(2\pi)^{N/2}\sigma_0^N} e^{-\frac{1}{2\sigma_0^2}\|\boldsymbol{y}'-\boldsymbol{x}_m'\|^2} \frac{1}{(2\pi)^{N/2}\sigma_1^N} e^{-\frac{1}{2\sigma_1^2}\|\boldsymbol{y}''-\boldsymbol{x}_m''\|^2},$$

which is equivalent to

$$\arg\min_m \left(\frac{1}{\sigma_0^2}\|\boldsymbol{y}' - \boldsymbol{x}_m'\|^2 + \frac{1}{\sigma_1^2}\|\boldsymbol{y}'' - \boldsymbol{x}_m''\|^2 \right).$$

Finally, we can express it also as

$$\arg\min_m \left(\|\boldsymbol{y}' - \boldsymbol{x}_m'\|^2 + \frac{\sigma_0^2}{\sigma_1^2}\|\boldsymbol{y}'' - \boldsymbol{x}_m''\|^2 \right).$$

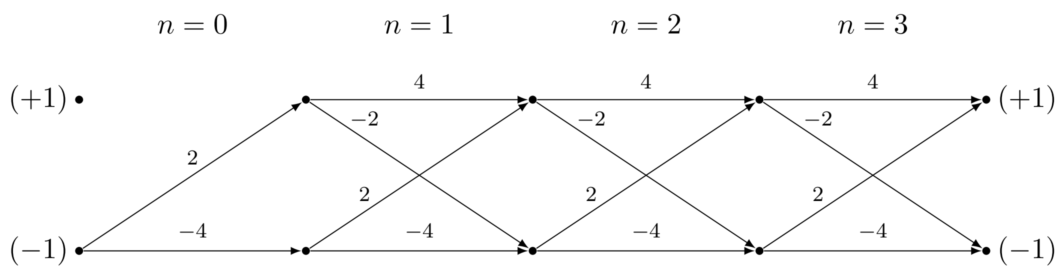
This has a nice intuition: Instead of minimizing the squared distance, we minimize a *weighted* version: On the first half of the vector, we take just squared distances, but on the second half, we multiply the squared distance by $\frac{\sigma_0^2}{\sigma_1^2}$.

For the special case where $\sigma_0^2 = 2$ and $\sigma_1^2 = 1$, we can express the ML detector as

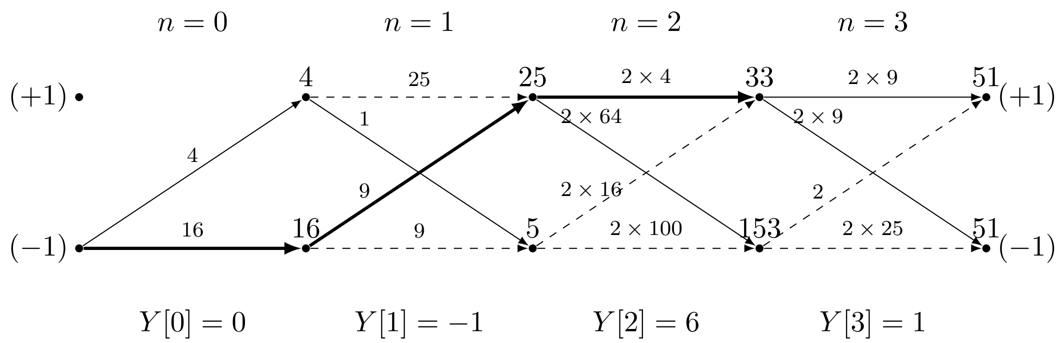
$$\arg\min_m \left(\|\boldsymbol{y}' - \boldsymbol{x}'_m\|^2 + 2\|\boldsymbol{y}'' - \boldsymbol{x}''_m\|^2 \right), \tag{1}$$

which is intuitive: Since the noise in the second part is only half as strong as in the first part, the second part is much more reliable, and hence, should have more impact on our decision. This is precisely what the formula says: The (squared) distance accumulated in the second part is *twice as important* for the decision than the (squared) distance accumulated in the first part.

(b) First, we simply draw the trellis. Recall that this is merely an efficient and convenient way to "list" all possible noiseless received sequences.



The whole point is now to find the one path through this trellis that is the ML path with respect to the received sequence. If all noises have the same variance, this is simply the one path that is closest (in Euclidean distance) to the received sequence. What to do when the noise variances are not all the same? We use part (a), Equation (1), to realize that for ML decoding, as we accumulate squared distances along the trellis, for the first two time instants, we should count the squared distance only *once*, but for times $n = 2, 3, 4, 5, \ldots$ we should count the squared distance *twice* (i.e., multiply it by two). Hence, for the usual trellis search, this looks as follows:



At this point, we observe that the optimal path **must** pass through the third node in the top row (the vertex with label "33") — the corresponding vertex in the bottom row (with label "153") is dead since it has no outgoing edges anymore! Note that if this were *not* the case, I would have to continue drawing the trellis. But this allows us to conclude that the ML path **must include** the segment from the root to the vertex labeled "16", then to "25", then to "33". Hence, the ML decision on the first three bits is $-1, 1, 1$.

(c) Looking at the trellis, we can see that the minimum distance between any two paths through the trellis is $d_{min}^2 = 36$ (for example consider the transmitted signal $(1, 1, 1)$ and

$(1,1,-1)$) and the variance of the noise satisfies $\sigma^2 \leq 2$. Hence, the error probability is upper bounded by $P_e \leq 7Q(d_{min}/(2\sqrt{\sigma^2})) \leq 7Q(6/(2\sqrt{2})) \approx 0.1186$.. However if we consider the "terminated-trellis", then the minimum distance is $d_{min} = \sqrt{6^2+2^2} = \sqrt{40}$, and the variance of the noise satisfies $\sigma^2 \leq 2$. Hence, the error probability is upper bounded by $P_e \leq 7Q(d_{min}/(2\sqrt{\sigma^2})) \leq 7Q(\sqrt{40}/(2\sqrt{2})) = 7Q(\sqrt{5}) \approx 0.0887$. That is, our decoded bits in part (b) are all correct with probability a little over 91 percent!

Solution 4.

(a) In the Z-domain, we could conveniently analyze what will happen to our channel:

$$\begin{aligned} U'_{ZF}(z) &= D'_{ZF}(z)U(z) \\ &= D'_{ZF}(z)\left(D(z)I(z)+V(z)\right) \\ &= (1-\beta z)(1-\beta z^{-1})I(z) + (1-\alpha z)(1-\alpha z^{-1})V(z) \\ &= (1+\beta^2-\beta(z+z^{-1}))I(z) + (1+\alpha^2-\alpha(z+z^{-1}))V(z). \end{aligned}$$

In the discrete-time domain, this would be

$$\begin{aligned} U'_{ZF}[n] &= (1+\beta^2)I[n] - \beta I[n-1] - \beta I[n+1] + V'[n] \\ V'[n] &= (1+\alpha^2)V[n] - \alpha V[n-1] - \alpha V[n+1]. \end{aligned}$$

(b) Again we do our analysis in the Z-domain.

$$\begin{aligned} S_{V'}(z) &= D'_{ZF}(z)D'^{*}_{ZF}(1/z^*)D(z)\frac{N_0}{2} \\ &= (1-\alpha z)(1-\alpha z^{-1})(1-\beta z)(1-\beta z^{-1})\frac{N_0}{2}. \end{aligned}$$

Recall that the power spectral density is the Z-transform of the autocorrelation. We can therefore find the noise power by looking at the coefficients corresponding to $z=0$ in the power spectral density:

$$\begin{aligned} \mathbb{E}[V'^2] &= R_{V'}[0] \\ &= (1+\alpha^2+2\alpha\beta+\beta^2+\alpha^2\beta^2)\frac{N_0}{2}. \end{aligned}$$

(c) We find our new noise term to be

$$G[n] = -\beta I[n-1] - \beta I[n+1] + V'[n].$$

Fortunately, $I[n]$ is both i.i.d. and independent of V'. This gives us the variance as just a linear combination of all the individual variances.

$$\mathbb{E}[G^2] = 2\beta^2\mathcal{E} + (1+\alpha^2+2\alpha\beta+\beta^2+\alpha^2\beta^2)\frac{N_0}{2}.$$

Then, we find P_e as

$$P_e \approx Q\left(\frac{(1+\beta^2)\sqrt{\mathcal{E}}}{\sqrt{2\beta^2\mathcal{E} + (1+\alpha^2+2\alpha\beta+\beta^2+\alpha^2\beta^2)\frac{N_0}{2}}}\right).$$

Notice how our pretend-Gaussian features the signal power $\mathcal{E}$ in its variance. In this expression, both the signal power and the noise power scale linearly in the signal power $\mathcal{E}$. Hence in the high power region the error probability will not vanish. Recall that a precise zero-forcing filter will be asymptotically optimal in high SNR, a mismatched (even slightly!) zero-forcing filter will have deteriorated performance.