Handout 9 Advanced Digital Communications
Solutions to Homework 3 Oct. 17, 2016

SOLUTION 1.

(a) Let's consider the case where $n = 2$ first, we have

$$\Pr\{Z = 0\} = \Pr\{X_1 \oplus X_2 = 0\} = \Pr\{X_1 = 0, X_2 = 0\} + \Pr\{X_1 = 1, X_2 = 1\} = \frac{1}{2},$$

where we used independence of X_1 and X_2. By induction, one could easily show that for arbitrary n, we have

$$\Pr\{Z = 0\} = \frac{1}{2}.$$

(i) For any x and z in $\{0, 1\}$,

$$\begin{aligned}
\Pr(Z = z|X_1 = x) &= \Pr\{X_1 \oplus X_2 \oplus \cdots \oplus X_n = z|X_1 = z\} \\
&= \Pr\{X_2 \oplus \cdots \oplus X_n = x \oplus z|X_1 = x\} \\
&= \Pr\{X_2 \oplus \cdots \oplus X_n = x \oplus z\} \\
&= \frac{1}{2} \\
&= \Pr\{Z = z\}
\end{aligned}$$

where the third equality follows since X_is are independent and the fourth since the parity check on any number of uniformly distributed binary RVs is, itself, uniformly distributed in $\{0, 1\}$ (as we showed above). We conclude that Z is independent of X_1

(ii) Similarly, for any $x_1, x_2, \ldots, x_{n-1}$ and z in $\{0, 1\}$,

$$\begin{aligned}
&\Pr\{Z = z|X_1 = x_1, \ldots, X_{n-1} = x_{n-1}\} \\
&\quad = \Pr\{X_1 \oplus X_2 \oplus \cdots \oplus X_n = z|X_1 = x_1, \ldots, X_{n-1} = x_{n-1}\} \\
&\quad = \Pr\{X_n = z \oplus x_1 \oplus \cdots \oplus x_{n-1}|X_1 = x_1, \ldots, X_{n-1} = x_{n-1}\} \\
&\quad = \Pr\{X_n = z \oplus x_1 \oplus \cdots \oplus x_{n-1}\} \\
&\quad = \frac{1}{2} \\
&\quad = \Pr\{Z = z\}.
\end{aligned}$$

We,hence, conclude that Z is independent of $X_1, \ldots, X_{n-1}$.

(iii) No, Z is a deterministic function of $X_1, \ldots, X_n$, more precisely,

$$\Pr\{Z = z|X_1 = x_1, \ldots, X_n = x_n\} = \mathbb{1}\{z = x_1 \oplus \ldots x_n\} \neq \Pr\{Z = z\}.$$

(iv) Suppose $\Pr\{X_i = 1\} = \frac{3}{4}$, we have

$$\begin{aligned}
\Pr\{Z = 0\} = \Pr\{X_1 \oplus X_2 = 0\} &= \Pr\{X_1 = 0, X_2 = 0\} \\
&\quad + \Pr\{X_1 = 1, X_2 = 1\} = \frac{9 + 1}{16} = \frac{5}{8},
\end{aligned}$$

but

$$\Pr\{Z = 0 | X_1 = 0\} = \Pr\{X_1 \oplus X_2 = 0 | X_1 = 0\} = \Pr\{X_2 = 0\} = \frac{1}{4} \neq \frac{5}{8}$$

where we used the independence of X_1 and X_2. Thus Z is not independent of X_1.

(b) Define the projection matrix

$$P := [\ \boldsymbol{\psi}_1^T \mid \boldsymbol{\psi}_2^T \mid \dots \mid \boldsymbol{\psi}_n^T\ ],$$

that is, the $n \times n$ matrix whose columns are the basis vectors $\{\boldsymbol{\psi}_1, \dots, \boldsymbol{\psi}_n\}$ (note that in this problem the vectors are assumed to be row vectors). Then, $\boldsymbol{W} = \boldsymbol{Z}P$ and, hence, like $\boldsymbol{Z}$, $\boldsymbol{W}$ is a zero-mean Gaussian vector. The covariance matrix of $\boldsymbol{W}$ is

$$\Sigma_{\boldsymbol{W}} = \mathbb{E}[\boldsymbol{W}^T \boldsymbol{W}] = \mathbb{E}[P^T \boldsymbol{Z}^T \boldsymbol{Z} P] = P^T \cdot \sigma^2 I \cdot P$$

since $\boldsymbol{Z}$ has independence components (hence covariance matrix $\Sigma_{\boldsymbol{Z}} = \sigma^2 I$). Finally since $\{\boldsymbol{\psi}_1, \dots, \boldsymbol{\psi}_n\}$ is an orthonormal basis $P^T P = I$ (it is easy to see that $(P^T P)_{i,j} = \boldsymbol{\psi}_i \boldsymbol{\psi}_j^T = \mathbb{1}\{i = j\}$). Consequently, $\Sigma_{\boldsymbol{W}} = \sigma^2 I$ which proves $\boldsymbol{W}$ has the same distribution as $\boldsymbol{Z}$.

Solution 2.

(a) Let D_0, D_1 be the MAP decision regions for hypotheses 0 and 1 when the apriori probabilities are $(\pi_0, 1 - \pi_0)$. Similarly, let D_0' and D_1' be the MAP decision regions for hypotheses 0 and 1 when the apriori probabilities are $(\pi_0', 1 - \pi_0')$ and D_0'', D_1'' be the MAP decision regions for hypotheses 0 and 1 when the apriori probabilities are $(\pi_0'', 1 - \pi_0'')$, where $\pi_0'' = \lambda \pi_0 + (1 - \lambda)\pi_0'$. Thus

$$\begin{aligned}
V(\pi_0) &= \pi_0 p_0(D_1) + (1 - \pi_0) p_1(D_0), \\
V(\pi_0') &= \pi_0 p_0(D_1') + (1 - \pi_0) p_1(D_0'), \\
V(\pi_0'') &= \pi_0 p_0(D_1'') + (1 - \pi_0) p_1(D_0''),
\end{aligned}$$

where p_0 (respectively p_1) denotes the distribution of the observable conditioned on the hypothesis 0 (resp. 1).

(b) Since the MAP rule minimizes the error probability, using any other decision regions in any of the above will increase the probability of error. So,

$$\begin{aligned}
V(\pi_0) &\leq \pi_0 p_0(D_1'') + (1 - \pi_0) p_1(D_0''), \\
V(\pi_0') &\leq \pi_0' p_0(D_1'') + (1 - \pi_0') p_1(D_0'').
\end{aligned}$$

Multiplying the first by λ and the second by $(1 - \lambda)$ and adding we get the desired result:

$$\begin{aligned}
\lambda V(\pi_0) + (1 - \lambda) V(\pi_0') &\leq (\lambda \pi_0 + (1 - \lambda)\pi_0') p_0(D_1'') + (1 - (\lambda \pi_0 + (1 - \lambda)\pi_0')) p_1(D_0'') \\
&= V(\lambda \pi_0 + (1 - \lambda)\pi_0')
\end{aligned}$$

SOLUTION 3.

(a) Let $P_e(A)$ denote the error probability of the optimum detector for the signal set A. Because of the additive nature of the channel, when the translated signal set A' is used at the transmitter, a receiver that adds $m_A(t)$ to the received signal and uses the same decision rules as that for the signal set A will have the error probability $P_e(A)$. Consequently, $P_e(A') \leq P_e(A)$. (Note that $P_e(A')$ is the error probability of the optimum receiver for signal set A' which we still don't know if it is the one we just described.) Swapping the roles of A and A' we also get $P_e(A) \leq P_e(A')$. Therefore we conclude that $P_e(A') = P_e(A)$.

(b)

$$
\begin{aligned}
\mathcal{E}_{A'} &= \sum_j p_i \|a_j(t) - m_A(t)\|^2 \\
&= \sum_j p_j \|a_j(t)\|^2 + \sum_j p_i \|m_A(t)\|^2 - \sum_j 2p_j \langle a_j(t), m_A(t)\rangle \\
&\overset{(*)}{=} \mathcal{E}_A + \|m_A(t)\|^2 + \sum_j 2p_j \langle a_j(t), m_A(t)\rangle \\
&= \mathcal{E}_A + \|m_A(t)\|^2 - 2\left\langle \sum_j p_j a_j(t), m_A(t)\right\rangle \\
&\overset{(\dagger)}{=} \mathcal{E}_A + \|m_A(t)\|^2 - 2\langle m_A(t), m_A(t)\rangle \\
&= \mathcal{E}_A - \|m_A(t)\|^2,
\end{aligned}
$$

where $(*)$ follows from the definition of average energy and that $\sum_j p_j = 1$ and $(\dagger)$ from the definition of the average signal $m_A(t)$. By part (a), adding a constant vector $(-m(A))$ does not change the error probability, but it reduces the average transmitted energy, so it is good.

SOLUTION 4.

(a) From the definition of white Gaussian noise (Section 3.2) it follows that the random variable $Z = \int z(\tau)h(T-\tau)\, dt$ has zero mean and variance equal to $N_0/2 \int h(T-\tau)^2\, dt = N_0/2\|h(t)\|^2$.

(b) Using Cauchy–Schwarz inequality,

$$
\begin{aligned}
|(h * s)(T)|^2 &= \left| \int_{-\infty}^{\infty} h(\tau)s(T-\tau)d\tau \right|^2 \\
&\leq \int_{-\infty}^{\infty} h^2(\tau)d\tau \int_{-\infty}^{\infty} s^2(T-\tau)d\tau = \|s(t)\|^2 \cdot \|h(t)\|^2
\end{aligned}
$$

with equality if and only if $h(t) = as(T-t)$ for some constant a. Consequently,

$$
\mathrm{SNR} = \frac{|(h * s)(T)|^2}{\mathrm{var}\big((h * Z)(T)\big)} \leq \frac{\|h(t)\|^2\|s(t)\|^2}{N_0/2\|h(t)\|^2} = \frac{\|s(t)\|^2}{N_0/2}
$$

with equality if and only if $h(t) = as(T-t)$ for some constant a. Hence, we can conclude that in order to maximize the SNR, we need to select the filter $h(t) = as(T-t)$, which is precisely the matched filter.

SOLUTION 5.

(a) Since the space spanned by $\{x_1(t), x_2(t)\}$ is the same as the space spanned by $\{\varphi_1(t), x_2(t)\}$, we can obtain $\varphi_2(t)$ by applying the Gram–Schmidt procedure on $\{\varphi_1(t), x_2(t)\}$:

$$\begin{aligned}
x_2(t) - \langle x_2(t), \varphi_1(t)\rangle\varphi_1(t) &= x_2(t) - \left\langle x_2(t), \frac{x_1(t)-x_2(t)}{\|x_1(t)-x_2(t)\|}\right\rangle \frac{x_1(t)-x_2(t)}{\|x_1(t)-x_2(t)\|} \\
&= x_2(t) - \frac{\langle x_1(t), x_2(t)\rangle - \|x_2(t)\|^2}{\|x_1(t)-x_2(t)\|^2}\cdot(x_1(t)-x_2(t)) \\
&= x_2(t) - \frac{\langle x_1(t), x_2(t)\rangle - \|x_2(t)\|^2}{\|x_1(t)\|^2 + \|x_2(t)\|^2 - 2\langle x_1(t), x_2(t)\rangle}\cdot(x_1(t)-x_2(t)) \\
&= x_2(t) - \frac{\langle x_1(t), x_2(t)\rangle - \mathcal{E}}{2\mathcal{E} - 2\langle x_1(t), x_2(t)\rangle}\cdot(x_1(t)-x_2(t)) \\
&= x_2(t) + \frac{1}{2}(x_1(t)-x_2(t)) = \frac{1}{2}\cdot(x_1(t)+x_2(t)).
\end{aligned}$$

Therefore,

$$\varphi_2(t) = \frac{x_2(t) - \langle x_2(t), \varphi_1(t)\rangle\varphi_1(t)}{\|x_2(t) - \langle x_2(t), \varphi_1(t)\rangle\varphi_1(t)\|} = \frac{x_1(t)+x_2(t)}{\|x_1(t)+x_2(t)\|}.$$

(b) As explained in Section 3.3.1 of your lecture notes, the optimal receiver needs to computes the sufficient statistic (Y_1, Y_2) where $Y_1 = \langle Y(t), \varphi_1(t)\rangle$ and $Y_2 = \langle Y(t), \varphi_2(t)\rangle$. Since the two hypotheses are equally likely, the optimal decision (i.e., the MAP decision) reduces to the ML decision which is the minimum distance decoding, that is with $\boldsymbol{Y} = (Y_1, Y_2)$, and $\boldsymbol{x}_i = \big(\langle x_i(t), \psi_1(t)\rangle, \langle x_i(t), \psi_2(t)\rangle\big)$, $i = 1, 2$,

$$\hat{H}_{\mathrm{ML}}(\boldsymbol{Y} = \boldsymbol{y}) = \underset{i\in\{1,2\}}{\arg\min}\|\boldsymbol{y} - \boldsymbol{x}_i\|.$$

Note that Y_1 and Y_2 can be computed by passing the received signals through filters with impulse response $h_1(t)$ and $h_2(t)$ and sampling their outputs at time $t = T$.

(c) Let $Z_1 = \langle Z(t), \varphi_1(t)\rangle$ and $Z_2 = \langle Z(t), \varphi_2(t)\rangle$. Z_1 and Z_2 are independent because $\varphi_1(t)$ and $\varphi_2(t)$ are orthogonal. We have:

$$\begin{aligned}
Y_2 = \langle Y(t), \varphi_2(t)\rangle &= \begin{cases} \langle x_1(t), \frac{x_1(t)+x_2(t)}{\|x_1(t)+x_2(t)\|}\rangle + Z_2 & \text{if } x_1 \text{ is sent,} \\ \langle x_2(t), \frac{x_1(t)+x_2(t)}{\|x_1(t)+x_2(t)\|}\rangle + Z_2 & \text{if } x_2 \text{ is sent.} \end{cases} \\
&= \begin{cases} \frac{\|x_1(t)\|^2 + \langle x_1(t), x_2(t)\rangle}{\|x_1(t)+x_2(t)\|} + Z_2 & \text{if } x_1 \text{ is sent,} \\ \frac{\langle x_2(t), x_1(t)\rangle + \|x_1(t)\|^2}{\|x_1(t)+x_2(t)\|} + Z_2 & \text{if } x_2 \text{ is sent.} \end{cases} \\
&= \begin{cases} \frac{\mathcal{E} + \langle x_1(t), x_2(t)\rangle}{\|x_1(t)+x_2(t)\|} + Z_2 & \text{if } x_1 \text{ is sent,} \\ \frac{\mathcal{E} + \langle x_1(t), x_2(t)\rangle}{\|x_1(t)+x_2(t)\|} + Z_2 & \text{if } x_2 \text{ is sent.} \end{cases}
\end{aligned}$$

This shows that the distribution of Y_2 is independent from the transmitted signal (and from Y_1). Therefore, Y_2 can be thrown away. Hence, Y_1 is sufficient statistics for the hypothesis testing problem. Thus, the receiver only needs to compute Y_1 for which the matched filter $h_1(t)$ is sufficient.

(d) We have:

$$Y_1 = \langle Y(t), \varphi_1(t) \rangle = \begin{cases} \langle x_1(t), \frac{x_1(t)-x_2(t)}{\|x_1(t)-x_2(t)\|} \rangle + Z_1 & \text{if } x_1 \text{ is sent,} \\ \langle x_2(t), \frac{x_1(t)-x_2(t)}{\|x_1(t)-x_2(t)\|} \rangle + Z_1 & \text{if } x_2 \text{ is sent.} \end{cases}$$

$$= \begin{cases} \frac{\|x_1(t)\|^2 - \langle x_1(t), x_2(t) \rangle}{\|x_1(t)-x_2(t)\|} + Z_1 & \text{if } x_1 \text{ is sent,} \\ \frac{\langle x_2(t), x_1(t) \rangle - \|x_1(t)\|^2}{\|x_1(t)-x_2(t)\|} + Z_1 & \text{if } x_2 \text{ is sent.} \end{cases}$$

$$= \begin{cases} \frac{\mathcal{E} - \langle x_1(t), x_2(t) \rangle}{\|x_1(t)-x_2(t)\|} + Z_1 & \text{if } x_1 \text{ is sent,} \\ \frac{\langle x_1(t), x_2(t) \rangle - \mathcal{E}}{\|x_1(t)-x_2(t)\|} + Z_1 & \text{if } x_2 \text{ is sent.} \end{cases}$$

Note that $\|x_1(t) - x_2(t)\|^2 = \|x_1(t)\|^2 + \|x_2(t)\|^2 - 2\langle x_1(t), x_2(t) \rangle = 2\mathcal{E} - 2\langle x_1(t), x_2(t) \rangle$. Therefore,

$$Y_1 = \begin{cases} \frac{\|x_1(t)-x_2(t)\|^2}{2\|x_1(t)-x_2(t)\|} + Z_0 & \text{if 0 is sent,} \\ \frac{-\|x_1(t)-x_2(t)\|^2}{2\|x_1(t)-x_2(t)\|} + Z_0 & \text{if 1 is sent.} \end{cases}$$

$$= \begin{cases} \frac{1}{2}\|x_1(t) - x_2(t)\| + Z_0 & \text{if 0 is sent,} \\ -\frac{1}{2}\|x_1(t) - x_2(t)\| + Z_0 & \text{if 1 is sent.} \end{cases}$$

Now since $Z_0 = \langle N, \varphi_1(t) \rangle \sim \mathcal{N}(0, \frac{N_0}{2})$, the probability of error of the MAP decoder is given by

$$P_e = Q\left(\frac{\frac{1}{2}\|x_1(t) - x_2(t)\|}{\sqrt{\frac{N_0}{2}}} \right) = Q\left(\frac{\|x_1(t) - x_2(t)\|}{\sqrt{2N_0}} \right).$$

The Cauchy–Schwarz inequality gives $|\langle x_1(t), x_2(t) \rangle| \leq \|x_1(t)\| \cdot \|x_2(t)\| = \mathcal{E}$. Therefore, $\langle x_1(t), x_2(t) \rangle \geq -\mathcal{E}$. Hence,

$$\|x_1(t) - x_2(t)\|^2 = 2\mathcal{E} - 2\langle x_1(t), x_2(t) \rangle \leq 2\mathcal{E} + 2\mathcal{E} = 4\mathcal{E}.$$

We conclude that $\|x_1(t) - x_2(t)\| \leq 2\sqrt{\mathcal{E}}$. Therefore, the probability of error of the MAP decoder is lower-bounded as follows:

$$P_e = Q\left(\frac{\|x_1(t) - x_2(t)\|}{\sqrt{2N_0}} \right) \overset{(\star)}{\geq} Q\left(\frac{2\sqrt{\mathcal{E}}}{\sqrt{2N_0}} \right) = Q\left(\sqrt{\frac{2\mathcal{E}}{N_0}} \right).$$

Moreover, $(\star)$ becomes an equality when $\langle x_1(t), x_2(t) \rangle = -\mathcal{E} = -\|x_1(t)\| \cdot \|x_2(t)\|$, which is true if $x_2(t) = -x_1(t)$.

SOLUTION 6. Before we give the solution using the whitening filter approach studied here, we point out that the ML detection based on a direct calculation of the likelihood ratio works too. The conditional probability density function is

$$f_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x}_i) = \frac{1}{(2\pi)^{-1}|\Sigma|^{-1/2}} \exp\left\{ -\frac{(\boldsymbol{y} - \boldsymbol{x}_i)^T \Sigma^{-1} (\boldsymbol{y} - \boldsymbol{x}_i)}{2} \right\}$$

for $i = 0, 1$. As we will show, Σ can be written as $\Sigma = CC^H$ for some C. Hence the log-likelihood ratio will look like

$$\text{log-likelihood ratio} = \ln\|C^{-1}\boldsymbol{x}_0 - C^{-1}\boldsymbol{y}\|^2 - \ln\|C^{-1}\boldsymbol{x}_1 - C^{-1}\boldsymbol{y}\|^2$$

We will make the decision according to the sign of the above quantity. It is easy to see that this is equivalent to the decision rule we derive using the whitening filter approach below.

(a) If we define $C = \Phi\Lambda^{1/2}$, where $\Lambda^{1/2}$ is also a diagonal matrix with its diagonal elements being square root of diagonal elements of Λ, then

$$\begin{aligned} CC^H &= \Phi\Lambda^{1/2}(\Lambda^{1/2})^H\Phi^H \\ &= \Phi\Lambda\Phi^H = \Sigma \end{aligned}$$

The covariance matrix of $C^{-1}\boldsymbol{Z}$ is

$$\begin{aligned} \mathbb{E}[C^{-1}\boldsymbol{Z}(C^{-1}\boldsymbol{Z})^H] &= C^{-1}\,\mathbb{E}[\boldsymbol{Z}\boldsymbol{Z}^H](C^{-1})^H \\ &= C^{-1}\Sigma(C^{-1})^H \\ &= C^{-1}CC^H(C^H)^{-1} = I \end{aligned}$$

(b) We process $\boldsymbol{Y}$ as

$$\boldsymbol{S} = C^{-1}\boldsymbol{Y} = C^{-1}\boldsymbol{x} + C^{-1}\boldsymbol{Z}$$

where $C = \Phi\Lambda^{1/2}$ as given in (a). This will make sure that equivalent noise $C^{-1}\boldsymbol{Z}$ is white. Our system can be equivalently written as

$$\boldsymbol{S} = \tilde{\boldsymbol{x}} + \boldsymbol{W}$$

with $\boldsymbol{W} := C^{-1}\boldsymbol{Z}$ a Gaussian vector with covariance matrix I and $\tilde{\mathbf{x}}$ chosen uniformly in $\{C^{-1}\boldsymbol{x}_0, C^{-1}\boldsymbol{x}_1\}$.

Now we have a standard AWGN vector problem with white Gaussian noise and as we have seen in lecture notes, Section 3.4, and we know the ML detection rule is

$$\hat{H}_{\mathrm{ML}}(\boldsymbol{Y} = \boldsymbol{y}) = \arg\min_{i\in\{0,1\}} \|C^{-1}\boldsymbol{x}_i - C^{-1}\boldsymbol{y}\|$$

Due to the symmetry, the error probability is the same for either hypotheses. When $\boldsymbol{x}_1$ is sent, $\boldsymbol{Y} = \boldsymbol{x}_1 + \boldsymbol{Z}$ and according to the detection rule above, we make an error if

$$\|C^{-1}\boldsymbol{x}_0 - C^{-1}\boldsymbol{Y}\|^2 \leq \|C^{-1}\boldsymbol{x}_1 - C^{-1}\boldsymbol{Y}\|^2$$

which is equivalent to

$$\langle C^{-1}(\boldsymbol{x}_0 - \boldsymbol{x}_1), C^{-1}\boldsymbol{Y}\rangle \geq \frac{\|C^{-1}\boldsymbol{x}_0\|^2 - \|C^{-1}\boldsymbol{x}_1\|^2}{2}$$

Following along the same lines as in the lecture notes, Section 3.4.2. (after some algebra) we have

$$P_e = Q\left(\frac{\|C^{-1}(\boldsymbol{x}_0 - \boldsymbol{x}_1)\|}{2}\right)$$

(c) For Σ_1, the noise is already white hence have $C^{-1} = I$, and the detection rule is

$$\hat{H}_{ML}(\boldsymbol{Y} = \boldsymbol{y}) = \begin{cases} 0 & \text{if } \|\boldsymbol{y} - (1,0)^T\| \leq \|\boldsymbol{y} - (0,-1)^T\|. \\ 1 & \text{otherwise,} \end{cases}$$

which can be simplified to

$$\hat{H}_{\mathrm{ML}}(\boldsymbol{Y} = \boldsymbol{y}) = \begin{cases} 0 & \text{if } y_1 \geq -y_2 \\ 1 & \text{otherwise.} \end{cases}$$

(In the above y_1 and y_2 denote the first and second entry of $\boldsymbol{y}$, respectively.) The decision regions are shown in the left figure below and and the error probability is

$$P_e = Q(\sqrt{2}/2) \approx 0.2398.$$

For the case with Σ_2, it can be diagonalized as $\Sigma_2 = \Phi\Lambda\Phi^H$ with

$$\Phi = \frac{1}{\sqrt{2}}\begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} \quad \Lambda = \begin{bmatrix} 2/3 & 0 \\ 0 & 4/3 \end{bmatrix}$$

and the whitening filter is

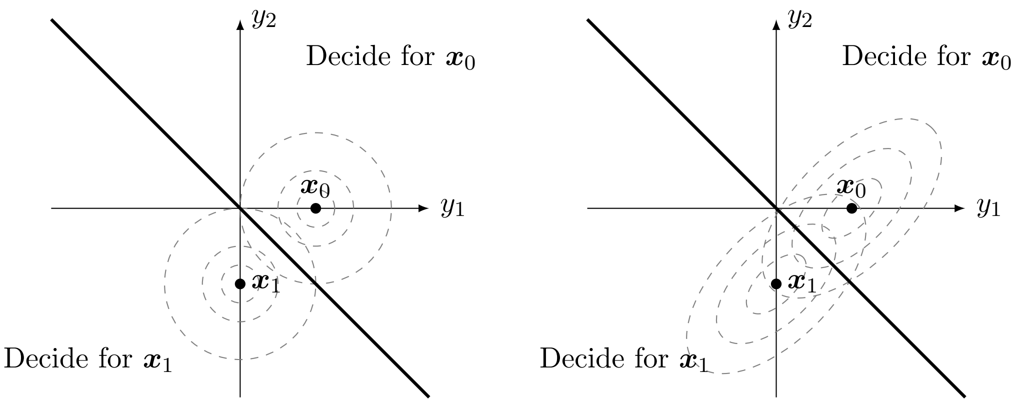
$$C^{-1} = (\Phi\Lambda^{1/2})^{-1} = \begin{bmatrix} -\frac{\sqrt{3}}{2} & \frac{\sqrt{3}}{2} \\ \sqrt{\frac{3}{8}} & \sqrt{\frac{3}{8}} \end{bmatrix}$$

and the ML detection rule in (b) can be simplified to

$$\hat{H}_{\mathrm{ML}}(\boldsymbol{Y} = \boldsymbol{y}) = \begin{cases} 0 & \text{if } y_1 \geq -y_2, \\ 1 & \text{otherwise.} \end{cases}$$

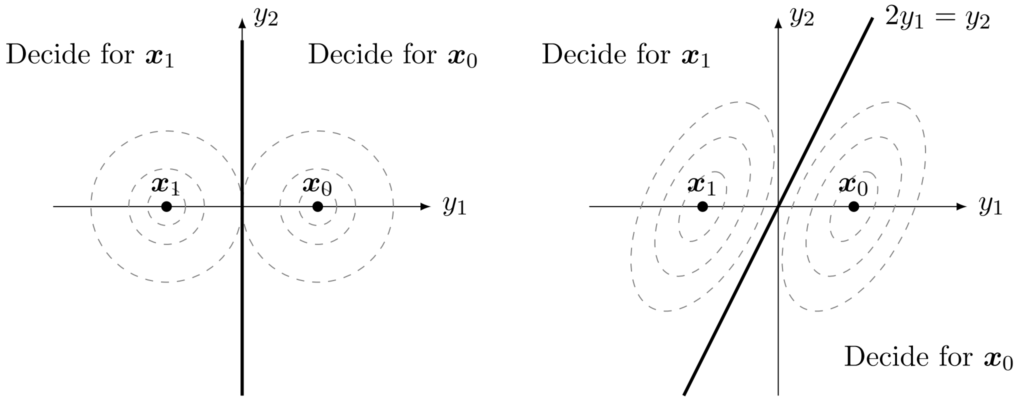
The decision regions are shown in the right figure below, and the error probability is

$$P_e = Q\left(\sqrt{\frac{3}{8}}\right) \approx 0.2701$$



The dashed curves are the contour line of the probability density function of the noise.

Remark. Notice that the detection rule here is the same for both cases, but the colored noise results a higher error probability. However, this is not always the case. If the two signal points are $\boldsymbol{x}_0 = (1, 0)$ and $\boldsymbol{x}_1 = (-1, 0)$. Using exactly the same calculation above we can find the following decision regions (the left figure shows the regions for Σ_1 and the right figure for Σ_2). We can show that the error probability with white noise is ≈ 0.159 and with colored noise is ≈ 0.125. In this case the error probability with the colored noise is *smaller* than that with the white noise. The intuition can already be seen from the contour line of the noise density in the plot.



The dashed curves are the contour line of the probability density function of the noise.