ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE
School of Computer and Communication Sciences

Principles of Digital Communications: Assignment date: March 7, 2012
Summer Semester 2012 Due date: March 14, 2012

Solution of Homework 4

Problem 1. *(Fisher-Neyman Factorization Theorem)*
The pdf $f_{Y|H}$ is a non-negative function. Hence without loss of generality we may assume that both g_i and h are non-negative.

1. The MAP decision rule can always be written as

$$\begin{aligned}\hat{H}(y) &= \arg\max_i f_{Y|H}(y|i)P_H(i)\\ &= \arg\max_i g_i(T(y))h(y)P_H(i)\\ &= \arg\max_i g_i(T(y))P_H(i)\end{aligned}$$

The last step is valid because $h(y)$ is non-negative constant which is independent of i and thus does not give any further information for our decision.

2. Recall, that if Y is a random variable with probability density function $f_Y(y)$ and $\mathcal{B}$ is an event, then

$$f_{Y|Y\in\mathcal{B}} = \frac{f_Y(y)1_{\mathcal{B}}(y)}{\int_{\mathcal{B}} f_Y(y)dy}, \tag{1}$$

where $1_{\mathcal{B}}(y)$ is the indicator function,

$$1_{\mathcal{B}}(y) = \begin{cases} 1 & \text{if } y \in \mathcal{B} \\ 0 & \text{otherwise.} \end{cases}$$

Now, consider our original problem where $T(Y)$ is a function of Y. Note that for every t, we can define the event $\mathcal{B}_t = \{y : T(y) = t\}$. Using (1), we have

$$f_{Y|H,T(Y)}(y|i,t) = \frac{f_{Y|H}(y|i)1_{\mathcal{B}_t}(y)}{\int_{\mathcal{B}_t} f_{Y|H}(y|i)dy}.$$

If $f_{Y|H}(y|i) = g_i(T(y))h(y)$, then

$$\begin{aligned} f_{Y|H,T(Y)}(y|i,t) &= \frac{g_i(T(y))h(y)1_{\mathcal{B}_t}(y)}{\int_{\mathcal{B}_t} g_i(T(y))h(y)dy} \\ &= \frac{g_i(t)h(y)1_{\mathcal{B}_t}(y)}{g_i(t)\int_{\mathcal{B}_t} h(y)dy} \\ &= \frac{h(y)1_{\mathcal{B}_t}(y)}{\int_{\mathcal{B}_t} h(y)dy}. \end{aligned}$$

Hence, we see that $f_{Y|H,T(Y)}(y|i,t)$ does not depend on i so $H \to T(Y) \to Y$.

In the following we verify the above results for two examples.

1. *(Example 1)* Note that $P_{Y_k|H}(1|i) = p_i, P_{Y_k|H}(0|i) = 1 - p_i$ and

$$P_{Y_1,\dots,Y_n|H}(y_1,\dots,y_n|i) = P_{Y_1|H}(y_1|i)\dots P_{Y_n|i}(y_n|i).$$

Thus, we have

$$P_{Y_1,\dots,Y_n|H}(y_1,\dots,y_n|i) = p_i^t(1-p_i)^{(n-t)},$$

where $t = \sum_k y_k$.

Choosing $g_i(t) = p_i^t(1-p_i)^{(n-t)}$ and $h(y) = 1$, we see that $P_{Y_1,\dots,Y_n|H}(y_1,\dots,y_n|i)$ fulfills the condition in the question.

2. *(Example 2)* We have $f_{Y_k|H}(y|i) = \frac{1}{\sqrt{2\pi}}e^{-\frac{(y-m_i)^2}{2}}$ and

$$f_{Y_1,\dots,Y_n|H}(y_1,\dots,y_n|i) = \prod_{k=1}^{n} f_{Y_k|H}(y_k|i)$$

since $Y_1,\dots,Y_n$ are independent. Thus,

$$\begin{aligned} f_{Y_1,\dots,Y_n|H}(y_1,\dots,y_n|i) &= \prod_{k=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{(y_k-m_i)^2}{2}} \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\sum_{k=1}^{n}\frac{(y_k-m_i)^2}{2}} \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{\sum_{k=1}^{n} y_k^2}{2}} e^{nm_i(\frac{1}{n}\sum_{k=1}^{n} y_k - \frac{m_i}{2})}. \end{aligned}$$

Choosing $g_i(t) = e^{nm_i(t-\frac{m_i}{2})}$ and $h(y_1,\dots,y_n) = \frac{1}{(2\pi)^{\frac{n}{2}}}e^{-\frac{\sum_{k=1}^{n} y_k^2}{2}}$, we see that

$$f_{Y_1,\dots,Y_n|H}(y_1,\dots.y_n|i) = g_i(T(y_1,\dots,y_n))h(y_1,\dots y_n)$$

hence the condition in the question is fulfilled.

Problem 2. *(Q-Function on Regions)*

1. One can see that the event $\{\mathbf{X} \in \text{Region}\}$ only depends on the first component X_1. Hence, we have

$$\begin{aligned}
\Pr\big\{\mathbf{X} \in \text{Region}\big\} &= \Pr\big\{\{X_1 \geq -2\} \cap \{X_1 \leq 1\}\big\} \\
&= 1 - \Pr\big\{\{X_1 < -2\} \cup \{X_1 > 1\}\big\} \\
&= 1 - Q(\frac{2}{\sigma}) - Q(\frac{1}{\sigma}),
\end{aligned}$$

where the last equality is true because $\{X_1 < -2\}$ and $\{X_1 > 1\}$ are disjoint events.

2. Since X_1 and X_2 are independent and have the **same** variance, rotating the vector $\mathbf{X}$ by any angle around the origin does not change its distribution. Equivalently, we can rotate the square region in Figure (b) by 45 degrees, and the probability of $\mathbf{X}$ being in the rotated region is the same as for the original region. The new region is a square whose edges are parallel to the axes of the coordinate system. The points where the edges of the square intersect the axes are $(\sqrt{2}, 0)$, $(-\sqrt{2}, 0)$, $(0, \sqrt{2})$ and $(0, -\sqrt{2})$. Hence,

$$\begin{aligned}
\Pr\big\{\mathbf{X} \in \text{Region}\big\} &= \Pr\big\{\{-\sqrt{2} \leq X_1 \leq \sqrt{2}\} \cap \{-\sqrt{2} \leq X_2 \leq \sqrt{2}\}\big\} \\
&\stackrel{(1)}{=} \Pr\big\{\{-\sqrt{2} \leq X_1 \leq \sqrt{2}\}\big\}^2 \\
&= \Big[1 - \Pr\big\{\{X_1 \leq -\sqrt{2}\} \cup \{X_1 \geq \sqrt{2}\}\big\}\Big]^2 \\
&= \big[1 - 2Q(\frac{\sqrt{2}}{\sigma})\big]^2,
\end{aligned}$$

where (1) holds because X_1 and X_2 are independent and identically distributed.

3. We solve this part using two different ways:

   (a) *First Solution:* From the same argument as in previous part, we can rotate $\mathbf{X}$ such that one of its components, say X_1, is perpendicular to the straight line that delimits the shaded region in Figure (c). Then, we need to know the shortest distance d of that line to the origin (the length of a segment that starts at $(0, 0)$ and is perpendicular to the line). Using standard trigonometric techniques, one finds that this length is $d = \frac{2}{\sqrt{5}}$ (An even more straight forward way to find d is to use the fact that corresponding sides of similar triangles have length in the same ratio so $\frac{d}{1} = \frac{2}{\sqrt{5}}$). Then, it follows that

$$\begin{aligned}
\Pr\big\{\mathbf{X} \in \text{Region}\big\} &= \Pr\big\{X_1 \geq \frac{2}{\sqrt{5}}\big\} \\
&= Q(\frac{2}{\sqrt{5}\sigma}).
\end{aligned}$$

(b) *Second Solution:* We are looking for the probability that $X_2 \geq 1 - \frac{1}{2}X_1$, i.e., the probability that $Z \triangleq X_2 + \frac{1}{2}X_1 - 1 \geq 0$. But $Z \sim \mathcal{N}(-1, \frac{5}{4}\sigma^2)$. Hence, $\Pr\{\mathbf{X} \in \text{Region}\} = \Pr\{Z \geq 0\} = Q(\frac{2}{\sqrt{5}\sigma})$.

Problem 3. *Comparison of 16-PAM and 16-QAM*

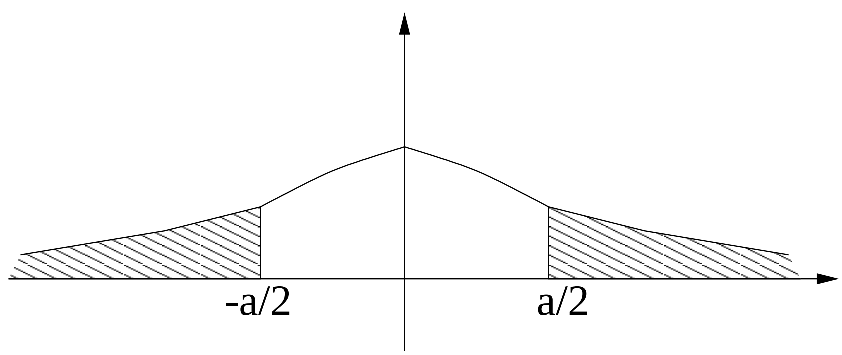
1. **16-PAM.** Denote the additive white Gaussian noise process by Z. Thus, Z is zero-mean Gaussian of variance σ^2, and the observation Y is also Gaussian of variance σ^2, but with mean corresponding to the particular signal point that is being transmitted. Label the signal points from left to right by $1, \ldots, 16$. Then,

$$\begin{aligned} Pr\{e|H=1\} &= Pr\{Y \geq -7a \mid H=1\} = Pr\{Z \geq \frac{a}{2}\} \\ &= Pr\{\frac{Z}{\sigma} \geq \frac{a}{2\sigma}\} = Q\left(\frac{a}{2\sigma}\right) \qquad (2) \end{aligned}$$

By symmetry, $Pr\{e|H=1\} = Pr\{e|H=16\}$. Moreover,

$$\begin{aligned} Pr\{e|H=2\} &= Pr\{Y \leq -7a \text{ or } Y \geq -6a \mid H=2\} \qquad (3) \\ &= Pr\{Z \leq -\frac{a}{2} \text{ or } Z \geq \frac{a}{2}\} \overset{(a)}{=} 2Pr\{Z \geq \frac{a}{2}\} \qquad (4) \\ &= 2Q\left(\frac{a}{2\sigma}\right). \qquad (5) \end{aligned}$$

The following schematic drawing should illustrate how we obtained equality in (a):



Again, by symmetry, $Pr\{e|H=i\} = Pr\{e|H=2\}$, for $i = 3, \ldots, 15$. Putting things together, we obtain

$$\begin{aligned} Pr\{e\} &= \sum_{i=1}^{16} p_H(i) Pr\{e|H=i\} = \sum_{i=1}^{16} \frac{1}{16} Pr\{e|H=i\} \qquad (6) \\ &= \frac{1}{16}\left(2 \cdot Q\left(\frac{a}{2\sigma}\right) + 14 \cdot 2Q\left(\frac{a}{2\sigma}\right)\right) \qquad (7) \\ &= \frac{15}{8} Q\left(\frac{a}{2\sigma}\right). \qquad (8) \end{aligned}$$

16-QAM. Denote the additive white Gaussian noise process in the x-direction by Z_1 and in the y-direction by Z_2. In our setup, both Z_1 and Z_2 are zero-mean Gaussian of variance σ^2. Label the signal points from left to right, top to bottom by $1, \ldots, 16$. Then, for the four corner points, we find

$$Pr\{e|H=1\} \quad = \quad Pr\{Y_1 \geq -b \text{ or } Y_2 \leq b \mid H=1\}. \tag{9}$$

However, the connection with "or" does not allow to decompose into two disjoint events. We may rewrite as follows to obtain a connection with "and":

$$\begin{aligned} Pr\{e|H=1\} &= 1 - Pr\{Y_1 \leq -b \text{ and } Y_2 \geq b \mid H=1\} && (10)\\ &= 1 - Pr\{Y_1 \leq -b \mid H=1\} \cdot Pr\{Y_2 \geq b \mid H=1\}. && (11) \end{aligned}$$

However, a simple way not to get trapped in this kind of logic is to consider the probability of *correct* decision rather than the probability of error. We will use this approach to derive the solution to the problem. Thus,

$$\begin{aligned} Pr\{correct|H=1\} &= Pr\{Y_1 \leq -b \text{ and } Y_2 \geq b \mid H=1\} && (12)\\ &= Pr\{Y_1 \leq -b \mid H=1\} \cdot Pr\{Y_2 \geq b \mid H=1\} && (13)\\ &= Pr\{Z_1 \leq \frac{b}{2}\} \cdot Pr\{Z_2 \geq \frac{-b}{2}\} && (14)\\ &= \left(1 - Q\left(\frac{b}{2\sigma}\right)\right) Q\left(-\frac{b}{2\sigma}\right) && (15)\\ &= \left(1 - Q\left(\frac{b}{2\sigma}\right)\right)^2. && (16) \end{aligned}$$

For the points on the edges (i.e. numbers $2, 3, 5, 8, 9, 12, 14, 15$), we find similarly

$$\begin{aligned} Pr\{correct|H=2\} &= Pr\{-b \leq Y_1 \leq 0 \text{ and } Y_2 \geq b \mid H=2\} && (17)\\ &= Pr\{-\frac{b}{2} \leq Z_1 \leq \frac{b}{2}\} \cdot Pr\{Z_2 \geq -\frac{b}{2}\}, && (18) \end{aligned}$$

where

$$\begin{aligned} Pr\{-\frac{b}{2} \leq Z_1 \leq \frac{b}{2}\} &= 1 - Pr\{Z_1 \leq -\frac{b}{2} \text{ or } Z_1 \geq \frac{b}{2}\} && (19)\\ &= 1 - 2Pr\{Z_1 \geq \frac{b}{2}\} && (20)\\ &= 1 - 2Q\left(\frac{b}{2\sigma}\right), && (21) \end{aligned}$$

thus,

$$Pr\{correct|H=2\} \quad = \quad \left(1 - 2Q\left(\frac{b}{2\sigma}\right)\right)\left(1 - Q\left(\frac{b}{2\sigma}\right)\right). \tag{22}$$

Finally, for the four points in the middle, we obtain

$$Pr\{correct|H=6\} = Pr\{-b \le Y_1 \le 0 \text{ and } 0 \le Y_2 \le b \mid H=6\} \tag{23}$$
$$= Pr\{-\frac{b}{2} \le Z_1 \le \frac{b}{2}\} \cdot Pr\{-\frac{b}{2} \le Z_2 \le \frac{b}{2}\} \tag{24}$$
$$= \left(1 - 2Q\left(\frac{b}{2\sigma}\right)\right)^2. \tag{25}$$

Putting things together, we find

$$Pr\{correct\} = \sum_{i=1}^{16} p_H(i) Pr\{correct|H=i\} = \sum_{i=1}^{16} \frac{1}{16} Pr\{correct|H=i\} \tag{26}$$
$$= \frac{1}{16}\Bigg[4 \cdot \left(1 - Q\left(\frac{b}{2\sigma}\right)\right)^2 + 8 \cdot \left(1 - Q\left(\frac{b}{2\sigma}\right)\right)\left(1 - 2Q\left(\frac{b}{2\sigma}\right)\right)$$
$$+4 \cdot \left(1 - 2Q\left(\frac{b}{2\sigma}\right)\right)\left(1 - 2Q\left(\frac{b}{2\sigma}\right)\right)\Bigg] \tag{27}$$
$$= 1 - 3Q\left(\frac{b}{2\sigma}\right) + \frac{9}{4}\left(Q\left(\frac{b}{2\sigma}\right)\right)^2. \tag{28}$$

From here, we find $Pr\{e\} = 1 - Pr\{correct\}$, thus

$$Pr\{e\} = 3Q\left(\frac{b}{2\sigma}\right) - \frac{9}{4}\left(Q\left(\frac{b}{2\sigma}\right)\right)^2. \tag{29}$$

2. **16-PAM.** By symmetry, we only consider the positive signals to find

$$E_s = 2\frac{1}{16}\left(\left(\frac{a}{2}\right)^2 + \left(\frac{3a}{2}\right)^2 + \ldots + \left(\frac{15a}{2}\right)^2\right) \tag{30}$$
$$= \frac{a^2}{32}\left(1 + 3^2 + 5^2 + \ldots + 15^2\right) = \frac{85a^2}{4}. \tag{31}$$

16-QAM. By symmetry, we only consider the positive quadrant to find

$$E_s = 4\frac{1}{16}\left(\left(\frac{b}{2}\right)^2 + \left(\frac{b}{2}\right)^2 + \left(\frac{3b}{2}\right)^2 + \left(\frac{3b}{2}\right)^2 + 2\left(\left(\frac{b}{2}\right)^2 + \left(\frac{3b}{2}\right)^2\right)\right) \tag{32}$$
$$= \frac{b^2}{16}\left(1 + 1 + 9 + 9 + 2(1 + 9)\right) = \frac{5b^2}{2}. \tag{33}$$

3. **16-PAM.** We find $a/2 = \sqrt{E_s/85}$, thus

$$Pr\{e\} = \frac{15}{8} Q\left(\sqrt{\frac{E_s}{85\sigma^2}}\right). \tag{34}$$

16-QAM. We find $b/2 = \sqrt{E_s/10}$, thus

$$Pr\{e\} = 3Q\left(\sqrt{\frac{E_s}{10\sigma^2}}\right) - \frac{9}{4}Q^2\left(\sqrt{\frac{E_s}{10\sigma^2}}\right). \tag{35}$$

To plot these functions, we use `matlab`. Unfortunately, `matlab` does not feature the Q-function directly; instead, there is

$$erfc(x) = \frac{2}{\sqrt{\pi}}\int_x^\infty e^{-t^2}dt. \tag{36}$$

By a change of variables, it is easy to show that

$$Q(x) = \frac{1}{2}erfc\left(\frac{x}{\sqrt{2}}\right). \tag{37}$$

The following `matlab` program does the job:

```
%
%  Principles of Digital Communications, Summer Semester 2001 (Prof. B. Rimoldi)
%
%  Michael Gastpar
%
logES = [ -2:0.1: 3];
ES = 10.^logES;  % this is E_s/\sigma^2

PrePAM = 15/8 *   1/2*erfc(  sqrt(ES/85)  /sqrt(2)); PreQAM =  3
*   1/2*erfc(  sqrt(ES/10)   /sqrt(2))
              - 9/4 *  1/2*erfc( sqrt(ES/10)   /sqrt(2)).^2;

loglog(ES, PrePAM, '--', ES, PreQAM); title('Comparison of 16-PAM
(--) and 16-QAM'); xlabel('E_s/\sigma^2'); ylabel('Pr\{e\}');
```

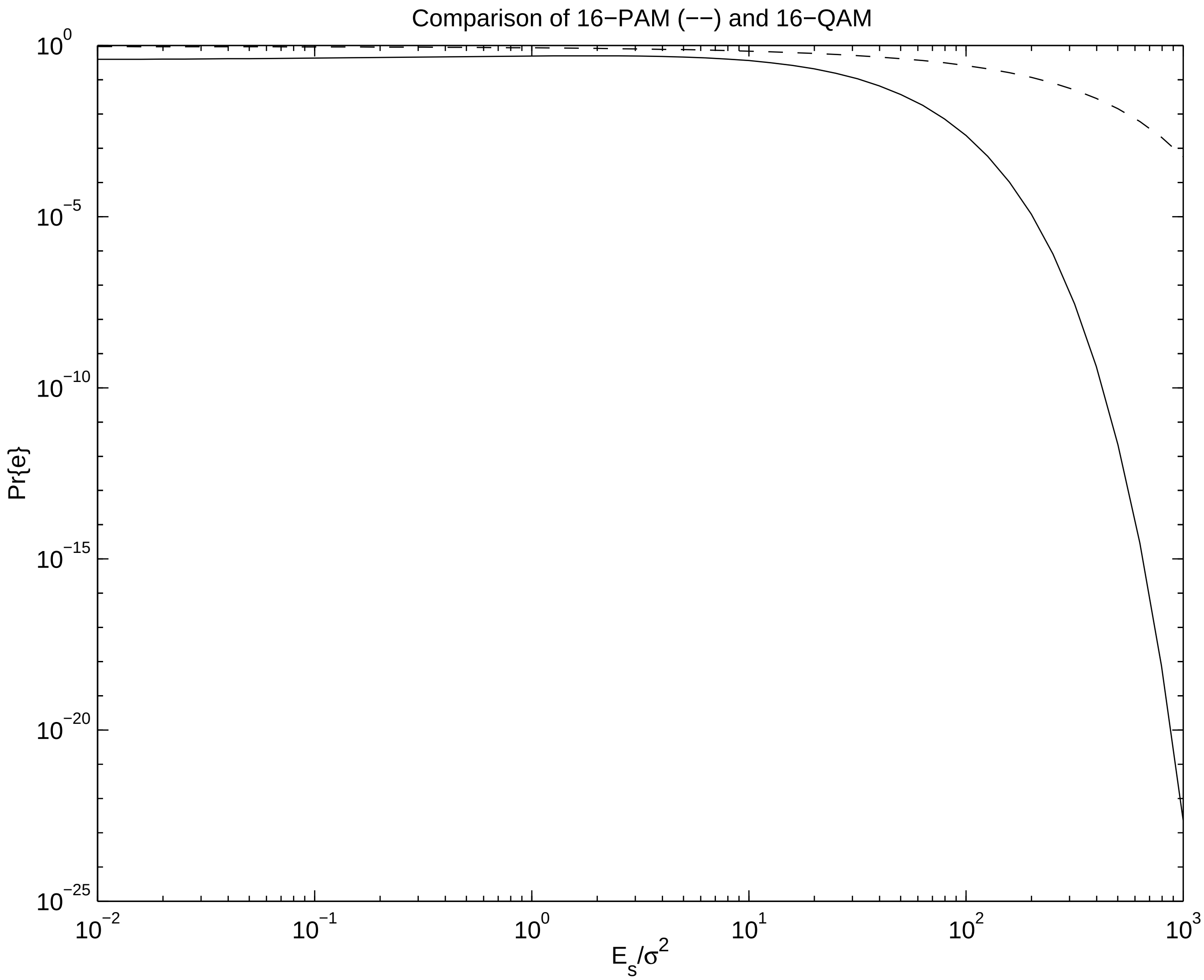
Problem 4. *(Antenna array)*

1. Let $\mathbf{Y} = [Y_1, Y_2]^T$. Then

$$f_{\mathbf{Y}|H_0}(\mathbf{y}) = \frac{1}{2\pi\sigma_1\sigma_2}\exp\left[-\frac{(y_1-A)^2}{2\sigma_1^2} - \frac{(y_2-A)^2}{2\sigma_2^2}\right]$$
$$f_{\mathbf{Y}|H_1}(\mathbf{y}) = \frac{1}{2\pi\sigma_1\sigma_2}\exp\left[-\frac{(y_1+A)^2}{2\sigma_1^2} - \frac{(y_2+A)^2}{2\sigma_2^2}\right].$$

The MAP decision rule is

$$\frac{f_{\mathbf{Y}|H_0}(\mathbf{y})}{f_{\mathbf{Y}|H_1}(\mathbf{y})} \underset{\hat{H}_1}{\overset{\hat{H}_0}{\gtrless}} \frac{P_H(1)}{P_H(0)}$$



or equivalently,

$$\mathrm{LLR}(\mathbf{y}) \underset{\hat{H}_1}{\overset{\hat{H}_0}{\gtrless}} \ln\left[\frac{P_H(1)}{P_H(0)}\right]$$

where LLR is the log-likelihood ratio. In this particular case the optimal decision rule is

$$\ln\left[\frac{f_{\mathbf{Y}|H_0}(\mathbf{y})}{f_{\mathbf{Y}|H_1}(\mathbf{y})}\right] \underset{\hat{H}_1}{\overset{\hat{H}_0}{\gtrless}} \ln\left[\frac{P_H(1)}{P_H(0)}\right] \text{ or equivalently,}$$

$$\frac{2Ay_1}{\sigma_1^2} + \frac{2Ay_2}{\sigma_2^2} \underset{\hat{H}_1}{\overset{\hat{H}_0}{\gtrless}} 0 \text{ or equivalently,}$$

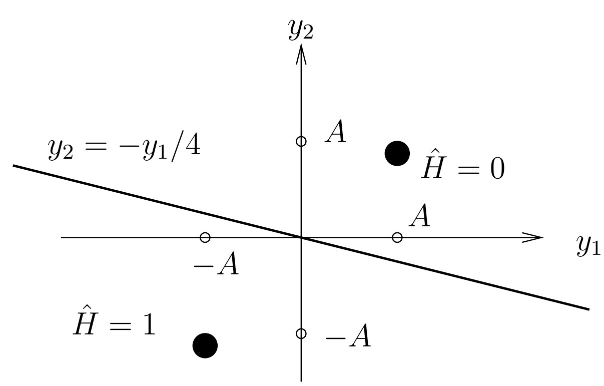
$$\sigma_2^2 y_1 + \sigma_1^2 y_2 \underset{\hat{H}_1}{\overset{\hat{H}_0}{\gtrless}} 0$$

2. When $\sigma_1 = 2\sigma_2$, the decision rule becomes

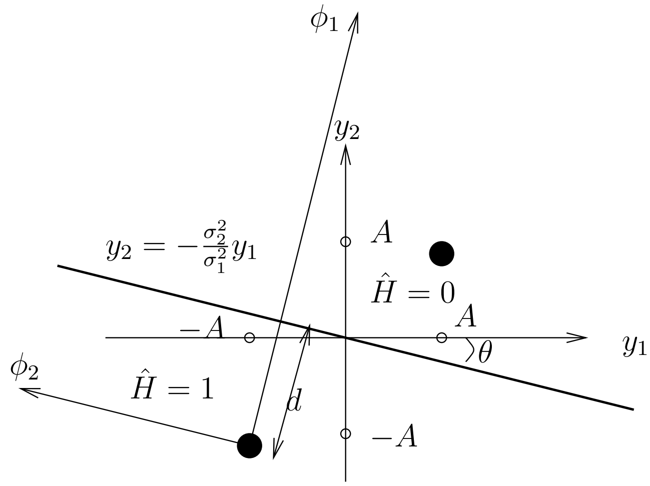
$$\sigma_2^2 y_1 + 4\sigma_2^2 y_2 \underset{\hat{H}_1}{\overset{\hat{H}_0}{\gtrless}} 0 \text{ or equivalently,}$$

$$y_2 \underset{\hat{H}_1}{\overset{\hat{H}_0}{\gtrless}} -\frac{y_1}{4}.$$

The decision regions are sketched below.



3. We work out two solutions. The first solution (the longer one) consists of finding the probability that $\mathbf{Y} = [Y_1, Y_2]^T \in \mathcal{R}_0$ when $H = 1$.



The projection of noise along ϕ_1 is $z_1 \sin\theta + z_2 \cos\theta$ where $\tan\theta = \sigma_2^2/\sigma_1^2$. The noise variance along ϕ_1 is $\sigma_{\phi_1}^2 = (\sin^2\theta)\sigma_1^2 + (\cos^2\theta)\sigma_2^2$. Thus the probability of error is

$$\begin{aligned}
\Pr(\text{error}) &= \frac{1}{2}\Pr(\text{error}|H_0) + \frac{1}{2}\Pr(\text{error}|H_1) \\
&= \Pr(\text{error}|H_1) \\
&= Q\left(\frac{d}{\sigma_{\phi_1}}\right). \qquad (38)
\end{aligned}$$

A little calculation shows that $d = \sqrt{2}A\cos(\pi/4 - \theta)$. We thus have

$$\begin{aligned}
\frac{d}{\sigma_{\phi_1}} &= \frac{\sqrt{2}A\{\cos(\pi/4)\cos\theta + \sin(\pi/4)\sin\theta\}}{\sqrt{(\sin^2\theta)\sigma_1^2 + (\cos^2\theta)\sigma_2^2}} \\
&= \frac{A(1+\tan\theta)}{\sqrt{1+\tan^2\theta}}.
\end{aligned}$$

Substituting $\tan\theta = \sigma_2^2/\sigma_1^2$ in the above expression, we have

$$\frac{d}{\sigma_{\phi_1}} = A\sqrt{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}}$$

and consequently

$$\begin{aligned}
\Pr(\text{error}) &= Q\left(\frac{d}{\sigma_{\phi_1}}\right) \\
&= Q\left(A\sqrt{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}}\right).
\end{aligned}$$

The second solution consists of finding the probability that $\sigma_2^2 Y_1 + \sigma_1^2 Y_2 > 0$ when $H = 1$. But when $H = 1$, $\sigma_2^2 Y_1 + \sigma_1^2 Y_2 = \sigma_2^2(-A + Z_1) + \sigma_1^2(-A + Z_2)$. We see immediately that this $\sim \mathcal{N}(-A(\sigma_2^2+\sigma_1^2),(\sigma_2^4\sigma_1^2+\sigma_1^4\sigma_2^2))$. Hence,

$$\begin{aligned}\Pr\{\text{error}\} &= \Pr\{\text{error}|H=1\}\\ &= Q\left(\frac{A(\sigma_2^2+\sigma_1^2)}{\sqrt{\sigma_2^4\sigma_1^2+\sigma_1^4\sigma_2^2}}\right)\\ &= Q\left(A\sqrt{\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2}}\right).\end{aligned}$$

Problem 5. *(Sufficient Statistics)*

1. We have

$$\begin{aligned}f_{(Y_1,\ldots,Y_n)|H}(y_1,\ldots,y_n|0) &= \prod_{i=1}^n f_{Y_i|H}(y_i|0)\\ &= \left(\frac{1}{4}\right)^{n-\sum_{i=1}^n y_i}\left(\frac{3}{4}\right)^{\sum_{i=1}^n y_i}\\ &= \left(\frac{1}{4}\right)^{n-T(y_1,\ldots,y_n)}\left(\frac{3}{4}\right)^{n-T(y_1,\ldots,y_n)}\end{aligned}$$

and similarly

$$f_{(Y_1,\ldots,Y_n)|H}(y_1,\ldots,y_n|1) = \left(\frac{1}{4}\right)^{\sum y_i}\left(\frac{3}{4}\right)^{n-\sum y_i} = \left(\frac{1}{4}\right)^{T(y_1,\ldots,y_n)}\left(\frac{3}{4}\right)^{T(y_1,\ldots,y_n)}$$

As we see above, the pdf $f_{Y_1,\ldots,Y_n|H}(\cdot|\cdot)$ is only a function of $T(y_1,\ldots,y_n)$. So $T(y_1,\ldots,y_n)$ has all the necessary information of $(Y_1,\ldots,Y_n)$ to predict H.

2. We have

$$\begin{aligned}f_{Y_1,\ldots,Y_n|H}(y_1,\ldots,y_n|H) &= \prod_{i=1}^n f_{Y_i|H}(y_i|H)\\ &= (1-H)^{\sum_{i=1}^n y_i}H^n = (1-H)^{T(y_1,\ldots,y_n)}H^n\end{aligned}$$

So again the probability $f(y_i,\ldots,y_n|H)$ is only dependent on $T(y_1,\ldots,y_n)$ and we can proceed as in part (1).

3. (a) We can simply check that when $H=0$, $L^H=1$ and the equality holds. Similarly, when $H=1$,

$$\begin{aligned}f(y_1,y_2,\ldots,y_n|H) &= f(y_1,y_2,\ldots,y_n|0)L(y_1,y_2,\ldots,y_n)\\ &= f(y_1,y_2,\ldots,y_n|0)\times\frac{f(y_1,y_2,\ldots,y_n|1)}{f(y_1,y_2,\ldots,y_n|0)}\\ &= f(y_1,y_2,\ldots,y_n|1).\end{aligned}$$

Hence the equality holds for both cases.

(b) We see that $f(y_1, y_2, \ldots, y_n|0)$ is a function of $(y_1, y_2, \ldots, y_n)$ which does not depend on H. The other part, $L(y_1, y_2, \ldots, y_n)^H$, is a function of H and L. Hence it depends on $(y_1, y_2, \ldots, y_n)$ through L. Using Fischer-Neyman factorization we obtain that L is a sufficient statistics.

(c) We know that MAP rule, which uses the likelihood ratio, is the optimal decision rule. this implies that it uses all of the information in the observation to decrease the error probability. The previous part shows that all of the relevant information about H is contained in L. Hence the MAP decision rule must depend on L.