8 Optimal Detection for Additive Noise Channels: 1-D Case

We now derive the optimal demodulator for the waveform channel. From the previous chapter, we have seen that instead of analyzing the waveform channel, we can convert it to an equivalent vector channel. The length of the vector is the same as the size K of the orthonormal basis for the waveforms $s_1(t), s_2(t), \ldots, s_M(t)$. In this chapter, we will assume $K = 1$. This is the case, for example, when we use PAM.

Definition 8.1. Detection Problem: When $K = 1$, our problem under consideration is simply that of **detecting** the scalar message S in the presence of additive noise N. The received signal R is given by

$$R = S + N.$$

- S is selected from an alphabet $\mathcal{S}$ containing M possible values $s^{(1)}$, $s^{(2)}$, ..., $s^{(M)}$.
- $p_S\left(s^{(j)}\right) = P\left[S = s^{(j)}\right] \equiv p_j$.
- S and N are independent.

A detector's job is to guess the value of the channel input S from the value of the received channel output R. We denote this guessed value by $\hat{S}$. An optimal detector is the one that minimizes the (symbol) error probability $P(\mathcal{E}) = P\left[\hat{S} \neq S\right]$.

8.2. The analysis here is very similar to what we have done in Chapter 3. Here, for clarity, we note some important differences:

- In Chapter 3, The channel input and output are denoted by X and Y, respectively. Here, they are denoted by S and R.
- In Chapter 3, the transition probabilities are arbitrary and summarized by the matrix $\mathbf{Q}$. Here, the transition probabilities is basically controlled by the additive noise.
- In Chapter 3, both X and Y are discrete. Here, S is discrete. However, because noise is continuous, R will be a continuous random variable.

Even with these differences, several techniques that we used in Chapter 3 will be applicable here.

Example 8.3. Review: To re-connect with what we studied in Chapter 3, let's try to find the **Q** matrix when the additive noise is discrete. Suppose

$$p_S(s) = \begin{cases} 0.3, & s = -1, \\ 0.7, & s = 1, \\ 0, & \text{otherwise}, \end{cases} \quad \text{and } p_N(n) = \begin{cases} 0.2, & n \in \{-0.5, +0.5\}, \\ 0.6, & n = 0, \\ 0, & \text{otherwise}. \end{cases}$$

Because $R = S + N$, we know that

(a) given $S = -1$, we have $R = -1 + N$:

(b) given $S = 1$, we have $R = 1 + N$:

The **Q** matrix is given by

Note that each row of the **Q** matrix is simply a shifted copy of the noise pmf. The amount of shift is the corresponding value of s for that row.

8.4. Formula-wise, when the additive noise is discrete, each row of the **Q** matrix (as in Example 8.3) is given by

$$p_{R|S}(r|s) = p_N(r - s). \tag{47}$$

8.5. When the additive noise is continuous, there are uncountably many possible values for the channel output R. Hence, representing conditional probabilities in the form of a matrix **Q** does not make sense here.

When R is continuous, the conditional pmf $p_{R|S}(r|s)$ is replaced by the conditional pdf $f_{R|S}(r|s)$. For additive noise N with pdf $f_N(n)$, we have

$$f_{R|S}(r|s) = f_N(r - s). \tag{48}$$

Example 8.6. Suppose the discrete additive noise in Example 8.3 is replaced by a continuous additive noise:

S	R
1	1.3800
-1	-1.7267
1	1.5429
1	1.9886
1	1.1556
1	1.8981
1	0.9856
1	0.5699
1	0.7173
-1	-0.9756
1	1.2335
-1	-1.1320
-1	-0.0612
-1	-2.1877
-1	-0.9035
1	0.6967
1	1.3127
1	0.1829
1	1.4026
-1	-1.6058

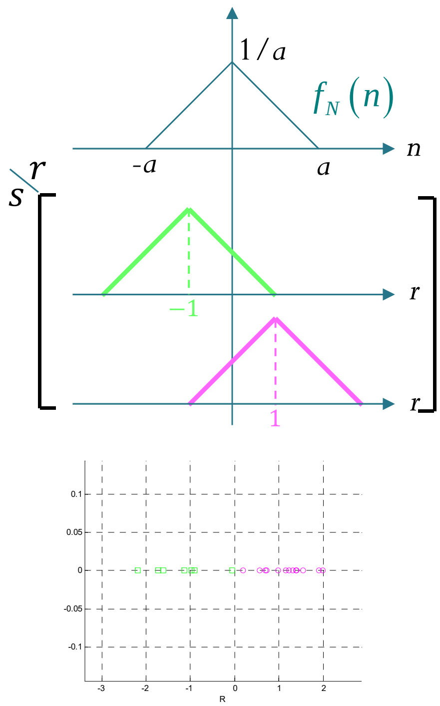

Figure 40: Binary PAM under "Triangular" Noise

8.7. The optimal detector, which minimizes the error probability, is the **MAP detector**:

$$\hat{s}_{\text{MAP}}(r) = \arg\max_{s \in \mathcal{S}} p_S(s) f_{R|S}(r|s) = \arg\max_{s \in \mathcal{S}} p_S(s) f_N(r - s). \tag{49}$$

Because event $[W = j]$ is the same as event $[S = s^{(j)}]$, we also have

$$\hat{w}_{\text{MAP}}(r) = \arg\max_{j \in \{1,2,\dots,M\}} p_j f_N\left(r - s^{(j)}\right). \tag{50}$$

When the prior probabilities are ignored, we have the (sub-optimal) **ML detector**:

$$\hat{s}_{\text{ML}}(r) = \arg\max_{s \in \mathcal{S}} f_{R|S}(r|s) = \arg\max_{s \in \mathcal{S}} f_N(r - s). \tag{51}$$

and

$$\hat{w}_{\mathrm{ML}}(r) = \underset{j\in\{1,2,\ldots,M\}}{\arg\max}\ f_N\left(r - s^{(j)}\right). \tag{52}$$

8.8. Graphically, here are the steps to find the MAP detector:

(a) Plot $p_1 f_N(r - s^{(1)})$, $p_2 f_N(r - s^{(2)})$, ..., $p_M f_N(r - s^{(M)})$.

- Note that they are functions of r.
- This is similar to scaling the rows of the **Q** matrix by the corresponding prior probabilities in Chapter 3 to get the **P** matrix.

(b) Select the maximum plot for each (observed) r value.

- If there are multiple max values, select any.
- The corresponding $s^{(j)}$ is the value of $\hat{s}_{\mathrm{MAP}}$ at r.

Example 8.9. Back to Example 8.6.

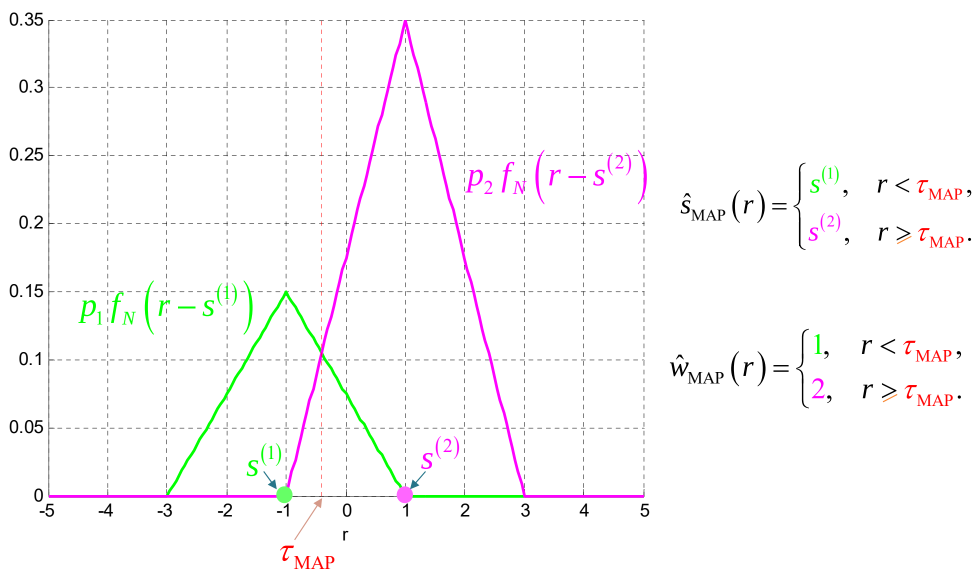

Figure 41: MAPD for Binary PAM under "Triangular" Noise

Definition 8.10. The ith **decision "region"**, denoted by $\mathcal{D}_i$ for a decoder $\hat{s}(r)$ is defined as the collection of all the r values at which r is decoded as $s^{(i)}$.

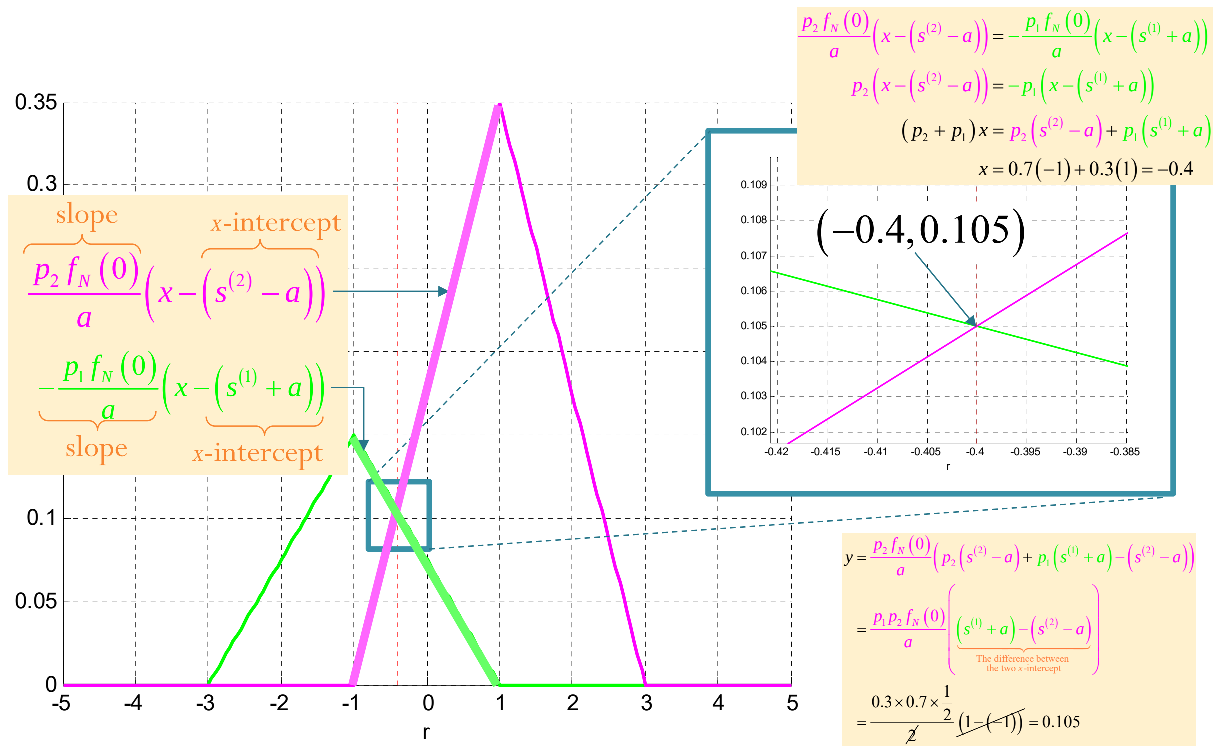

Figure 42: Solving for τ_{MAP} in MAPD for Binary PAM under "Triangular" Noise

- The collection $\mathcal{D}_1$, $\mathcal{D}_2$, …, $\mathcal{D}_M$ should partition the whole observable values (support) of R.

Example 8.11. Back to Example 8.6.

$$\hat{s}_{MAP}(r)=\begin{cases} s^{(1)}, & r\in\mathcal{D}_1, \\ s^{(2)}, & r\in\mathcal{D}_2, \end{cases} \quad \text{where} \quad \begin{aligned} \mathcal{D}_1 &= (-\infty,\tau_{\text{MAP}}), \\ \mathcal{D}_2 &= [\tau_{\text{MAP}},\infty). \end{aligned}$$

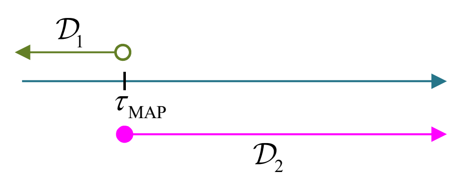

Figure 43: Decision Regions in MAPD for Binary PAM under "Triangular" Noise

8.12. The error probability of a detector can be found via its success probability

$$
\begin{aligned}
P(\mathcal{C}) &= \sum_{i=1}^{M} P\left(\mathcal{C} | S = s^{(i)}\right) P\left[S = s^{(i)}\right] = \sum_{i=1}^{M} P\left[R \in D_i \middle| S = s^{(i)}\right] p_i \\
&= \sum_{i=1}^{M} p_i P\left[S + N \in D_i \middle| S = s^{(i)}\right] = \sum_{i=1}^{M} p_i P\left[N + s^{(i)} \in D_i\right] \\
&= \sum_{i=1}^{M} p_i \int_{D_i} f_N\left(r - s^{(i)}\right) dr = \sum_{i=1}^{M} \int_{D_i} p_i f_N\left(r - s^{(i)}\right) dr.
\end{aligned}
$$

This gives

$$
\begin{aligned}
P(\mathcal{E}) &= 1 - P(\mathcal{C}) \\
&= \sum_{i=1}^{M} p_i \int_{D_i^c} f_N\left(r - s^{(i)}\right) dr = \sum_{i=1}^{M} \int_{D_i^c} p_i f_N\left(r - s^{(i)}\right) dr.
\end{aligned}
$$

Although, at first, the above expressions may look complicated, it is similar to what we did in Chapter 3: graphically, the area under the max (selected) plot is $P(\mathcal{C})$.

Example 8.13. Back to Example 8.6.

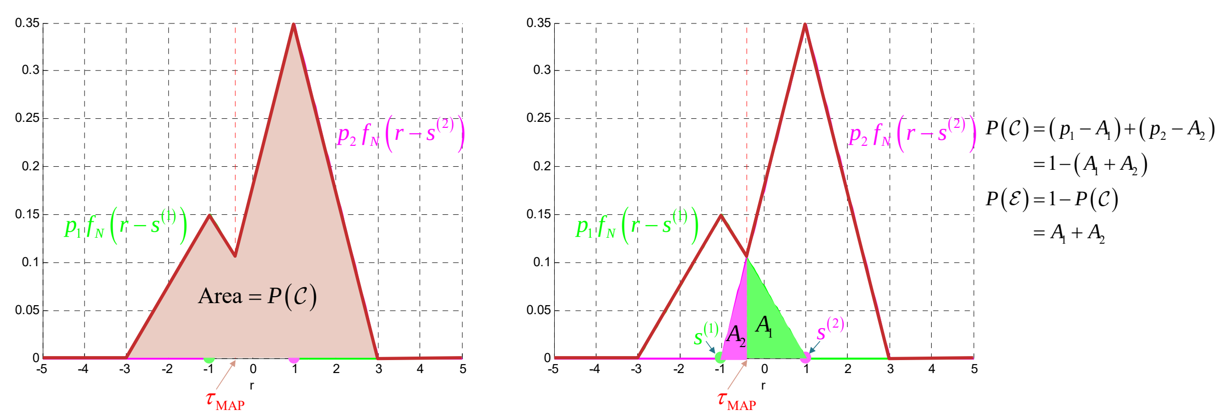

Figure 44: Probability of Successful Detection for Binary PAM under “Triangular” Noise