Advanced Digital Communications (EQ2410)

Additional Material for Lecture 3, Period 2, 2016

Performance Analysis for Equalization with the MLSE

Background

Assume in the following that A and B denote two discrete random variables with $A \in \mathcal{A}$ and $B \in \mathcal{B}$. $\Pr(A = a)$ and $\Pr(B = b)$ denote the probabilities for A and B, respectively, and the joint probability is given by $\Pr(A = a, B = b)$. From probability theory we know the following:

- *Complete probability:*
$$\Pr(A = a) = \sum_{b \in \mathcal{B}} \Pr(A = a, B = b) \tag{1}$$

- *Conditional probability:*
$$\Pr(A = a | B = b) = \frac{\Pr(A = a, B = b)}{\Pr(B = b)} \tag{2}$$

Derivation of the Error Probability

We are interested in deriving the error probability $P_e(k)$ for the k-th bit under maximum likelihood sequence estimation (MLSE). The error probability is defined as

$$P_e(k) = \Pr(b[k] \neq \hat{b}_{ML}[k]), \tag{3}$$

where $\hat{b}_{ML}[k]$ denotes the k-th element of the ML sequence $\hat{\boldsymbol{b}}_{\boldsymbol{ML}}$ and $b[k]$ denotes the k-th element of the transmitted sequence $\boldsymbol{b}$. Let furthermore

- $\mathcal{E}$ denote the set of all existing error sequences $\boldsymbol{e}$,
- $\mathcal{E}_k$ denote the set of all existing error sequences $\boldsymbol{e}$ which lead to an error event at symbol position k, and
- $\mathcal{B}$ denote the set of all possible symbol sequences $\boldsymbol{b}$ (with length N).

The derivation is presented in four steps: in the first step, the error probability is expressed in terms of the pairwise error probability conditioned on a valid error sequence and the probability that an error sequence is valid. In step 2 and 3, the latter probabilities are derived. In the final step, all results are combined.

Step 1: Decomposition of $P_e(k)$

The derivation of $P_{e(k)}$ is shown in the following, and the steps are explained below:

$$\begin{aligned}
P_e(k) &= \Pr(b[k] \neq \hat{b}_{ML}[k]) && (4)\\
&= \sum_{\boldsymbol{e} \in \mathcal{E}} \sum_{\boldsymbol{b} \in \mathcal{B}} \Pr(b[k] \neq \hat{b}_{ML}[k], \boldsymbol{e}, \boldsymbol{b}) && (5)\\
&= \sum_{\boldsymbol{e} \in \mathcal{E}_k} \sum_{\boldsymbol{b} \in \mathcal{B}} \Pr(\hat{\boldsymbol{b}}_{\boldsymbol{ML}} = \boldsymbol{b} - 2\boldsymbol{e}, \boldsymbol{e} \text{ is valid for } \boldsymbol{b}, \boldsymbol{e}, \boldsymbol{b}) && (6)\\
&= \sum_{\boldsymbol{e} \in \mathcal{E}_k} \sum_{\boldsymbol{b} \in \mathcal{B}} \Pr(\hat{\boldsymbol{b}}_{\boldsymbol{ML}} = \boldsymbol{b} - 2\boldsymbol{e} | \boldsymbol{e} \text{ is valid for } \boldsymbol{b}, \boldsymbol{e}, \boldsymbol{b}) \Pr(\boldsymbol{e} \text{ is valid for } \boldsymbol{b}, \boldsymbol{e}, \boldsymbol{b}) && (7)\\
&\leq \sum_{\boldsymbol{e} \in \mathcal{E}_k} Q\left(\frac{\|s(\boldsymbol{e})\|}{\sigma}\right) \sum_{\boldsymbol{b} \in \mathcal{B}} \Pr(\boldsymbol{e} \text{ is valid for } \boldsymbol{b}, \boldsymbol{e}, \boldsymbol{b}) && (8)
\end{aligned}$$

1. Equation (5) follows directly from (1). We are marginalizing over all existing error sequences $\boldsymbol{e} \in \mathcal{E}$ and all hypotheses for the transmitted symbol sequence $\boldsymbol{b} \in \mathcal{B}$.

2. Error sequences $\boldsymbol{e}$ which do not lead to error events at time k will lead to zero probability in the argument of the sum in Equation (6). Therefore, we can replace the sum over $\boldsymbol{e} \in \mathcal{E}$ by $\boldsymbol{e} \in \mathcal{E}_k$. The event "$b[k] \neq \hat{b}_{ML}[k]$" is then equivalent to the event "$\hat{\boldsymbol{b}}_{\boldsymbol{ML}} = \boldsymbol{b} - 2\boldsymbol{e}, \boldsymbol{e}$ is valid for $\boldsymbol{b}$" and can be substituted in the argument of the probability.

3. Equation (7) follows directly from (2).

4. In Equation (8), we use the fact that

$$\Pr(\hat{\boldsymbol{b}}_{\boldsymbol{ML}} = \boldsymbol{b} - 2\boldsymbol{e} | \boldsymbol{e} \text{ is valid for } \boldsymbol{b}, \boldsymbol{e}, \boldsymbol{b}) \leq Q\left(\frac{\|s(\boldsymbol{e})\|}{\sigma}\right),$$

as shown in the book. Using furthermore the fact that this term does not depend on $\boldsymbol{b}$, we arrive at Equation (8).

Step 2: Probability for a Valid Error Sequence

We continue with the derivation of the second term in the sum in Equation (8). It corresponds to the probability of a valid error sequence, and the steps are explained below.

$$\sum_{\boldsymbol{b} \in \mathcal{B}} \Pr(\boldsymbol{e} \text{ is valid for } \boldsymbol{b}, \boldsymbol{e}, \boldsymbol{b}) = \sum_{\boldsymbol{b} \in \mathcal{B}} \Pr(\boldsymbol{e} \text{ is valid for } \boldsymbol{b}, \boldsymbol{e} | \boldsymbol{b}) \Pr(\boldsymbol{b}) \tag{9}$$

$$= \sum_{\boldsymbol{b} \in \mathcal{B}} \prod_{i=1}^{N} \Pr(e[i], e[i] \text{ is valid for } b[i] \,|\, b[i]) \Pr(b[i]) \tag{10}$$

$$= \sum_{\substack{\boldsymbol{b} \in \mathcal{B}, \\ \boldsymbol{e} \text{ is valid for } \boldsymbol{b}}} \prod_{\substack{i=1, \\ e[i]=0}}^{N} \Pr(b[i]) \prod_{\substack{i=1, \\ e[i]\neq 0}}^{N} \Pr(b[i] = e[i]) \tag{11}$$

$$= \prod_{\substack{i=1, \\ e[i]\neq 0}}^{N} \Pr(b[i] = e[i]) \sum_{\substack{\boldsymbol{b} \in \mathcal{B}, \\ \boldsymbol{e} \text{ is valid for } \boldsymbol{b}}} \prod_{\substack{i=1, \\ e[i]=0}}^{N} \Pr(b[i]) \tag{12}$$

$$= 2^{-w(\boldsymbol{e})} \sum_{\substack{\boldsymbol{b} \in \mathcal{B}, \\ \boldsymbol{e} \text{ is valid for } \boldsymbol{b}}} 2^{-(N-w(\boldsymbol{e}))} \tag{13}$$

$$= 2^{-w(\boldsymbol{e})} \tag{14}$$

1. Equation (9) follows directly from (2).

2. Since the transmitted symbols $b[i]$ in $\boldsymbol{b}$ are independent and the error symbols $e[i]$ in $\boldsymbol{e}$ are independent, the probabilities for the sequences in Equation (9) can be expressed as products of the probabilities of the respective symbols, as shown in Equation (10). Note that we introduced N as the lengths of the sequences $\boldsymbol{e}, \boldsymbol{b}$.

3. In Equation (11), we made use of the observation that the event "$e[i], e[i]$ is valid for $b[i] \,|\, b[i]$" is deterministic; i.e.,

$$\Pr(e[i], e[i] \text{ is valid for } b[i] \,|\, b[i]) = \begin{cases} 1, & \text{if } e[i] = 0 \text{ or } e[i] = b[i], \\ 0, & \text{if } e[i] \neq 0 \text{ and } e[i] \neq b[i]. \end{cases}$$

Since this "probability" is zero whenever $e[i] \neq 0$ and $e[i] \neq b[i]$, it is sufficient to sum only over those bit sequences $\boldsymbol{b}$ for which $\boldsymbol{e}$ is valid for $\boldsymbol{b}$.

4. Since all sequences $\boldsymbol{b} \in \mathcal{B}$ for which "$\boldsymbol{e}$ is valid for $\boldsymbol{b}$" holds have identical bits $b[i]$ at positions i where $e[i] \neq 0$ (namely $b[i] = e[i]$), the second product in Equation (11) is a constant w.r.t. the summation. It can be put in front of the sum, as shown in Equation (12).

5. Under the assumption of uniformly distributed bits (i.e., $\Pr(b[i] = +1) = \Pr(b[i] = -1) = 1/2$) and by expressing the Hamming weight[1] of the error sequences as $w(\boldsymbol{e})$, we can simplify Equation (12) to Equation (13).

6. The cardinality of the set $\mathcal{B}' = \{\boldsymbol{b} \in \mathcal{B} : \boldsymbol{e} \text{ is valid for } \boldsymbol{b}\}$ is $|\mathcal{B}'| = 2^{N-w(\boldsymbol{e})}$ (the bit sequences in $\mathcal{B}'$ differ in $N - w(\boldsymbol{e})$ bit positions and are identical in $w(\boldsymbol{e})$ bit positions). Since the argument in the sum in Equation (13) is constant, it is simple to verify that it sums up to 1. What remains in Equation (14) is the result given in the book.

Step 3: Pairwise Error Probability Conditioned on an Error Sequence

What remains to be shown is the relation

$$\Pr(\hat{\boldsymbol{b}}_{\boldsymbol{ML}} = \boldsymbol{b} - 2\boldsymbol{e} | \boldsymbol{e} \text{ is valid for } \boldsymbol{b}, \boldsymbol{e}, \boldsymbol{b}) \leq Q\left(\frac{\|s(\boldsymbol{e})\|}{\sigma}\right). \tag{15}$$

In a first step, we can bound the term on the LHS of the previous equation by the pairwise error probability for detecting the signal $s(\boldsymbol{b} - 2\boldsymbol{e})$ assuming that $s(\boldsymbol{b})$ was transmitted. We get

$$\begin{aligned}
\Pr(\hat{\boldsymbol{b}}_{\boldsymbol{ML}} = \boldsymbol{b} - 2\boldsymbol{e} | \boldsymbol{e} \text{ is valid for } \boldsymbol{b}, \boldsymbol{e}, \boldsymbol{b}) &= \Pr(\boldsymbol{b} - 2\boldsymbol{e} = \arg\max_{\boldsymbol{a}} \Lambda(\boldsymbol{a}) | \boldsymbol{e} \text{ is valid for } \boldsymbol{b}, \boldsymbol{e}, \boldsymbol{b}) \\
&\leq \Pr(\Lambda(\boldsymbol{b} - 2\boldsymbol{e}) > \Lambda(\boldsymbol{b}) | \boldsymbol{e} \text{ is valid for } \boldsymbol{b}, \boldsymbol{e}, \boldsymbol{b}).
\end{aligned}$$

From Chapter 3.5.1 in the textbook it is however known that error probability for two signals $s_1(t)$ and $s_2(t)$ in Gaussian noise is given as

$$P_{e,pw} = Q\left(\frac{\|s_1 - s_2\|}{2\sigma}\right).$$

Replacing s_1 and s_2 by $s(\boldsymbol{b} - 2\boldsymbol{e})$ and $s(\boldsymbol{b})$ and taking into account that in our case $s(\boldsymbol{b} - 2\boldsymbol{e}) = s(\boldsymbol{b}) - 2s(\boldsymbol{e})$ proves Equation (15).

Step 4: Combining the Results

Combining the results from Equations (8), (14), and (15) leads to the results from the textbook,

$$P_e(k) \leq \sum_{\boldsymbol{e} \in \mathcal{E}_k} Q\left(\frac{\|s(\boldsymbol{e})\|}{\sigma}\right) 2^{-w(\boldsymbol{e})}.$$

[1] The Hamming weight $w(\boldsymbol{e})$ of the error sequence $\boldsymbol{e}$ gives the number of non-zero entries in $\boldsymbol{e}$.