

# ROBUSTNESS OF THE FINITE-LENGTH MMSE-DFE WITH RESPECT TO CHANNEL AND SECOND-ORDER STATISTICS ESTIMATION ERRORS

Athanasios P. Liavas

Department of Mathematics, University of the Aegean, Karlovassi, 83200, GREECE

E-mail: liavas@aegean.gr

**Abstract** - The filters of the finite-length minimum mean-square error decision-feedback equalizer (MMSE-DFE) can be computed by assuming perfect knowledge of the channel impulse response and the input and noise second-order statistics. In practice, we estimate the unknown quantities and thus inevitable estimation errors arise. In this work, we model the estimation errors as small perturbations and we derive a second-order approximation to the excess MSE. Then, assuming that the input and noise SOS are perfectly known, we derive an expression for the mean excess MSE in terms of the channel estimation error covariance matrix. Analogous expressions involving the noise and input SOS estimation error covariance matrices appear on [1].

## 1. INTRODUCTION

The finite-length MMSE-DFE has proved to be an efficient structure toward intersymbol interference (ISI) mitigation in packet-based communication systems [2]. It is determined by two optimal filters, namely, the feedforward and the feedback filter. These filters, as well as related performance measures, can be computed by assuming perfect knowledge of the channel impulse response and the input and additive channel noise second-order statistics (SOS) [2]. In practice, we estimate the unknown quantities and thus inevitable estimation errors arise. Consequently, the analysis of the robustness of the finite-length MMSE-DFE with respect to mismatch is of great importance.

This problem was first considered in [3], where the authors developed closed-form expressions for the perturbed MMSE-DFE filters and the corresponding performance measures. In this work, we present a detailed second-order perturbation analysis, that explicitly reveals the factors that govern the performance of the MMSE-DFE under mismatch.

## 2. FINITE-LENGTH MMSE-DFE

In this section, we recall known results concerning the finite-length MMSE-DFE [2]. We assume that the channel impulse response and the input and noise SOS are perfectly known.

### 2.1. Channel model

Let us consider a baseband discrete-time fractionally-spaced noisy communication channel modeled by the  $\nu$ -th order 1-input/ $p$ -output

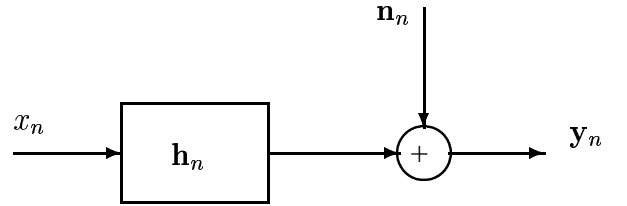


Fig. 1. Channel model.

linear time-invariant system depicted in Fig. 1. Its input-output relation is given by the convolution  $y_n = \sum_{i=0}^{\nu} \mathbf{h}_i x_{n-i} + \mathbf{n}_n$ , where  $x_n$  denotes the input sequence and the  $p$ -dimensional vectors  $\mathbf{y}_n$ ,  $\mathbf{n}_n$  and  $\mathbf{h}_i$  denote, respectively, the terms of the output, noise and channel finite impulse response sequences. The impulse response terms  $\mathbf{h}_i$  are vectors composed of the samples of the continuous-time impulse response modeling the combined effect of the transmit filter, the physical channel and the receiver filter [2]. By grouping the impulse response terms, we construct the impulse response vector  $\mathcal{H}_\nu \triangleq [\mathbf{h}_0^T \dots \mathbf{h}_\nu^T]^T$ , where superscript  $T$  denotes transpose.

By stacking  $N_f$  successive output samples, we construct the data vector  $\mathbf{y}_{n:n-N_f+1} \triangleq [\mathbf{y}_n^T \dots \mathbf{y}_{n-N_f+1}^T]^T$ , which can be expressed as  $\mathbf{y}_{n:n-N_f+1} = \mathbf{H} \mathbf{x}_{n:n-N_f-\nu+1} + \mathbf{n}_{n:n-N_f+1}$  where the  $pN_f \times (\nu + N_f)$  filtering matrix  $\mathbf{H}$  is defined as

$$\mathbf{H} \triangleq \begin{bmatrix} \mathbf{h}_0 & \dots & \dots & \mathbf{h}_\nu \\ & & \ddots & \\ & & & \mathbf{h}_0 & \dots & \dots & \mathbf{h}_\nu \end{bmatrix}$$

and the definitions of  $\mathbf{x}_{n:n-N_f-\nu+1}$  and  $\mathbf{n}_{n:n-N_f+1}$  are obvious.

### 2.2. Finite-length MMSE-DFE

Our aim is to recover (a delayed version of) the input sequence  $x_n$  by passing the noisy output data  $\mathbf{y}_n$  through an equalizer structure. To this end, we employ the finite-length DFE depicted in Fig. 2. The DFE is determined by the following parameter vectors:

1.  $\mathbf{w} \triangleq [\mathbf{w}_0^T \dots \mathbf{w}_{N_f-1}^T]^T$ , which denotes the  $p$ -input/1-output length- $N_f$  feedforward filter;
2.  $\mathbf{b} \triangleq [1 \ b_1 \ \dots \ b_{N_b}]^T$ , which determines the 1-input/1-output length- $N_b$  strictly causal feedback filter. The settings of the feedback filter are  $\{-b_1, \dots, -b_{N_b}\}$ .

Assuming that *the past decisions are correct* and considering the delay  $\Delta$ , the error between the desired output  $x_{n-\Delta}$  and the input to

This work was supported in part by the EPETII Program of the Greek Secretariat for Research and Technology.

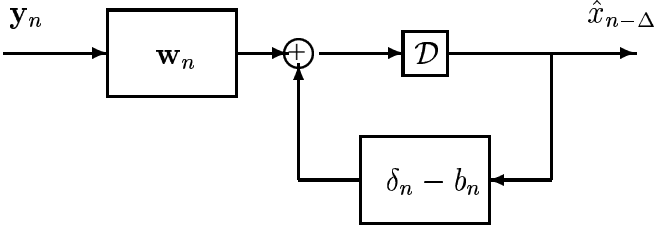


Fig. 2. Finite-length DFE.

the decision device  $\mathcal{D}$  is given by

$$\begin{aligned}
 e_n &\triangleq x_{n-\Delta} - \left( \sum_{i=0}^{N_f-1} \mathbf{w}_i^T \mathbf{y}_{n-i} - \sum_{i=1}^{N_b} b_i x_{n-\Delta-i} \right) \\
 &= \mathbf{b}^T \mathbf{x}_{n-\Delta:n-\Delta-N_b} - \mathbf{w}^T \mathbf{y}_{n:n-N_f+1} \\
 &= \tilde{\mathbf{b}}^T \mathbf{x}_{n:n-N_f-\nu+1} - \mathbf{w}^T \mathbf{y}_{n:n-N_f+1}
 \end{aligned} \quad (1)$$

where we defined  $\tilde{\mathbf{b}} \triangleq [\mathbf{0}_{1 \times \Delta} \ \mathbf{b}^T \ \mathbf{0}_{1 \times s}]^T$ , with  $\mathbf{0}_{i \times j}$  denoting the  $i \times j$  zero matrix and  $s \triangleq N_f + \nu - \Delta - N_b - 1$ . In order to simplify notation, we shall omit the subscripts from  $\mathbf{x}_{n:n-N_f-\nu+1}$ ,  $\mathbf{y}_{n:n-N_f+1}$  and  $\mathbf{n}_{n:n-N_f+1}$ .

The MMSE-DFE settings are computed by minimizing the mean square error (MSE)  $\mathcal{E}[e_n^2]$ , which can be expressed as

$$\begin{aligned}
 \text{MSE} &\triangleq \mathcal{E} \left[ (\tilde{\mathbf{b}}^T \mathbf{x} - \mathbf{w}^T \mathbf{y}) (\mathbf{x}^T \tilde{\mathbf{b}} - \mathbf{y}^T \mathbf{w}) \right] \\
 &= \tilde{\mathbf{b}}^T \mathbf{R}_{xx} \tilde{\mathbf{b}} - \tilde{\mathbf{b}}^T \mathbf{R}_{xy} \mathbf{w} - \mathbf{w}^T \mathbf{R}_{yx} \tilde{\mathbf{b}} + \mathbf{w}^T \mathbf{R}_{yy} \mathbf{w}
 \end{aligned} \quad (2)$$

where

$$\mathbf{R}_{xx} \triangleq \mathcal{E}[\mathbf{x}\mathbf{x}^T], \quad \mathbf{R}_{xy} \triangleq \mathcal{E}[\mathbf{x}\mathbf{y}^T] = \mathbf{R}_{xx} \mathbf{H}^T = \mathbf{R}_{yy}^T \quad (3)$$

$$\mathbf{R}_{yy} \triangleq \mathcal{E}[\mathbf{y}\mathbf{y}^T] = \mathbf{H}\mathbf{R}_{xx}\mathbf{H}^T + \mathbf{R}_{nn}, \quad \mathbf{R}_{nn} \triangleq \mathcal{E}[\mathbf{n}\mathbf{n}^T]. \quad (4)$$

At the optimal settings, the error  $e_n$  is uncorrelated with the data vector  $\mathbf{y}$ , i.e.,  $\mathcal{E}[e_n \mathbf{y}^T] = \mathbf{0}_{1 \times p N_f}$ , yielding [2]  $\mathbf{w} = \mathbf{R}_{yy}^{-1} \mathbf{R}_{yx} \tilde{\mathbf{b}}$ . Substituting this expression for  $\mathbf{w}$  into (2), we obtain  $\text{MSE} = \tilde{\mathbf{b}}^T \mathbf{R} \tilde{\mathbf{b}}$ , where

$$\mathbf{R} \triangleq \mathbf{R}_{xx} - \mathbf{R}_{xy} \mathbf{R}_{yy}^{-1} \mathbf{R}_{yx}. \quad (5)$$

If we define

$$\mathbf{R}_\Delta \triangleq \begin{bmatrix} \mathbf{0}_{\Delta \times (N_b+1)} & \\ & \mathbf{I}_{N_b+1} \\ & & \mathbf{0}_{s \times (N_b+1)} \end{bmatrix} \mathbf{R}$$

where  $\mathbf{I}_i$  denotes the  $i \times i$  identity matrix, then the MSE is given by  $\text{MSE} = \mathbf{b}^T \mathbf{R}_\Delta \mathbf{b}$  and it can be shown [2] that it is minimized for

$$\mathbf{b}_o = \frac{\mathbf{R}_\Delta^{-1} \mathbf{e}_o}{\mathbf{e}_o^T \mathbf{R}_\Delta^{-1} \mathbf{e}_o}, \quad \mathbf{w}_o = \mathbf{R}_{yy}^{-1} \mathbf{R}_{yx} \tilde{\mathbf{b}}_o \quad (6)$$

where  $\mathbf{e}_o$  is the vector with 1 at its first position and zeros elsewhere. The corresponding minimum MSE,  $\mathcal{M}$ , is derived by putting in (2)  $\tilde{\mathbf{b}} = \tilde{\mathbf{b}}_o$  and  $\mathbf{w} = \mathbf{w}_o$ . Alternative expressions are

$$\mathcal{M} = \tilde{\mathbf{b}}_o^T \mathbf{R} \tilde{\mathbf{b}}_o = \mathbf{b}_o^T \mathbf{R}_\Delta \mathbf{b}_o = \frac{1}{\mathbf{e}_o^T \mathbf{R}_\Delta^{-1} \mathbf{e}_o}. \quad (7)$$

### 3. MMSE-DFE: PERFORMANCE ANALYSIS UNDER MISMATCH

#### 3.1. The framework

Let us assume that an estimation procedure has furnished the estimates  $\{\hat{\mathbf{h}}_i\}_{i=0}^{\hat{\nu}}$ , leading to the impulse response vector estimate  $\hat{\mathcal{H}}_{\hat{\nu}} \triangleq [\hat{\mathbf{h}}_0^T \cdots \hat{\mathbf{h}}_{\hat{\nu}}^T]^T$ . We consider the case  $\hat{\nu} < \nu$ , that is, the estimated channel order  $\hat{\nu}$  is less than the true channel order  $\nu$ . The case  $\hat{\nu} \geq \nu$  is analogous. For the purposes of analysis, we have to compare the true and estimated impulse response vectors,  $\mathcal{H}_\nu$  and  $\hat{\mathcal{H}}_{\hat{\nu}}$ . Since  $\hat{\nu} < \nu$ , we take into account the fact that, due to pulse shaping and dispersive effects in the transmission medium,  $\mathcal{H}_\nu$  is usually composed of tails of small leading and trailing terms, and we augment the estimated impulse response vector  $\hat{\mathcal{H}}_{\hat{\nu}}$  with leading and trailing zeros, obtaining the vector

$$\hat{\mathcal{H}}_{\hat{\nu}}^{m_1} \triangleq \begin{bmatrix} \mathbf{0}_{1 \times pm_1} & \hat{\mathcal{H}}_{\hat{\nu}}^T & \mathbf{0}_{1 \times p(\nu - \hat{\nu} - m_1)} \end{bmatrix}^T$$

whose length equals the length of  $\mathcal{H}_\nu$ . Then, we define

$$m_1^* \triangleq \arg \min_{m_1} \|\mathcal{H}_\nu - \hat{\mathcal{H}}_{\hat{\nu}}^{m_1}\|_2,$$

where  $\|\cdot\|_2$  denotes, depending on the argument, the matrix or vector 2-norm. That is,  $pm_1^*$  is the number of leading zeros we must insert in front of  $\hat{\mathcal{H}}_{\hat{\nu}}$ , so that the augmented impulse response vector estimate becomes closest to  $\mathcal{H}_\nu$ . In the sequel, we shall work with  $\hat{\mathcal{H}}_\nu \triangleq \hat{\mathcal{H}}_{\hat{\nu}}^{m_1^*}$ . We note that working with  $\hat{\mathcal{H}}_\nu$  instead of  $\hat{\mathcal{H}}_{\hat{\nu}}$  amounts to insertion of an extra delay of  $m_1^*$  time units. We consider our channel estimate as being good if  $\mathcal{H}_\nu$  and  $\hat{\mathcal{H}}_\nu$  are close to each other. In terms of the associated filtering matrices, we express this condition as:

$$\Delta \mathbf{H} \triangleq \hat{\mathbf{H}} - \mathbf{H}, \quad \|\Delta \mathbf{H}\|_2 \ll \|\mathbf{H}\|_2.$$

The errors in the input and noise SOS can be expressed as:

$$\Delta \mathbf{R}_{xx} \triangleq \hat{\mathbf{R}}_{xx} - \mathbf{R}_{xx}, \quad \|\Delta \mathbf{R}_{xx}\|_2 \ll \|\mathbf{R}_{xx}\|_2$$

$$\Delta \mathbf{R}_{nn} \triangleq \hat{\mathbf{R}}_{nn} - \mathbf{R}_{nn}, \quad \|\Delta \mathbf{R}_{nn}\|_2 \ll \|\mathbf{R}_{nn}\|_2.$$

#### 3.2. MMSE-DFE: Perturbation analysis

Under mismatch, efforts toward computation of  $\mathbf{R}_{xy}$ ,  $\mathbf{R}_{yx}$  and  $\mathbf{R}_{yy}$  lead to:  $\hat{\mathbf{R}}_{xy} \triangleq \hat{\mathbf{R}}_{xx} \hat{\mathbf{H}}^T = \hat{\mathbf{R}}_{yx}^T$ ,  $\hat{\mathbf{R}}_{yy} \triangleq \hat{\mathbf{H}} \hat{\mathbf{R}}_{xx} \hat{\mathbf{H}}^T + \hat{\mathbf{R}}_{nn}$ .<sup>1</sup> The corresponding first-order perturbations are computed as:

$$\Delta \mathbf{R}_{xy} = \Delta \mathbf{R}_{xx} \mathbf{H}^T + \mathbf{R}_{xx} \Delta \mathbf{H}^T = \Delta \mathbf{R}_{yx}^T \quad (8)$$

$$\Delta \mathbf{R}_{yy} = \Delta \mathbf{H} \mathbf{R}_{xx} \mathbf{H}^T + \mathbf{H} \Delta \mathbf{R}_{xx} \mathbf{H}^T + \mathbf{H} \mathbf{R}_{xx} \Delta \mathbf{H}^T + \Delta \mathbf{R}_{nn}. \quad (9)$$

Efforts for computing  $\mathbf{R}$  and  $\mathbf{R}_\Delta$  yield  $\hat{\mathbf{R}} \triangleq \hat{\mathbf{R}}_{xx} - \hat{\mathbf{R}}_{xy} \hat{\mathbf{R}}_{yy}^{-1} \hat{\mathbf{R}}_{yx}$

$$\hat{\mathbf{R}}_\Delta \triangleq \begin{bmatrix} \mathbf{0}_{\Delta \times (N_b+1)} & \\ & \mathbf{I}_{N_b+1} \\ & & \mathbf{0}_{s \times (N_b+1)} \end{bmatrix} \hat{\mathbf{R}}.$$

The first-order perturbation  $\Delta \mathbf{R}$  is given by [1]

$$\begin{aligned}
 \Delta \mathbf{R} &= \Delta \mathbf{R}_{xx} - \mathbf{R}_{xy} \mathbf{R}_{yy}^{-1} \Delta \mathbf{R}_{yx} - \Delta \mathbf{R}_{xy} \mathbf{R}_{yy}^{-1} \mathbf{R}_{yx} \\
 &\quad + \mathbf{R}_{xy} \mathbf{R}_{yy}^{-1} \Delta \mathbf{R}_{yy} \mathbf{R}_{yy}^{-1} \mathbf{R}_{yx}.
 \end{aligned} \quad (10)$$

<sup>1</sup>Another way to estimate  $\mathbf{R}_{yy}$  is through the output data samples.

$\Delta \mathbf{R}_\Delta$  can be easily derived from the definitions of  $\widehat{\mathbf{R}}_\Delta$  and  $\Delta \mathbf{R}$ . The resulting ‘‘optimal’’ filters are given by

$$\widehat{\mathbf{b}}_o = \frac{\widehat{\mathbf{R}}_\Delta^{-1} \mathbf{e}_0}{\mathbf{e}_0^T \widehat{\mathbf{R}}_\Delta^{-1} \mathbf{e}_0}, \quad \widehat{\mathbf{w}}_o = \widehat{\mathbf{R}}_{yy}^{-1} \widehat{\mathbf{R}}_{yx} \widehat{\mathbf{b}}_o \quad (11)$$

where  $\widehat{\mathbf{b}}_o$  is the appropriately zero-padded version of  $\widetilde{\mathbf{b}}_o$ . Assuming correct past decisions, the corresponding MSE can be expressed as

$$\begin{aligned} \widehat{\mathcal{M}} &\triangleq \mathcal{E} \left[ \left( \widehat{\mathbf{b}}_o^T \mathbf{x} - \widehat{\mathbf{w}}_o^T \mathbf{y} \right) \left( \mathbf{x}^T \widetilde{\mathbf{b}}_o - \mathbf{y}^T \widetilde{\mathbf{w}}_o \right) \right] \\ &= \widetilde{\mathbf{b}}_o^T \mathbf{R}_{xx} \widetilde{\mathbf{b}}_o - \widetilde{\mathbf{b}}_o^T \mathbf{R}_{xy} \widetilde{\mathbf{w}}_o - \widetilde{\mathbf{w}}_o^T \mathbf{R}_{yx} \widetilde{\mathbf{b}}_o + \widetilde{\mathbf{w}}_o^T \mathbf{R}_{yy} \widetilde{\mathbf{w}}_o. \end{aligned} \quad (12)$$

By inspection of (2) and (12), we deduce that  $\widehat{\mathcal{M}}$  is the value of the constrained quadratic function  $\text{MSE}$  at the point  $[\widetilde{\mathbf{b}}_o^T \widetilde{\mathbf{w}}_o^T]^T$ , which is ‘‘close’’ to the optimal point  $[\widetilde{\mathbf{b}}_o^T \mathbf{w}_o^T]^T$ . If we denote with  $\Delta \widetilde{\mathbf{b}}_o$  and  $\Delta \mathbf{w}_o$  the first-order perturbations in quantities  $\widetilde{\mathbf{b}}_o$  and  $\mathbf{w}_o$ , respectively, we obtain the second-order approximation

$$\widehat{\mathcal{M}} = \mathcal{M} + \text{f.o.t.} + \begin{bmatrix} \Delta \widetilde{\mathbf{b}}_o^T & \Delta \mathbf{w}_o^T \end{bmatrix} \begin{bmatrix} \mathbf{R}_{xx} - \mathbf{R}_{xy} \\ -\mathbf{R}_{yx} & \mathbf{R}_{yy} \end{bmatrix} \begin{bmatrix} \Delta \widetilde{\mathbf{b}}_o \\ \Delta \mathbf{w}_o \end{bmatrix}. \quad (13)$$

The summand ‘‘f.o.t’’ (first-order-terms) is zero due to the optimality of the point  $[\widetilde{\mathbf{b}}_o^T \mathbf{w}_o^T]^T$ . Thus, the excess MSE is approximated by the second-order error terms. We must note that an expression analogous to (13) has appeared in [3]. However, expression (13) is not very informative because it does not explicitly reveal the factors which govern the size of the excess MSE. This is our subject in the sequel. At first, we derive first-order approximations to the perturbations on  $\mathbf{b}_o$  and  $\mathbf{w}_o$ . Then, these expressions will be used for the derivation of the second-order approximation to the excess MSE.

**Result 1 [1]:** Let  $\Delta \mathbf{b}_o$  and  $\Delta \mathbf{w}_o$  be the first-order perturbations in quantities  $\mathbf{b}_o$  and  $\mathbf{w}_o$ , respectively. Then

$$\Delta \mathbf{b}_o = \left( \mathbf{b}_o^T \Delta \mathbf{R}_\Delta \mathbf{b}_o \right) \mathbf{R}_\Delta^{-1} \mathbf{e}_0 - \mathbf{R}_\Delta^{-1} \Delta \mathbf{R}_\Delta \mathbf{b}_o \quad (14)$$

$$\Delta \mathbf{w}_o = \mathbf{R}_{yy}^{-1} \left( \mathbf{R}_{yx} \Delta \widetilde{\mathbf{b}}_o + \Delta \mathbf{R}_{yx} \widetilde{\mathbf{b}}_o - \Delta \mathbf{R}_{yy} \mathbf{w}_o \right) \quad (15)$$

where  $\Delta \widetilde{\mathbf{b}}_o$  is the appropriately zero-padded version of  $\Delta \mathbf{b}_o$ .

**Result 2 [1]:** Let SOT denote the second-order error terms in (13). Then

$$\begin{aligned} \text{SOT} &= \left( \widetilde{\mathbf{b}}_o^T \Delta \mathbf{R}_{xy} - \mathbf{w}_o^T \Delta \mathbf{R}_{yy} \right) \mathbf{R}_{yy}^{-1} \\ &\quad \times \left( \Delta \mathbf{R}_{yx} \widetilde{\mathbf{b}}_o - \Delta \mathbf{R}_{yy} \mathbf{w}_o \right) + \Delta \widetilde{\mathbf{b}}_o^T \mathbf{R}_\Delta \Delta \widetilde{\mathbf{b}}_o. \end{aligned} \quad (16)$$

Using (14) and the relation  $\Delta \widetilde{\mathbf{b}}_o^T \mathbf{R}_\Delta \Delta \widetilde{\mathbf{b}}_o = \Delta \mathbf{b}_o^T \mathbf{R}_\Delta \Delta \mathbf{b}_o$ , we derive an alternative expression for the second term of (16), as [1]:

$$\Delta \widetilde{\mathbf{b}}_o^T \mathbf{R}_\Delta \Delta \widetilde{\mathbf{b}}_o = \mathbf{b}_o^T \Delta \mathbf{R}_\Delta \mathbf{R}_\Delta^{-1} \Delta \mathbf{R}_\Delta \mathbf{b}_o - \frac{\left( \mathbf{b}_o^T \Delta \mathbf{R}_\Delta \mathbf{b}_o \right)^2}{\mathcal{M}}. \quad (17)$$

#### 4. MEAN EXCESS MSE: CHANNEL ESTIMATION ERRORS

In this section, we assume that the input and noise SOS are perfectly known and we derive an analytic expression for the mean excess MSE in terms of the channel estimation error covariance matrix  $\mathbf{R}_{\Delta \mathcal{H}_\nu} \triangleq \mathcal{E} \left[ \Delta \mathcal{H}_\nu \Delta \mathcal{H}_\nu^T \right]$ , with  $\Delta \mathcal{H}_\nu \triangleq \widehat{\mathcal{H}}_\nu - \mathcal{H}_\nu$ . We simplify the

results by making the assumption that  $\mathbf{R}_{xx} = \mathbf{I}_{N_f + \nu}$ . Then, from (3), (5), (8) and (10), we obtain

$$\mathbf{R}_{xy} = \mathbf{H}^T = \mathbf{R}_{yx}^T, \quad \mathbf{R} = \mathbf{I}_{N_f + \nu} - \mathbf{H}^T \mathbf{R}_{yy}^{-1} \mathbf{H} \quad (18)$$

$$\Delta \mathbf{R}_{xy} = \Delta \mathbf{H}^T = \Delta \mathbf{R}_{yx}^T, \quad \Delta \mathbf{R}_{yy} = \mathbf{H} \Delta \mathbf{H}^T + \Delta \mathbf{H} \mathbf{H}^T \quad (19)$$

$$\Delta \mathbf{R} = -\mathbf{R}_{xy} \mathbf{R}_{yy}^{-1} \Delta \mathbf{H} \mathbf{R} - \mathbf{R} \Delta \mathbf{H}^T \mathbf{R}_{yy}^{-1} \mathbf{R}_{yx}. \quad (20)$$

In the sequel, we shall use the following relations and definitions:

1. The product  $\mathbf{w}_o^T \mathbf{H}$  expresses the convolution of the multi-channel  $\mathcal{H}_\nu$  and the feed-forward filter  $\mathbf{w}_o$  and, due to the commutativity property of the convolution, is equal to  $\mathcal{H}_\nu^T \mathcal{W}_o$ , where  $\mathcal{W}_o$  is the  $p(\nu + 1) \times (N_f + \nu)$  filtering matrix constructed by the block vector  $\mathbf{w}_o$ . Thus:

$$\mathbf{w}_o^T \mathbf{H} = \mathcal{H}_\nu^T \mathcal{W}_o. \quad (21)$$

2. We define the combined impulse response  $\mathbf{f}_o \triangleq \mathbf{H}^T \mathbf{w}_o$  and the residual impulse response

$$\mathbf{g}_o \triangleq \widetilde{\mathbf{b}}_o - \mathbf{f}_o \quad (22)$$

with elements  $g_{o,i}$ , for  $i = 1, \dots, N_f + \nu$ . It is easy to show that under the above assumptions  $\mathbf{g}_o = \mathbf{R} \widetilde{\mathbf{b}}_o$ .

3. Simple calculations show that

$$\mathbf{g}_o^T \mathbf{H}^T = \mathcal{H}_\nu^T \mathcal{G}_o \quad (23)$$

where  $\mathcal{G}_o$  is the block Hankel matrix defined as

$$\mathcal{G}_o \triangleq \begin{bmatrix} g_{o,1} \mathbf{I}_p & g_{o,2} \mathbf{I}_p & \cdots & g_{o,N_f} \mathbf{I}_p \\ g_{o,2} \mathbf{I}_p & g_{o,3} \mathbf{I}_p & \cdots & g_{o,N_f+1} \mathbf{I}_p \\ \vdots & \vdots & \ddots & \vdots \\ g_{o,\nu+1} \mathbf{I}_p & g_{o,\nu+2} \mathbf{I}_p & \cdots & g_{o,N_f+\nu} \mathbf{I}_p \end{bmatrix}.$$

Now, we consider each term of SOT separately.

A) *Term SOT1:* Let us consider the first term of SOT in (16), denoted SOT1. Using (19), (21), (22) and (23), we obtain [1]

$$\widetilde{\mathbf{b}}_o^T \Delta \mathbf{R}_{xy} - \mathbf{w}_o^T \Delta \mathbf{R}_{yy} = \Delta \mathcal{H}_\nu^T \underbrace{(\mathcal{G}_o - \mathcal{W}_o \mathbf{H}^T)}_{\mathcal{A}_1}.$$

Thus, the term SOT1 can be expressed as

$$\text{SOT1} = \Delta \mathcal{H}_\nu^T \mathcal{A}_1 \mathbf{R}_{yy}^{-1} \mathcal{A}_1^T \Delta \mathcal{H}_\nu. \quad (24)$$

B) *Term SOT2<sub>1</sub>:* The first term of the right-hand side of (17), denoted SOT2<sub>1</sub>, is expressed as SOT2<sub>1</sub> =  $\widetilde{\mathbf{b}}_o^T \Delta \mathbf{R} \mathbf{M} \Delta \mathbf{R} \widetilde{\mathbf{b}}_o$ , with

$$\mathbf{M} \triangleq \begin{bmatrix} \mathbf{0}_{(N_b+1) \times \Delta} \\ \mathbf{I}_{N_b+1} \\ \mathbf{0}_{(N_b+1) \times s} \end{bmatrix} \mathbf{R}_\Delta^{-1} \begin{bmatrix} \mathbf{0}_{(N_b+1) \times \Delta} & \mathbf{I}_{N_b+1} & \mathbf{0}_{(N_b+1) \times s} \end{bmatrix}.$$

Using (6), (20), (21) and (23) we obtain [1]

$$\begin{aligned} \text{SOT2}_1 &= \Delta \mathcal{H}_\nu^T \underbrace{(\mathcal{W}_o \mathbf{R} + \mathcal{G}_o \mathbf{R}_{yy}^{-1} \mathbf{R}_{yx})}_{\mathcal{A}_2} \mathbf{M} \\ &\quad \times (\mathbf{R} \mathcal{W}_o^T + \mathbf{R}_{xy} \mathbf{R}_{yy}^{-1} \mathcal{G}_o^T) \Delta \mathcal{H}_\nu. \end{aligned} \quad (25)$$

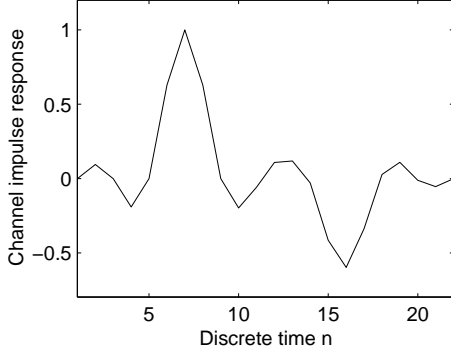


Fig. 3. Channel impulse response.

C) Term SOT<sub>2</sub>: The second term of the right-hand side of (17), denoted SOT<sub>2</sub>, is given by [1]

$$\text{SOT}_2 = \frac{4}{\mathcal{M}} \Delta \mathcal{H}_\nu^T \underbrace{\mathcal{W}_o \mathbf{g}_o \mathbf{g}_o^T \mathcal{W}_o^T}_{\mathcal{A}_3} \Delta \mathcal{H}_\nu. \quad (26)$$

Using (24), (25) and (26), we obtain the second-order approximation

$$\mathcal{E}[\text{SOT}] = \mathcal{E}[\text{Tr}(\Delta \mathcal{H}_\nu^T \mathcal{S} \Delta \mathcal{H}_\nu)] = \text{Tr}(\mathcal{S} \mathbf{R}_{\Delta \mathcal{H}_\nu}) \quad (27)$$

where Tr denotes the matrix trace and

$$\mathcal{S} \triangleq \mathcal{A}_1 \mathbf{R}_{yy}^{-1} \mathcal{A}_1^T + \mathcal{A}_2 \mathbf{M} \mathcal{A}_2^T - \frac{4 \mathcal{A}_3}{\mathcal{M}}. \quad (28)$$

Using some simplifications and approximations it is shown in [1] that

$$\mathcal{E}[\text{SOT}] \leq 2(\nu + 1) \|\mathbf{w}_o\|_2^2 \|\mathbf{R}_{\Delta \mathcal{H}_\nu}\|_F \quad (29)$$

where  $\|\cdot\|_F$  denotes the matrix Frobenious norm. For the analogous results relating  $\mathcal{E}[\text{SOT}]$  and the noise and input SOS estimation error covariance matrices the reader is referred to [1].

## 5. SIMULATIONS

In our simulations, we use the communication channel whose impulse response is depicted in Fig. 3. It models a multipath scenario resulting in severe ISI, and is derived by oversampling, by a factor of 2, the impulse response  $h(t) = p(t) - 0.6p(t - 4.45T)$ , where  $p(t)$  is the (truncated) pulse with a raised-cosine spectrum and roll-off factor  $\beta = 0.22$ . The truncation interval is  $[-3T, 3T]$ , where  $T$  denotes the symbol period, and the sampling instants are the integer multiples of  $T/2$ . The vector impulse response  $\{\mathbf{h}_i\}_{i=0}^{10}$  of the corresponding 1-input/2-output system is constructed by grouping together the even and the odd terms of this oversampled impulse response. The input is a BPSK signal, taking, with equal probability, the values  $\pm 1$ , yielding  $\mathbf{R}_{xx} = \mathbf{I}_{N_f + \nu}$ . At the multi-channel output, we add temporally and spatially white Gaussian noise with variance  $\sigma_n^2$ . Hence,  $\mathbf{R}_{nn} = \sigma_n^2 \mathbf{I}_{2N_f}$ . We define the SNR as  $\text{SNR} \triangleq 10 \log_{10} (\mathcal{E}[\|\mathbf{w}_n\|_2^2] / \mathcal{E}[\|\mathbf{n}_n\|_2^2])$ , where  $\mathbf{w}_n$  is the noiseless channel output, defined as  $\mathbf{w}_n \triangleq \sum_{i=0}^{\nu} \mathbf{h}_i x_{n-i}$ .

In the sequel, we consider the performance of the MMSE-DFE with filter lengths  $N_f = 8$  and  $N_b = 6$ , for the delay  $\Delta = 5$ . Due to space limitation, we consider only the channel estimation errors.

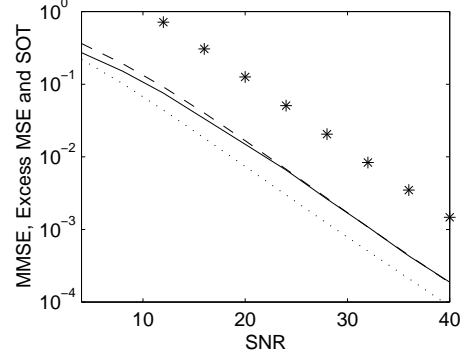


Fig. 4. MMSE  $\mathcal{M}$  (dotted line), mean excess MSE  $\mathcal{E}[\widehat{\mathcal{M}} - \mathcal{M}]$  (solid line),  $\mathcal{E}[\text{SOT}]$  (dashed line) and bound(29) (\*-), versus SNR.

Thus, we assume that the input and additive white Gaussian channel noise SOS are perfectly known. Also, we assume that the channel order is perfectly known. The channel is estimated by the application of the maximum-likelihood method to the training sequence consisted of  $N_{tr} = 26$  consecutive training symbols (a packet consists of  $N_p = 138$  data symbols) [4, Sect. 15.2]. The channel estimate is used for the computation of the DFE filters. In Fig. 4, we plot the MMSE  $\mathcal{M}$  (dotted line), the mean of the theoretical excess MSE  $\mathcal{E}[\widehat{\mathcal{M}} - \mathcal{M}]$  (solid line),  $\mathcal{E}[\text{SOT}]$  (dashed line) and the bound (29). Quantity  $\mathcal{E}[\widehat{\mathcal{M}} - \mathcal{M}]$  is computed experimentally over 500 independent input and additive noise realizations as follows:

$$\mathcal{E}[\widehat{\mathcal{M}} - \mathcal{M}] = \frac{1}{500} \sum_{i=1}^{500} (\widehat{\mathcal{M}}_i - \mathcal{M})$$

where  $\widehat{\mathcal{M}}_i - \mathcal{M}$  is the excess MSE of the  $i$ -th realization. The channel estimation error covariance matrix  $\mathbf{R}_{\Delta \mathcal{H}_\nu}$ , used for the computation of  $\mathcal{E}[\text{SOT}]$ , can be computed by extending results of [4, Sect. 15.2] to the 1-input/2-output channel setting. We observe that:

1.  $\mathcal{E}[\text{SOT}]$  provides an accurate measure of the mean excess MSE. This has been observed in many simulations validating the usefulness of our second-order approximation. We also observe that the bound (29) is pessimistic.
2. The excess MSE is larger than the MMSE. The same has been observed in simulations with severe ISI channels, while for less severe channels it is usually smaller than the MMSE.

## 6. REFERENCES

- [1] A. P. Liavas, "On the robustness of the finite-length MMSE-DFE with respect to channel and second-order estimation errors," submitted to the *IEEE Trans. Signal Processing*.
- [2] N. Al-Dhahir and J. M. Cioffi, "MMSE Decision-Feedback equalizers: Finite-length results," *IEEE Trans. Information Theory*, vol. 41, no. 4, pp. 961–975, July 1995.
- [3] N. Al-Dhahir and J. M. Cioffi, "Mismatched finite-complexity MMSE decision feedback equalizers," *IEEE Trans. Signal Processing*, vol. 45, no. 4, pp. 935–944, April 1997.
- [4] H. Meyr, M. Moeneclaey and S. Fechtel. *Digital Communication Receivers: Synchronization, Channel Estimation and Signal Processing*. John Wiley, 1998.