How to choose the number of taps in a DFE ?

Naofal Al-Dhahir *and Christina Fragouli †
AT&T shannon Laboratory
Florham Park, NJ 07932

Abstract

In this paper, we address the following problem. Given a fixed *total* number of feedforward and feedback filter taps in a decision feedback equalizer, what is the optimum number of taps for each filter and what is the optimum decision delay setting ? We propose a simple algorithm that reduces the solution search space from being 2–dimensional to 1–dimensional at a small performance loss. We apply our algorithm to indoor wireless channels.

1 Introduction

In an increasing number of applications, a minimum–mean–square–error decision–feedback equalizer (MMSE–DFE) is implemented in the receiver to mitigate intersymbol interference (ISI) on bandlimited dispersive channels. In most practical situations, computational and implementational complexity considerations often place a constraint on the maximum number of total (feedforward and feedback) filter taps that can be used. In addition, for short filters, the decision delay setting could affect performance significantly. Therefore, it becomes of chief importance to understand the dependence of the performance of the finite–length MMSE–DFE on the number of its feedforward and feedback filters and on the decision delay in order to set them properly.

Currently–used techniques for setting these 3 parameters are either ad–hoc or computationally intense (require an exhaustive search over these three parameters and inversion of a matrix whose size is equal to the total number of feedforward and feedback taps for each step in the search). In

*Contact author. Phone :(973)360 – 8407, FAX :(973)360 – 8178, Email : naofal@research.att.com
†C. Fragouli is now with the National Capodistrian University of Athens.

this paper, we present a simple characterization of the MMSE–DFE's performance that results in a simple order–recursive algorithm. Furthermore, we reduce the exhaustive search space from a 2–dimensional to a 1–dimensional search for joint selection of the number of MMSE–DFE taps and the decision delay setting.

The rest of this paper is organized as follows. Section 2 presents closed–form expressions for computing the optimum DFE taps and evaluating its performance given the number of feedforward and feedback taps and a decision delay setting. Section 3 proposes a simple 1–dimensional suboptimal search algorithm. Numerical results illustrating the near–optimum performance of the proposed algorithm are given in Section 4 and the paper is concluded in Section 5.

2 The Finite-Length MMSE-DFE

The channel impulse response (CIR) is assumed to be a finite impulse response (FIR) filter with memory ν, i.e., $\mathbf{h} = \begin{bmatrix} h_0 & h_1 & \cdots & h_\nu \end{bmatrix}^t$, where $(.)^t$ denotes the transpose. Over a block of $(N+1)$ channel output symbols, denoted by $\mathbf{y}_{i-N}^{i}$, the input–output relationship can be cast in matrix form as follows (see also Equation (12) in the Appendix)

$$\mathbf{y}_{i-N}^{i} = \mathbf{H}\mathbf{x}_{i-N-\nu}^{i} + \mathbf{n}_{i-N}^{i} \ , \tag{1}$$

where $\mathbf{H}$ is the $(N+1) \times (N+\nu+1)$ Toeplitz [1] channel convolution matrix. The input and noise sequences (denoted by $\mathbf{x}_{i-N-\nu}^{i}$ and $\mathbf{n}_{i-N}^{i}$, respectively) are assumed to be zero–mean, independent of each other, and have positive–definite (non–singular) auto–correlation matrices denoted by $\mathbf{R}_{xx}$ and $\mathbf{R}_{nn}$, respectively. Using (1), the input–output cross–correlation matrix and the output auto–correlation matrix are given by

$$\mathbf{R}_{xy} \stackrel{def}{=} E[\mathbf{x}_{i-N-\nu}^{i}\mathbf{y}_{i-N}^{*i}] = \mathbf{R}_{xx}\mathbf{H}^* \tag{2}$$

$$\mathbf{R}_{yy} \stackrel{def}{=} E[\mathbf{y}_{i-N}^{i}\mathbf{y}_{i-N}^{*i}] = \mathbf{H}\mathbf{R}_{xx}\mathbf{H}^* + \mathbf{R}_{nn} \ , \tag{3}$$

where E[.] and $(.)^*$ denote the expected value and the complex–conjugate transpose, respectively.

2.1 Optimum 2-Dimensional Search Algorithm

A block diagram of the MMSE–DFE is shown in Figure 1. The error symbol at time i is given by $e_i \stackrel{def}{=} \hat{x}_{i-\Delta} - z_i = \tilde{\mathbf{b}}^*\mathbf{x}_{i-N-\nu}^{i} - \mathbf{w}^*\mathbf{y}_{i-N}^{i}$, where Δ is the decision delay, $\tilde{\mathbf{b}}^* \stackrel{def}{=} \begin{bmatrix} \mathbf{0}_{1\times\Delta} & \mathbf{b}^* & \mathbf{0}_{1\times s} \end{bmatrix}$,

[1] The channel is assumed *time-invariant* over the transmission block.

$\mathbf{0}_{m\times n}$ is the all–zeros matrix with m rows and n columns, and $s \stackrel{def}{=} N+\nu-\Delta-N_b$. We assume that the feedforward filter consists of $(N+1)$ taps, i.e., $\mathbf{w} \stackrel{def}{=} \left[\begin{array}{ccc} w_0 & \cdots & w_N \end{array}\right]^t$, and the *strictly–causal* feedback filter $\mathbf{e}_\Delta^* - \tilde{\mathbf{b}}^* = \left[\begin{array}{ccccc} \mathbf{0}_{1\times(\Delta+1)} & -b_1 & \cdots & -b_{N_b} & \mathbf{0}_{1\times s} \end{array}\right]$ consists of N_b non–zero taps, where $\mathbf{b} \stackrel{def}{=} \left[\begin{array}{cccc} 1 & b_1 & \cdots & b_{N_b} \end{array}\right]^t$ and $\mathbf{e}_\Delta^* = \left[\begin{array}{ccc} \mathbf{0}_{1\times\Delta} & 1 & \mathbf{0}_{1\times(N+\nu-\Delta)} \end{array}\right]$.

Using the *Orthogonality Principle* of linear least squares estimation, it can be shown that the mean square of the (assumed stationary) error sequence is equal to

$$\begin{aligned} MSE &\stackrel{def}{=} \mathrm{E}[\mid e_i \mid^2] \\ &= \tilde{\mathbf{b}}^*(\mathbf{R}_{xx} - \mathbf{R}_{xy}\mathbf{R}_{yy}^{-1}\mathbf{R}_{yx})\tilde{\mathbf{b}} \\ &\stackrel{def}{=} \tilde{\mathbf{b}}^*\mathbf{R}_{x/y}^{\perp}\tilde{\mathbf{b}} \qquad (4) \\ &= \mathbf{b}^*\mathbf{R}_\Delta\mathbf{b}\ . \qquad (5) \end{aligned}$$

We defined

$$\mathbf{R}_\Delta \stackrel{def}{=} \left[\begin{array}{ccc} \mathbf{0}_{(N_b+1)\times\Delta} & \mathbf{I}_{N_b+1} & \mathbf{0}_{(N_b+1)\times s} \end{array}\right] \mathbf{R}_{x/y}^{\perp} \left[\begin{array}{c} \mathbf{0}_{\Delta\times(N_b+1)} \\ \mathbf{I}_{N_b+1} \\ \mathbf{0}_{s\times(N_b+1)} \end{array}\right], \qquad (6)$$

where $\mathbf{I}_m$ is the identity matrix of size m. It can be shown that the optimum feedback and feedforward filters that minimize (5) are given by

$$\begin{aligned} \mathbf{b}_{opt} &= \frac{\mathbf{R}_\Delta^{-1}\left[\begin{array}{c} 1 \\ \mathbf{0}_{N_b\times 1} \end{array}\right]}{\mathbf{R}_\Delta^{-1}(1,1)} \qquad (7) \\ \mathbf{w}_{opt}^* &= \tilde{\mathbf{b}}_{opt}^*\mathbf{R}_{xy}\mathbf{R}_{yy}^{-1}\ , \qquad (8) \end{aligned}$$

where $\mathbf{R}_\Delta^{-1}(1,1)$ is the $(1,1)$ entry of $\mathbf{R}_\Delta^{-1}$. The resulting minimum mean square error (MMSE) is given by

$$MMSE(N, N_b, \Delta) = \frac{1}{\mathbf{R}_\Delta^{-1}(1,1)}\ . \qquad (9)$$

Therefore, the *unbiased* [3] decision–point SNR of the MMSE–DFE is equal to

$$SNR_{MMSE-DFE,U}(N, N_b, \Delta) = \frac{1}{MMSE(N, N_b, \Delta)} - 1\ . \qquad (10)$$

Hence, for any choice of the triplet (N, N_b, Δ), Equations (7) and (8) are used to determine the optimum MMSE–DFE settings and Equation (10) is used to evaluate its performance. Under the implementation constraint $N+N_b = N_{tot}$ (where N_{tot} is a given constant), determining the optimum triplet (N, N_b, Δ) is computationally intense since it requires an exhaustive 2–dimensional [2] search

[2]The linear constraint $N+N_b = N_{tot}$ reduces the exhaustive search space for (N, N_b, Δ) from being 3–dimensional to 2–dimensional.

and computation of the inverse $\mathbf{R}_\Delta^{-1}$ for each step of the search. Next, we show how to reduce this computational complexity at a small performance loss.

2.2 Recursive Formula for MMSE

Using Kramer's rule for matrix inversion [4], we get the following alternative MMSE expression

$$MMSE(N, N_b, \Delta) = \frac{|\mathbf{R}_\Delta|}{|\bar{\mathbf{R}}_\Delta^{1,1}|} \,, \tag{11}$$

where $\bar{\mathbf{R}}_\Delta^{1,1}$ is obtained from $\mathbf{R}_\Delta$ by deleting its first row and first column and $|.|$ denotes the determinant. This compact characterization of the MMSE offers further insight into the effect of the triplet (N, N_b, Δ) on the MMSE–DFE performance as illustrated in Figure 2 and described in the sequel :

- Given $\mathbf{h}$, $\mathbf{R}_{xx}$, and $\mathbf{R}_{nn}$, the number of feedforward taps $(N+1)$ determines $\mathbf{R}_{x/y}^{\perp}$ as defined in (4).
- The number of feedback taps N_b determines the *size* of $\mathbf{R}_\Delta$ as defined in (6).
- The delay of the feedforward filter Δ determines the *location* of the $(1,1)$ entry of $\mathbf{R}_\Delta$. Hence, changing the delay corresponds to *shifting* $\mathbf{R}_\Delta$ along the main diagonal of $\mathbf{R}_{x/y}^{\perp}$.
- For a given N, optimizing N_b and Δ is equivalent to locating the (N_b+1)–dimensional matrix $\mathbf{R}_\Delta$ whose upper–left corner is on the main diagonal of $\mathbf{R}_{x/y}^{\perp}$ such that the determinant ratio [3] in (11) is minimized. Note that (11) can be computed using the well–known *order–recursive* formula for the determinant $|\mathbf{R}_\Delta| = \sum_{j=1}^{N_b+1}(-1)^{j+1}\mathbf{R}_\Delta(1,j)|\bar{\mathbf{R}}_\Delta^{1,j}|$.

Finally, applying Kramer's rule to (7), we get the following alternative expression for the optimum feedback filter coefficients

$$\mathbf{b}_{opt} = \left[\begin{array}{ccccc} 1 & \frac{|\bar{\mathbf{R}}_\Delta^{1,2}|}{|\bar{\mathbf{R}}_\Delta^{1,1}|} & \frac{|\bar{\mathbf{R}}_\Delta^{1,3}|}{|\bar{\mathbf{R}}_\Delta^{1,1}|} & \cdots & \frac{|\bar{\mathbf{R}}_\Delta^{1,(N_b+1)}|}{|\bar{\mathbf{R}}_\Delta^{1,1}|} \end{array} \right] .$$

[3]This ratio is also equal to the first diagonal element in an upper–diagonal–lower Cholesky factorization of $\mathbf{R}_\Delta$.

3 Suboptimum 1–Dimensional Search Algorithm

Problem Statement

Given the CIR $\mathbf{h}$ of memory ν and a fixed total number of DFE taps $N_{tot} \geq \nu$ [4]. Assume that the noise is white ($\mathbf{R}_{nn} = \frac{1}{SNR}\mathbf{I}_{N+1}$ where SNR is the signal to noise ratio) and the input data is white ($\mathbf{R}_{xx} = \mathbf{I}_{N+\nu+1}$). We want to solve the non–linear integer programming problem

$$min_{N,N_b,\Delta} MMSE(N, N_b, \Delta) \quad \textit{subject to} \quad N + N_b = N_{tot}$$

where the *integers* $N, N_b,$ and Δ satisfy $\quad 0 \leq N \leq N_{tot} - 1 \; ; \; 1 \leq N_b \leq N_{tot} \; ; \; 0 \leq \Delta \leq N + \nu$.

An exhaustive search would require $\sum_{N=0}^{N_{tot}-1}(N + \nu + 1) = \frac{(N_{tot}+1)N_{tot}}{2} + N_{tot}\nu$ steps since there are $(N + \nu + 1)$ possible delay values. We can further restrict the search space by noting that

- $1 \leq N_b \leq \nu$ (we prove in the Appendix that for $N_b > \nu$ the last $(N_b - \nu)$ feedback taps are equal to zero).

- $N \leq \Delta \leq N + \nu - N_b$ (since for $N > \Delta$ the feedforward filter coefficients $\mathbf{w}_{\Delta+1}^{N}$ are equal to zero. Moreover, we do not need $\Delta > (N + \nu - N_b)$ because of the feedback filter post–cursor cancellation action).

Over these reduced ranges, the exhaustive search size is reduced [5] to $\sum_{N_b=1}^{\nu}(\nu - N_b + 1) = \frac{\nu^2+\nu}{2}$ steps but still has quadratic complexity. The following suboptimum algorithm reduces the search to a 1–dimensional search with ν steps.

Proposed Algorithm

Input : Total number of DFE tap N_{tot} and channel memory ν

1. For $N_b = 1, 2, \cdots, \nu$.

2. Set $N = N_{tot} - N_b$.

3. Set $\Delta = N + \nu - N_b = N_{tot} - 2N_b + \nu$. In other words, *we assume that the feedback filter cancels all post-cursor ISI.*[6] Figure 3 illustrates this assumed relationship between the DFE parameters.

[4]When $N_{tot} < \nu$, the DFE performance is degraded significantly. Hence, this case is not useful in practice and is not considered.

[5]Assuming that $N_{tot} \geq \nu$.

[6]The optimum solution is not always on this line but we found the performance loss from this assumption to be small, as it will be shown in the simulations section.

4. Choose the value of N_b that results in the minimum MMSE, as given in (11),i.e.

$$N_b = argmin_{1\le N_b \le \nu} MMSE(N = N_{tot} - N_b, N_b, \Delta = N_{tot} - 2N_b + \nu) \ .$$

   The corresponding values of N and Δ are determined from Steps 2 and 3 above. Figure 4 depicts an example of the variation of $SNR_{MMSE-DFE,U}$ versus N_b for a wireless indoor channel at input SNR of 15 dB with $N_{tot} = 10$. The existence of several local optima justifies the need for the 1–dimensional search in general. However, for large N_{tot} (compared to ν) no search is needed as detailed next.

Special Case

When $N_{tot} \gg \nu$, the optimum values for the triplet (N, N_b, Δ) are $N_b = \nu$, $N = N_{tot} - \nu$, and $\Delta = N$.[7] For this special case, the error sequence of the FIR MMSE–DFE is *white* when its filter taps are optimized (see the Appendix for a proof). [8].

4 Numerical Results

We consider the wireless local area network (LAN) environment and compute the loss in $SNR_{MMSE-DFE,U}$ (as given in (10)) when using the suboptimal 1–dimensional search algorithm instead of optimal 2–dimensional exhaustive search. The results are shown in Figures 5 and 6 for Channels A and B as specified in the ETSI HIPERLAN standard [2]. which model an office environment with maximum delay spread of 390 nsec and 730 nsec, respectively. This corresponds to a channel memory of 8 and 15 for Channels A and B, respectively. The input SNR is set to 15 dB and the results are averaged over 1000 channel realizations. It can be seen that for $N_{tot} \ge \nu$, the resulting performance loss is less than 1.0 and 1.4 dB for Channels A and B, respectively. This loss diminishes quickly to less than 0.25 dB for both channels for $N_{tot} \ge 1.5\nu$.

[7] The assumption $N_b = \nu$ was termed the *key assumption* in [1] where it was shown to reduce the MMSE–DFE filter computation complexity from cubic to quadratic by using *Cholesky factorization* (instead of inversion) and exploiting matrix structure.

[8] Note that this whiteness property was shown for the infinite–length (i.e. $N = \infty$) MMSE–DFE in [3] but it does not hold for arbitrary finite (N, N_b, Δ) values.

5 Conclusions

Assuming a fixed total number of taps in the MMSE-DFE, dividing it between feedforward and feedback taps and determining the decision delay that optimize performance require a 2-dimensional exhaustive search. We presented a suboptimal 1-dimensional search algorithm. Simulation results in a wireless LAN environment showed that the proposed reduced-complexity algorithm results in a small performance loss from the optimal solution.

A Special Case of $N_b = \nu$

A vector of $N+1$ channel outputs can be expressed as

$$\underbrace{\begin{bmatrix} y(i) \\ y(i-1) \\ \vdots \\ y(i-N) \end{bmatrix}}_{\mathbf{y}_{i-N}^{i}} = \underbrace{\begin{bmatrix} h_0 & h_1 & \dots & h_\nu & 0 & \dots & 0 \\ 0 & h_0 & h_1 & \dots & h_\nu & \dots & 0 \\ & & \ddots & \ddots & & \ddots & \vdots \\ 0 & \dots & 0 & h_0 & h_1 & \dots & h_\nu \end{bmatrix}}_{\mathbf{H}_{(N+1)\times(N+\nu+1)}} \underbrace{\begin{bmatrix} x(i) \\ x(i-1) \\ \vdots \\ x(i-N-\nu) \end{bmatrix}}_{\mathbf{x}_{i-N-\nu}^{i}} . \tag{12}$$

We make the following assumptions

1. The feedback filter length is equal to the channel memory ($N_b = \nu$). The result also holds for $N_b > \nu$, but the additional memory is redundant.

2. The input data sequence is white ($\mathbf{R}_{xx} = \mathbf{I}_{N+\nu+1}$).

3. The input noise sequence is white ($\mathbf{R}_{nn} = \frac{1}{SNR}\mathbf{I}_N$).

4. Noise and data are uncorrelated.

Under these assumptions, we will show that

1. The error sequence is white

2. If the decision delay is Δ, feedforward filter taps with indices greater than $\Delta + 1$ are equal to zero, i.e., the optimum delay is $\Delta = N$.

We start with the error symbol e_i which can be expressed as

$$
\begin{aligned}
e_i &= x(i-\Delta) - (\underbrace{\left[\begin{array}{cccc} w_0^* & w_1^* & \cdots & w_N^* \end{array}\right]}_{\mathbf{w}_0^{*N}} \underbrace{\left[\begin{array}{c} y(i) \\ y(i-1) \\ \vdots \\ y(i-N) \end{array}\right]}_{\mathbf{y}_{i-N}^{i}} - \underbrace{\left[\begin{array}{ccc} b_1^* & \cdots & b_{N_b}^* \end{array}\right]}_{\mathbf{b}_1^{*N_b}} \underbrace{\left[\begin{array}{c} x(i-\Delta-1) \\ x(i-\Delta-2) \\ \vdots \\ x(i-\Delta-N_b) \end{array}\right]}_{\mathbf{x}_{i-\Delta-N_b}^{i-\Delta-1}}) \\
&= x(i-\Delta) - \underbrace{\left[\begin{array}{ccc} w_0^* & \cdots & w_\Delta^* \end{array}\right]}_{\mathbf{w}_0^{*\Delta}} \underbrace{\left[\begin{array}{c} y(i) \\ \vdots \\ y(i-\Delta) \end{array}\right]}_{\mathbf{y}_{i-\Delta}^{i}} - \underbrace{\left[\begin{array}{ccc} w_{\Delta+1}^* & \cdots & w_N^* \end{array}\right]}_{\mathbf{w}_{\Delta+1}^{*N}} \underbrace{\left[\begin{array}{c} y(i-\Delta-1) \\ \vdots \\ y(i-N) \end{array}\right]}_{\mathbf{y}_{i-N}^{i-\Delta-1}} \\
&+ \underbrace{\left[\begin{array}{ccc} b_1^* & \ldots & b_{N_b}^* \end{array}\right]}_{\mathbf{b}_1^{*N_b}} \underbrace{\left[\begin{array}{c} x(i-\Delta-1) \\ x(i-\Delta-2) \\ \vdots \\ x(i-\Delta-N_b) \end{array}\right]}_{\mathbf{x}_{i-\Delta-N_b}^{i-\Delta-1}} \\
&= x(i-\Delta) - \mathbf{w}_0^{*\Delta}(\mathbf{H}\mathbf{x}_{i-\Delta-\nu}^{i} + \mathbf{n}_{i-\Delta}^{i}) - \mathbf{w}_{\Delta+1}^{*N}(\mathbf{H}\mathbf{x}_{i-N-\nu}^{i-\Delta-1} + \mathbf{n}_{i-N}^{i-\Delta-1}) + \mathbf{b}_1^{*N_b}\mathbf{x}_{i-\Delta-N_b}^{i-\Delta-1} \\
&= x(i-\Delta) - \mathbf{w}_0^{*\Delta}(\left[\begin{array}{cc} \mathbf{H}_1 & \mathbf{H}_2 \end{array}\right] \left[\begin{array}{c} \mathbf{x}_{i-\Delta}^{i} \\ \mathbf{x}_{i-\Delta-\nu}^{i-\Delta-1} \end{array}\right] + \mathbf{n}_{i-\Delta}^{i}) - \mathbf{w}_{\Delta+1}^{*N}(\left[\begin{array}{cc} \mathbf{H}_3 & \mathbf{H}_4 \end{array}\right] \left[\begin{array}{c} \mathbf{x}_{i-\Delta-\nu}^{i-\Delta-1} \\ \mathbf{x}_{i-N-\nu}^{i-\Delta-\nu-1} \end{array}\right] + \mathbf{n}_{i-N}^{i-\Delta-1}) \\
&+ \mathbf{b}_1^{*N_b}\mathbf{x}_{i-\Delta-N_b}^{i-\Delta-1} \ , \qquad (13)
\end{aligned}
$$

where $\mathbf{H}_1$, $\mathbf{H}_2$, $\mathbf{H}_3$, and $\mathbf{H}_4$ are submatrices of $\mathbf{H}$. Using the first assumption ($N_b = \nu$), the last equation above becomes

$$
\begin{aligned}
e_i &= \underbrace{x(i-\Delta) - \mathbf{w}_0^{*\Delta}(\mathbf{H}_1\mathbf{x}_{i-\Delta}^{i} + \mathbf{n}_{i-\Delta}^{i})}_{\alpha} \\
&+ \underbrace{(-\mathbf{w}_0^{*\Delta}\mathbf{H}_2 - \mathbf{w}_{\Delta+1}^{*N}\mathbf{H}_3 + \mathbf{b}_1^{*N_b})}_{\beta^*}\mathbf{x}_{i-\Delta-\nu}^{i-\Delta-1} - \mathbf{w}_{\Delta+1}^{*N}(\mathbf{H}_4\mathbf{x}_{i-N-\nu}^{i-\Delta-\nu-1} + \mathbf{n}_{i-N}^{i-\Delta-1}) \ . \qquad (14)
\end{aligned}
$$

We want to minimize the MMSE cost $E[|e_i|^2]$ over all $\mathbf{w}_0^{\Delta}$, $\mathbf{w}_{\Delta+1}^{N}$ and $\mathbf{b}_1^{*N_b}$. Using the whiteness assumptions on input and noise, the MMSE cost can be expressed as

$$
E[|e_i|^2] = \alpha^2 + \|\beta\|^2 + \mathbf{w}_{\Delta+1}^{*N}(\mathbf{H}_4\mathbf{H}_4^* + \frac{1}{SNR}\mathbf{I}_{N-\Delta})\mathbf{w}_{\Delta+1}^{N} \ , \qquad (15)
$$

where α is a scalar constant and $\|.\|$ denotes the norm of a vector. It is clear that MMSE attains its minimum value α^2 when

$$\mathbf{w}_{\Delta+1}^{N} = \mathbf{0} \text{ and } \beta = \mathbf{0} \Rightarrow \mathbf{b}_1^{*N_b} = \mathbf{w}_0^{*\Delta}\mathbf{H}_2 \ . \tag{16}$$

Therefore, the feedforward filter has at most $\Delta + 1$ nonzero taps, i.e. $N = \Delta$. Note that the same conclusion holds had we assumed $N_b > \nu$ since in that case only the term $\|\mathbf{b}_{\nu+1}^{*N_b}\|^2$ is added to (15). Hence, to minimize MMSE, we would set $\mathbf{b}_{\nu+1}^{N_b} = \mathbf{0}$, i.e., we do not need more than ν feedback taps.

From (14) and (16) the error symbol e_i can be expressed as

$$e_i = x(i-\Delta) - \mathbf{w}_0^{*\Delta}(\mathbf{H}_1\mathbf{x}_{i-\Delta}^{i} + \mathbf{n}_{i-\Delta}^{i}) \ . \tag{17}$$

The optimum $\mathbf{w}_0^{\Delta}$ that minimizes (15) can be calculated from (17) as

$$\mathbf{w}_0^{*\Delta} = \mathbf{e}_{\Delta}^{*}\mathbf{H}_1^{*}(\mathbf{H}_1\mathbf{H}_1^{*} + \frac{1}{SNR}\mathbf{I}_{\Delta+1})^{-1} \ , \tag{18}$$

where $\mathbf{e}_{\Delta}^{*}$ is the $1\times(\Delta+1)$ row vector $\begin{bmatrix} 0 & \cdots & 0 & 1 \end{bmatrix}$. We will now show that $E[e_j e_i^*] = 0$. Assume that $j > i$. From (17), we have

$$E[e_j e_i^*] = E[e_j(x(i-\Delta) - \mathbf{w}_0^{*\Delta}(\mathbf{H}_1\mathbf{x}_{i-\Delta}^{i} + \mathbf{n}_{i-\Delta}^{i}))^*] \ . \tag{19}$$

But since $j > i$, we get

$$E[e_j x^*(i-\Delta)] = 0 \ . \tag{20}$$

Assume that, $i - \Delta < j - \Delta \le i$ (or else the second term in (19) also becomes zero), then

$$\begin{aligned} E[e_j e_i^*] &= -\mathbf{w}_0^{*\Delta}E[e_j(\mathbf{x}_{i-\Delta}^{*i}\mathbf{H}_1^* + \mathbf{n}_{i-\Delta}^{*i})] \\ &= -\mathbf{w}_0^{*\Delta}E[x(j-\Delta)\mathbf{x}_{i-\Delta}^{*i}\mathbf{H}_1^* - \mathbf{w}_0^{*\Delta}(\mathbf{H}_1\mathbf{x}_{j-\Delta}^{j}\mathbf{x}_{i-\Delta}^{*i}\mathbf{H}_1^* + \mathbf{n}_{j-\Delta}^{j}\mathbf{n}_{i-\Delta}^{*i})] \ . \end{aligned} \tag{21}$$

The following relations hold

$$E[\mathbf{x}_{j-\Delta}^{j}\mathbf{x}_{i-\Delta}^{*i}] = \mathbf{J}, \quad E[\mathbf{n}_{j-\Delta}^{j}\mathbf{n}_{i-\Delta}^{*i}] = \frac{1}{SNR}\mathbf{J}, \quad E[x(j-\Delta)\mathbf{x}_{i-\Delta}^{*i}] = \mathbf{e}_{\Delta}^{*}\mathbf{J} \ , \tag{22}$$

where the matrix $\mathbf{J}$ has the form $\mathbf{J} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{I}_{\Delta+1-(j-i)} & \mathbf{0} \end{bmatrix}$. Note that under the assumption $j - \Delta \le i$, we have $\Delta + 1 - (j-i) > 0$ (or else $\mathbf{J}$ becomes a zero matrix). This is because

$$E[\mathbf{x}_{j-\Delta}^{j}\mathbf{x}_{i-\Delta}^{*i}] = E\begin{bmatrix} x(j) \\ \vdots \\ x(j-\Delta) \end{bmatrix}\begin{bmatrix} x^*(i) & \cdots & x^*(j-\Delta) & \cdots & x^*(i-\Delta) \end{bmatrix}$$

$$= \; E\begin{bmatrix} x(j) \\ \vdots \\ x(j-\Delta) \end{bmatrix}\begin{bmatrix} x^*(i) & \cdots & x^*(j-\Delta) & 0 & \cdots & 0 \end{bmatrix} = E[\mathbf{x}_{j-\Delta}^{j}\mathbf{x}_{j-\Delta}^{*j}]\mathbf{J} \; .$$

Therefore, (21) becomes

$$E[e_j e_i^*] = -\mathbf{w}_0^{*\Delta}(\mathbf{e}_\Delta^*\mathbf{J}\mathbf{H}_1^* - \mathbf{w}_0^{*\Delta}(\mathbf{H}_1\mathbf{J}\mathbf{H}_1^* + \frac{1}{SNR}\mathbf{J})) \; . \tag{23}$$

But $\mathbf{J}\mathbf{H}_1^* = \mathbf{H}_1^*\mathbf{J}$ since $\mathbf{H}_1^*$ is a square lower–triangular Toeplitz matrix. Therefore,

$$E[e_j e_i^*] = -\mathbf{w}_0^{*\Delta}(\mathbf{e}_\Delta^*\mathbf{H}_1^* - \mathbf{w}_0^{*\Delta}(\mathbf{H}_1\mathbf{H}_1^* + \frac{1}{SNR}\mathbf{I}_{\Delta+1}))\mathbf{J} \; . \tag{24}$$

Substituting for $\mathbf{w}_0^{\Delta}$ from (18) we get

$$E[e_j e_i^*] = -\mathbf{w}_0^{*\Delta}(\mathbf{e}_\Delta^*\mathbf{H}_1^* - \mathbf{e}_\Delta^*\mathbf{H}_1^*(\mathbf{H}_1\mathbf{H}_1^* + \frac{1}{SNR}\mathbf{I}_{\Delta+1})^{-1}(\mathbf{H}_1\mathbf{H}_1^* + \frac{1}{SNR}\mathbf{I}_{\Delta+1}))\mathbf{J} = 0 \tag{25}$$

Similarly, for $j < i$, $E[e_j e_i^*] = (E[e_i e_j^*])^* = 0$. Finally, for $j = i$, we get:

$$E[e_i e_i^*] = E[e_i x^*(i-\Delta)] = 1 - \mathbf{w}_0^{*\Delta}\mathbf{H}_1\mathbf{e}_\Delta = 1 - \mathbf{e}_\Delta^*\mathbf{H}_1^*(\mathbf{H}_1\mathbf{H}_1^* + \frac{1}{SNR}\mathbf{I}_{\Delta+1})^{-1}\mathbf{H}_1\mathbf{e}_\Delta \; . \tag{26}$$

In summary, the error correlation sequence is equal to

$$E[e_i e_j^*] = (1 - \mathbf{e}_\Delta^*\mathbf{H}_1^*(\mathbf{H}_1\mathbf{H}_1^* + \frac{1}{SNR}\mathbf{I}_{\Delta+1})^{-1}\mathbf{H}_1\mathbf{e}_\Delta)\delta_{ij} \; ,$$

where δ_{ij} is the Kronecker delta function (equal to zero for $i \neq j$ and 1 when $i = j$).

References

[1] N. Al-Dhahir and J.M.Cioffi. MMSE Decision Feedback Equalizers : Finite-Length Results. *IEEE Transactions on Information Theory*, 41(4):961–976, July 1995.

[2] ETSI Normalization Committee. Channel Models for HIPERLAN/2 In Different Indoor Scenarios. Document No. 3ERI085B. Sophia-Antiplois, France, 1998.

[3] J.M.Cioffi and G. Dudevoir and M.V. Eyuboglu and G.D. Forney, Jr. Minimum Mean-Square-Error Decision Feedback Equalization and Coding - Parts I and II. *IEEE Transactions on Communications*, October 1995.

[4] G. Strang. *Linear Algebra and Its Applications*. Academic Press, 1988.

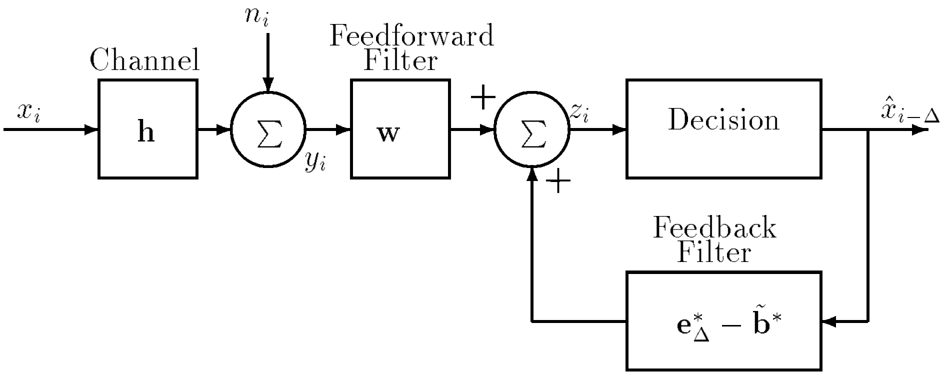

Figure 1: Block diagram of the MMSE–DFE

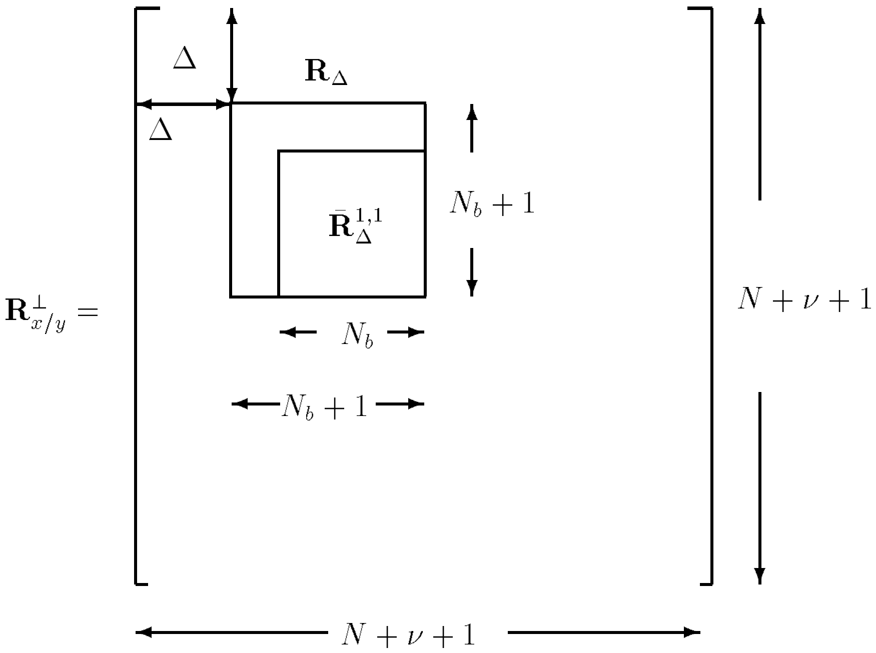

Figure 2: Algebraic interpretation of the parameters N, N_b, and Δ

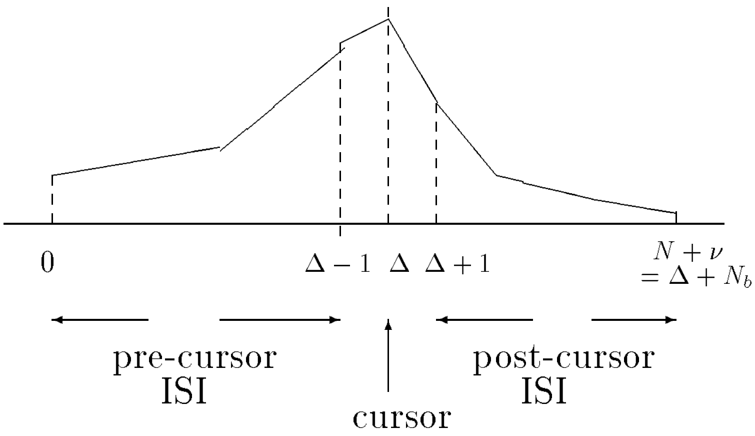

Figure 3: A typical impulse response at the output of the feedforward filter illustrating the assumed relation between the parameters N, ν, N_b, and Δ

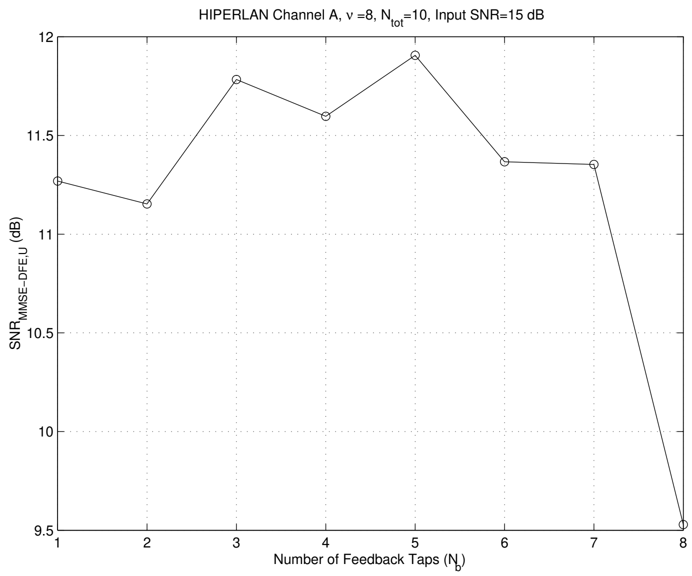

Figure 4: Example of the Variation of Optimal $SNR_{MMSE-DFE,U}$ with N_b for $N_{tot} = 10$, Input SNR=15 dB on a HIPERLAN A Channel

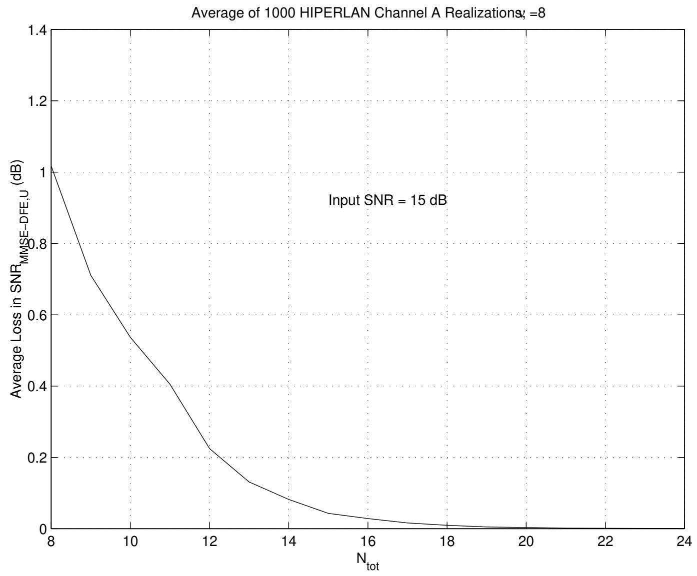

Figure 5: Average Loss in $SNR_{MMSE-DFE,U}$ for the Suboptimal Algorithm from Exhaustive Search for the HIPERLAN A Channel

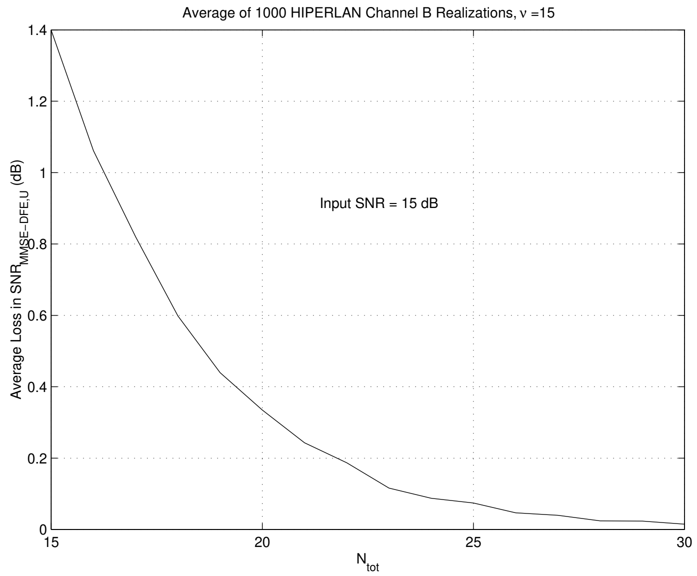

Figure 6: Average Loss in $SNR_{MMSE-DFE,U}$ for the Suboptimal Algorithm from Exhaustive Search for the HIPERLAN B Channel