

Optimal MSE DFE for Multicarrier Communication Systems

Vinod M. Prabhakaran
 Dept. of Electrical Engineering
 Indian Institute of Science
 Bangalore 560012, India
 Email: mp@ee.iisc.ernet.in

V. Umapathi Reddy
 Dept. of Electrical Communication
 Engineering
 Indian Institute of Science
 Bangalore 560012, India
 Email: vur@ece.iisc.ernet.in

Abstract — Block transmission systems which consist of a precoder transmit-filterbank and a block receiver have been used to combat frequency selective fading in wireless channels. Several authors have addressed the problem of decision feedback equalizer (DFE) receivers for such systems. We attempt this problem in a general framework and obtain the optimal solution for the minimum mean-square error (MMSE) DFEs of finite complexity. We show that several previously derived results are special cases of our general solution.

I. INTRODUCTION

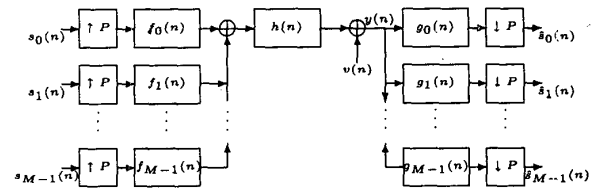
Block transmission systems, which consist of a precoder transmit filterbank and a block receiver, are at the heart of many modern communication techniques for digital audio broadcast and digital subscriber line applications. The most popular among such systems is the OFDM/DMT used in Digital Audio Broadcasting.

In systems using an FIR transmit filter, decision feedback equalizers (DFEs) are known to offer low bit-error rates [1]. Previous results with DFEs in block transmission schemes were obtained with some assumptions on the transmitter or receiver structures. An infinite-length DFE receiver was assumed in [2] and only an iterative solution was obtained. In [3], the author considers a specific form for the precoder where redundancy in the form of a known block of symbols is introduced between data blocks. DFE for a special class of precoder filterbanks was derived in [4]. The block length was chosen to be at least equal to the sum of the impulse response lengths of the channel and the precoder filter. In this paper, we attempt a general framework for minimum mean-square error (MMSE) DFEs of finite complexity for block transmission systems. No assumptions are made on the precoder filterbank, the channel (except that it is FIR) and the block size.

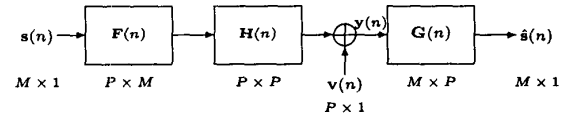
Our work is closely related to [5] in which FIR MMSE DFEs are derived for systems with an FIR transmit filter. We show that our solution reduces to that in [5] when the filterbank is made of just one filter. We also show that our solution reduces to the previously known results in [4] under the conditions imposed there on the precoder filterbank.

II. PROBLEM STATEMENT

Consider Fig. 1(a) which shows the discrete-time model of a baseband block transmission scheme with a linear equalizer. It was shown in [6] that most of the currently used block transmission schemes can be put in this form. It is straightforward to show [6] that Fig. 1(b) is an equivalent model. Here $\mathbf{s}(n) = [s_0(n) \cdots s_{M-1}(n)]^T$, $\{\mathbf{F}(n)\}_{ij} = f_j(nP + i)$; $0 \leq i \leq P - 1$; $0 \leq j \leq M - 1$, $\mathbf{H}(n) =$



(a) Block Transmission System



(b) Equivalent Model

Figure 1: Linear Equalizer

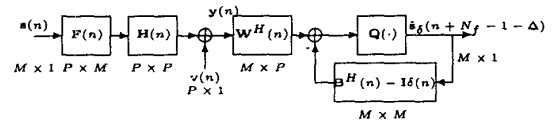


Figure 2: Decision Feedback Equalizer

$$\begin{pmatrix} h(nP) & h(nP+1) & \cdots & h(nP+P-1) \\ h(nP-1) & h(nP) & \cdots & h(nP+P-2) \\ \vdots & \vdots & \ddots & \vdots \\ h(nP-P+1) & h(nP-P+2) & \cdots & h(nP) \end{pmatrix},$$

$\mathbf{v}(n) = [v(nP) \cdots v(nP+P-1)]^T$, and $\{\mathbf{G}(n)\}_{ij} = g_i(nP - j)$; $0 \leq i \leq M - 1$; $0 \leq j \leq P - 1$. It may be noted that the receiver section in Fig. 1(a) is in a form different from that given in [6], but the expression for $\mathbf{G}(n)$ can similarly be found.

In Fig. 2, the linear equalizer is replaced by a DFE. It is easy to show that practical implementation is possible only if we assume that $\mathbf{B}(n) = \mathbf{0}$, $n < 0$, and $\mathbf{B}(0)$ is of the form $\mathbf{B}(0) = \mathbf{P}^T \mathbf{L} \mathbf{P}$, where \mathbf{L} is lower-triangular with ones along the diagonal and \mathbf{P} is a permutation matrix. The order in which the decisions are made for a block of data is determined by the permutation matrix \mathbf{P} . We will also make the following finite complexity assumption: the feed-forward filter, $\mathbf{W}(n) = \mathbf{0}$, $n \notin \{-(N_f - 1), \dots, 0\}$ and the feedback filter, $\mathbf{B}(n) = \mathbf{0}$, $n \notin \{0, \dots, N_b\}$.

Let $\mathbf{A}(n) \triangleq \mathbf{H}(n) * \mathbf{F}(n) = \sum_{k=-\infty}^{\infty} \mathbf{H}(n-k)\mathbf{F}(k)$. We will assume that $\mathbf{A}(n) = \mathbf{0}$, $n \notin \{0, \dots, \nu\}$. This amounts to

assuming that the transmit filters and the channel are causal FIR. If the channel impulse response length is L_h , and $L_f = \max_{0 \leq i \leq M-1} \{\text{length of } f_i(n)\}$, then $\nu = (L_f + L_h - 1)/P - 1$. The parameters Δ ($0 \leq \Delta \leq N_f + \nu - 1$) and δ ($0 \leq \delta \leq M - 1$) together determine the delay of the feed-forward filter. We define

$$\mathbf{s}_\delta(n) = [s_\delta(n) \ s_{\delta+1}(n) \ \dots \ s_{M-1}(n) \ s_0(n-1) \ \dots \ s_{\delta-1}(n-1)]^T$$

We also assume that $\mathbf{s}(n)$ and $\mathbf{v}(n)$ are zero-mean and uncorrelated vector random processes.

Referring to Fig. 2, we have

$$\mathbf{y}(n) = \mathbf{A}(n) * \mathbf{s}(n) + \mathbf{v}(n) \quad (1)$$

Our goal here is to obtain the optimal solution for the DFE. Though the criterion of optimization should be the minimization of bit-error rate, for reasons of tractability, we minimize the arithmetic mean-square error (MSE) defined as

$$\eta \triangleq \mathbf{E}(\mathbf{e}^H(n)\mathbf{e}(n)) \quad (2)$$

where

$$\mathbf{e}(n) \triangleq \hat{\mathbf{s}}(n + N_f - 1 - \Delta) - \hat{\mathbf{s}}_\delta(n + N_f - 1 - \Delta) \quad (3)$$

Thus, the problem is to choose $\mathbf{W}(n)$, $\mathbf{B}(n)$, Δ , δ , and \mathbf{P} so as to minimize η

$$\eta_{min} \triangleq \min_{\mathbf{W}(n), \mathbf{B}(n), \Delta, \delta, \mathbf{P}} \mathbf{E}(\mathbf{e}^H(n)\mathbf{e}(n))$$

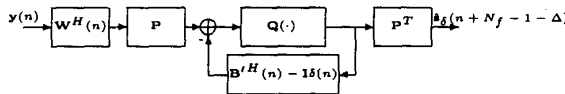


Figure 3: Decision Feedback Equalizer - An equivalent form

It is easy to show that Fig. 2 is equivalent to Fig. 3 if $\mathbf{B}'(n) = \mathbf{P}\mathbf{B}(n)\mathbf{P}^T$. \mathbf{P} is a permutation matrix and $\mathbf{B}'(0)$ is lower-triangular with ones along the diagonal. We can interpret this figure as follows. The permutation matrix \mathbf{P} decides the order in which symbol decisions are made in each block by permuting the sub-channels. The feedback loop makes decisions in order - i.e., the higher indexed decisions are made use of in making decisions of the lower indexed channels. \mathbf{P}^T restores the order. \mathbf{P} is also a parameter to be optimized because, in general, the choice of the order in which decisions are made will affect the MSE. When $\mathbf{P} = \mathbf{I}$ and $\delta = 0$, the problem reduces to one of the problems solved in [9] in the context of multi-user communication¹.

III. OPTIMAL SOLUTION

Assuming that the past decisions are correct, we can express (3) as

$$\mathbf{e}(n) = \bar{\mathbf{B}}^H \bar{\mathbf{s}}(n) - \bar{\mathbf{W}}^H \bar{\mathbf{y}}(n) \quad (4)$$

where $\bar{\mathbf{B}} \triangleq [0_{M \times \Delta M + \delta} \ \mathbf{B}^H(0) \ \mathbf{B}^H(1) \ \dots \ \mathbf{B}^H(N_b) \ \mathbf{0}_{M \times S}]^H$, $\bar{\mathbf{W}} \triangleq [\mathbf{W}^H(-N_f - 1) \ \dots \ \mathbf{W}^H(0)]^H$, $\bar{\mathbf{s}}(n) \triangleq [\mathbf{s}^H(n + (N_f -$

¹[9] appeared after this paper was submitted. Revisions were made to include connections with this important reference.

$1) \ \dots \ \mathbf{s}^H(n - \nu)]^H$, $\bar{\mathbf{y}}(n) \triangleq [\mathbf{y}^H(n + (N_f - 1)) \ \dots \ \mathbf{y}^H(n)]^H$, and $S \triangleq (N_f - 1 + \nu - N_b)M - \Delta M - \delta$ ²

From (1) and the above definitions, we can write

$$\bar{\mathbf{y}}(n) = \bar{\mathbf{A}}\bar{\mathbf{s}}(n) + \bar{\mathbf{v}}(n) \quad (5)$$

where

$$\bar{\mathbf{A}} \triangleq \begin{pmatrix} \mathbf{A}(0) & \mathbf{A}(1) & \dots & \mathbf{A}(\nu) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}(0) & \mathbf{A}(1) & \dots & \mathbf{A}(\nu) & \dots & \mathbf{0} \\ \vdots & & & \ddots & & & \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{A}(0) & \mathbf{A}(1) & \dots & \mathbf{A}(\nu) \end{pmatrix}$$

and

$$\bar{\mathbf{v}}(n) \triangleq [\mathbf{v}^H(n + (N_f - 1)) \ \dots \ \mathbf{v}^H(n)]^H$$

Substituting (4) and (5) in (2), we get

$$\begin{aligned} \eta &= \mathbf{E}(\mathbf{e}^H(n)\mathbf{e}(n)) = \text{tr}(\mathbf{E}(\mathbf{e}(n)\mathbf{e}^H(n))) \\ &= \text{tr}(\bar{\mathbf{B}}^H \mathbf{R}_{\bar{\mathbf{s}}\bar{\mathbf{s}}} \bar{\mathbf{B}} + \bar{\mathbf{W}}^H \mathbf{R}_{\bar{\mathbf{y}}\bar{\mathbf{y}}} \bar{\mathbf{W}} - \bar{\mathbf{B}}^H \mathbf{R}_{\bar{\mathbf{y}}\bar{\mathbf{s}}} \bar{\mathbf{W}} - \bar{\mathbf{W}}^H \mathbf{R}_{\bar{\mathbf{s}}\bar{\mathbf{y}}} \bar{\mathbf{B}}) \end{aligned} \quad (6)$$

where $\mathbf{R}_{\bar{\mathbf{s}}\bar{\mathbf{s}}} \triangleq \mathbf{E}(\bar{\mathbf{s}}(n)\bar{\mathbf{s}}^H(n))$, $\mathbf{R}_{\bar{\mathbf{y}}\bar{\mathbf{y}}} \triangleq \mathbf{E}(\bar{\mathbf{y}}(n)\bar{\mathbf{y}}^H(n)) = \bar{\mathbf{A}}\mathbf{R}_{\bar{\mathbf{s}}\bar{\mathbf{s}}}\bar{\mathbf{A}}^H + \mathbf{R}_{\bar{\mathbf{v}}\bar{\mathbf{v}}}$, $\mathbf{R}_{\bar{\mathbf{y}}\bar{\mathbf{s}}} \triangleq \mathbf{E}(\bar{\mathbf{y}}(n)\bar{\mathbf{s}}^H(n)) = \mathbf{R}_{\bar{\mathbf{s}}\bar{\mathbf{y}}}\bar{\mathbf{A}}^H$, $\mathbf{R}_{\bar{\mathbf{s}}\bar{\mathbf{y}}} \triangleq \mathbf{E}(\bar{\mathbf{s}}(n)\bar{\mathbf{y}}^H(n)) = \mathbf{R}_{\bar{\mathbf{y}}\bar{\mathbf{s}}}\bar{\mathbf{A}}$, and $\mathbf{R}_{\bar{\mathbf{v}}\bar{\mathbf{v}}} \triangleq \mathbf{E}(\bar{\mathbf{v}}(n)\bar{\mathbf{v}}^H(n))$. We will assume that these correlation matrices are non-singular.

Taking the gradient of η with respect to $\bar{\mathbf{W}}$ and equating it to the zero matrix, we get the optimum $\bar{\mathbf{W}}$

$$\bar{\mathbf{W}}_{opt} = \mathbf{R}_{\bar{\mathbf{y}}\bar{\mathbf{y}}}^{-1} \mathbf{R}_{\bar{\mathbf{y}}\bar{\mathbf{s}}} \bar{\mathbf{B}}_{opt} \quad (7)$$

Substituting the optimal feedforward filter of (7) in (6) and using the matrix inversion lemma, we can show that the MMSE for a given $\bar{\mathbf{B}}$ is

$$\eta = \text{tr}(\bar{\mathbf{B}}^H (\mathbf{R}_{\bar{\mathbf{s}}\bar{\mathbf{s}}} - \mathbf{R}_{\bar{\mathbf{y}}\bar{\mathbf{s}}}\mathbf{R}_{\bar{\mathbf{y}}\bar{\mathbf{y}}}^{-1}\mathbf{R}_{\bar{\mathbf{s}}\bar{\mathbf{y}}})\bar{\mathbf{B}}) \quad (8)$$

$$= \text{tr}(\bar{\mathbf{B}}^H (\mathbf{R}_{\bar{\mathbf{s}}\bar{\mathbf{s}}}^{-1} + \bar{\mathbf{A}}^H \mathbf{R}_{\bar{\mathbf{v}}\bar{\mathbf{v}}}^{-1} \bar{\mathbf{A}})^{-1} \bar{\mathbf{B}}) \quad (9)$$

$$= \text{tr}(\mathbf{B}^H \mathbf{R}_{\Delta, \delta, \mathbf{P}} \mathbf{B}) \quad (10)$$

where $\bar{\mathbf{B}} \triangleq [\mathbf{B}'^H(0) \ \dots \ \mathbf{B}'^H(N_b)]^H$, and

$$\mathbf{R}_{\Delta, \delta, \mathbf{P}} \triangleq \begin{bmatrix} \mathbf{0}_{(N_b+1)M \times \Delta M + \delta} & \mathbf{I}_{N_b+1} \otimes \mathbf{P} & \mathbf{0}_{(N_b+1)M \times S} \\ (\mathbf{R}_{\bar{\mathbf{s}}\bar{\mathbf{s}}}^{-1} + \bar{\mathbf{A}}^H \mathbf{R}_{\bar{\mathbf{v}}\bar{\mathbf{v}}}^{-1} \bar{\mathbf{A}})^{-1} & \begin{pmatrix} \mathbf{0}_{\Delta M + \delta \times (N_b+1)M} \\ \mathbf{I}_{N_b+1} \otimes \mathbf{P}^T \\ \mathbf{0}_{S \times (N_b+1)M} \end{pmatrix} \end{bmatrix} \quad (11)$$

\otimes denotes the Kronecker tensor product. Note that $\mathbf{R}_{\Delta, \delta, \mathbf{P}}$ is Hermitian.

Our objective is to find \mathbf{B} such that (10) is minimized. $\mathbf{B}'(0)$ is restricted to be lower-triangular with ones along the diagonal. The solution is given by the following theorem.

Theorem 1 (Modified Linear Vector Predictor) *If $\mathbf{z}(n)$ is a non-deterministic [8] zero-mean $M \times 1$ vector random process, then the optimal predictor $\Lambda(n)$, such that $\Lambda(n) = \mathbf{0}$, $n \notin \{0, \dots, N\}$ and $\Lambda(0)$ is lower-triangular with ones along the diagonal, which minimizes*

$$\zeta = \text{tr}(\mathbf{E}(\varepsilon(n)\varepsilon^H(n))),$$

²A negative value for S indicates that the feedback filter is longer than necessary. In that case, we redefine N_b to remove the redundant terms in $\mathbf{B}(n)$ thereby making $S = 0$.

where $\varepsilon(n) = \Lambda^H(n) * \mathbf{z}(n)$, is

$$[\Lambda_{opt}^H(0) \cdots \Lambda_{opt}^H(N)]^H = \begin{pmatrix} \mathbf{I}_M \\ -\mathbf{R}_{1:N,1:N}^{-1} \mathbf{R}_{1:N,0} \end{pmatrix} \mathbf{L}^{-1}$$

Here

$$\mathbf{R}_{1:N,1:N} = \begin{pmatrix} \mathbf{R}_{11} & \cdots & \mathbf{R}_{1N} \\ \vdots & \ddots & \vdots \\ \mathbf{R}_{N1} & \cdots & \mathbf{R}_{NN} \end{pmatrix}$$

$\mathbf{R}_{1:N,0} = [\mathbf{R}_{10}^H \cdots \mathbf{R}_{N0}^H]^H$, and $\mathbf{R}_{ij} = \mathbf{E}(\mathbf{z}(n-i)\mathbf{z}^H(n-j))$. \mathbf{L} is given by the Cholesky decomposition

$$\mathbf{R}_{00} - \mathbf{R}_{1:N,0}^H \mathbf{R}_{1:N,1:N}^{-1} \mathbf{R}_{1:N,0} = \mathbf{L}^H \mathbf{D} \mathbf{L} \quad (12)$$

Also

$$\zeta_{min} = \text{tr}(\mathbf{D})$$

Proof: We have

$$\zeta = \text{tr} \left(\sum_{i=0}^N \sum_{j=0}^N \Lambda^H(i) \mathbf{R}_{ij} \Lambda(j) \right) \quad (13)$$

$\nabla_{\Lambda(k)} \zeta = \mathbf{0}$, $1 \leq k \leq N$ gives

$$\sum_{j=0}^N \mathbf{R}_{kj} \Lambda_{opt}(j) = \mathbf{0}, \quad 1 \leq k \leq N$$

where we have used the fact that $\mathbf{R}_{ij} = \mathbf{R}_{ji}^H$, $0 \leq i \leq N$, $0 \leq j \leq N$. We can put this in the form

$$\mathbf{R}_{1:N,1:N} \begin{pmatrix} \Lambda_{opt}(1) \\ \vdots \\ \Lambda_{opt}(N) \end{pmatrix} = -\mathbf{R}_{1:N,0} \Lambda_{opt}(0)$$

Since $\mathbf{z}(n)$ is non-deterministic, $\mathbf{R}_{1:N,1:N}$ is positive-definite [8] and hence non-singular. So we have

$$\begin{pmatrix} \Lambda_{opt}(1) \\ \vdots \\ \Lambda_{opt}(N) \end{pmatrix} = -\mathbf{R}_{1:N,1:N}^{-1} \mathbf{R}_{1:N,0} \Lambda_{opt}(0)$$

Using the optimum setting for $\Lambda(1), \dots, \Lambda(N)$ in (13),

$$\zeta = \text{tr}(\Lambda^H(0)(\mathbf{R}_{00} - \mathbf{R}_{1:N,0}^H \mathbf{R}_{1:N,1:N}^{-1} \mathbf{R}_{1:N,0}) \Lambda(0)) \quad (14)$$

Since $\mathbf{z}(n)$ is non-deterministic, the $(N+1)M \times (N+1)M$ matrix $\begin{pmatrix} \mathbf{R}_{00} & \mathbf{R}_{1:N,0}^H \\ \mathbf{R}_{1:N,0} & \mathbf{R}_{1:N,1:N} \end{pmatrix}$ is positive definite [8]. Hence, the Schur complement of this matrix with respect to $\mathbf{R}_{1:N,1:N}$, namely $\mathbf{R}_{00} - \mathbf{R}_{1:N,0}^H \mathbf{R}_{1:N,1:N}^{-1} \mathbf{R}_{1:N,0}$ is also positive-definite. So the Cholesky factorization in (12) exists. It is now easy to see that $\Lambda_{opt}(0)$ is \mathbf{L}^{-1} . Substituting in (14),

$$\zeta_{min} = \text{tr}(\mathbf{D})$$

The optimal \mathbf{B} can now be written by rewriting (13) as

$$\zeta = \text{tr}(\bar{\Lambda}^H \begin{pmatrix} \mathbf{R}_{00} & \mathbf{R}_{1:N,0}^H \\ \mathbf{R}_{1:N,0} & \mathbf{R}_{1:N,1:N} \end{pmatrix} \bar{\Lambda})$$

where $\bar{\Lambda} = [\Lambda^H(0) \cdots \Lambda^H(N)]^H$ and making the following identifications:

$$\begin{aligned} N &= N_b \\ [\Lambda^H(0) \cdots \Lambda^H(N_b)]^H &= \mathbf{B} \\ \begin{pmatrix} \mathbf{R}_{00} & \mathbf{R}_{1:N_b,0}^H \\ \mathbf{R}_{1:N_b,0} & \mathbf{R}_{1:N_b,1:N_b} \end{pmatrix} &= \mathbf{R}_{\Delta,\delta,\mathbf{P}} \end{aligned}$$

The final equation is possible since $\mathbf{R}_{\Delta,\delta,\mathbf{P}}$ is Hermitian. We make the assumption that $\mathbf{R}_{\Delta,\delta,\mathbf{P}}$ and $\mathbf{R}_{1:N_b,1:N_b}$ are positive-definite. Then, the optimum \mathbf{B} for a fixed Δ, δ , and \mathbf{P} is³

$$\mathbf{B}_{\Delta,\delta,\mathbf{P}} = \begin{pmatrix} \mathbf{I}_M \\ -\mathbf{R}_{1:N_b,1:N_b}^{-1} \mathbf{R}_{1:N_b,0} \end{pmatrix} \mathbf{L}_{\Delta,\delta,\mathbf{P}}^{-1} \quad (15)$$

where $\mathbf{L}_{\Delta,\delta,\mathbf{P}}$ is given by the Cholesky decomposition

$$\mathbf{R}_{00} - \mathbf{R}_{1:N_b,0}^H \mathbf{R}_{1:N_b,1:N_b}^{-1} \mathbf{R}_{1:N_b,0} = \mathbf{L}_{\Delta,\delta,\mathbf{P}}^H \mathbf{D}_{\Delta,\delta,\mathbf{P}} \mathbf{L}_{\Delta,\delta,\mathbf{P}} \quad (16)$$

Also

$$\eta_{\Delta,\delta,\mathbf{P}} = \text{tr}(\mathbf{D}_{\Delta,\delta,\mathbf{P}}) \quad (17)$$

Now, we have to solve for $(\Delta_{opt}, \delta_{opt}, \mathbf{P}_{opt})$ which minimizes $\eta_{\Delta,\delta,\mathbf{P}}$

$$\eta_{min} = \min_{\substack{0 \leq \Delta M + \delta \leq (N_f - 1 + \nu)M \\ \mathbf{P}}} \text{tr}(\mathbf{D}_{\Delta,\delta,\mathbf{P}}) \quad (18)$$

It appears that there is no closed-form expression possible for the optimum settings of these parameters. For a given \mathbf{P} , (18) suggests an exhaustive search for the optimum value of Δ and δ .

Since even for small values of M , the number of possible \mathbf{P} 's is large ($M!$), an exhaustive search for the optimal \mathbf{P} may not be possible. A heuristic approach can be used to get a reasonable \mathbf{P} . For instance, in multicarrier systems, if the inter-block interference is not excessive, estimates of the sub-channel SNRs, if available, can be used as a criterion for choosing the order in which decisions are made within a block. We may choose to make decisions on the strongest sub-channel first, followed by the next strongest, and so on.

If an exhaustive search is to be carried out for \mathbf{P} , we need only calculate the $\mathbf{R}_{00} - \mathbf{R}_{1:N_b,0}^H \mathbf{R}_{1:N_b,1:N_b}^{-1} \mathbf{R}_{1:N_b,0}$ matrix for $\mathbf{P} = \mathbf{I}$. We can show that for other values of \mathbf{P} it can be obtained by pre-multiplying the above matrix by \mathbf{P} and post-multiplying by \mathbf{P}^T .

If $N_b = \nu$, the search for optimal Δ and δ for a given \mathbf{P} involves performing only one Cholesky decomposition (that of $(\mathbf{I}_{N_f+\nu} \otimes \mathbf{P})(\mathbf{R}_{\bar{\delta}\bar{\delta}}^{-1} + \bar{\mathbf{A}}^H \mathbf{R}_{\bar{\nu}\bar{\nu}}^{-1} \bar{\mathbf{A}})^{-1} (\mathbf{I}_{N_f+\nu} \otimes \mathbf{P}^T)$). This can be shown along the lines of a similar result in [5].

The optimal feedback filter \mathbf{B}_{opt} is given by (15) when the optimal settings are substituted. From (7), the optimum feed-forward filter is

$$\bar{\mathbf{W}}_{opt} = (\bar{\mathbf{A}} \mathbf{R}_{\bar{\delta}\bar{\delta}} \bar{\mathbf{A}}^H + \mathbf{R}_{\bar{\nu}\bar{\nu}})^{-1} \bar{\mathbf{A}} \mathbf{R}_{\bar{\delta}\bar{\delta}} \bar{\mathbf{B}}_{opt} \quad (19)$$

where

$$\bar{\mathbf{B}}_{opt} = [\mathbf{0}_{M \times \Delta_{opt} M + \delta} \quad (\mathbf{I}_{N_b+1} \otimes \mathbf{P}_{opt}) \mathbf{B}_{opt}^H (\mathbf{I}_{N_b+1} \otimes \mathbf{P}_{opt}^T) \quad \mathbf{0}_{M \times S_{opt}}]^H$$

and $S_{opt} = (N_f - 1 + \nu - N_b - \Delta_{opt})M - \delta_{opt}$.

The MMSE is given by (17) with the optimal settings substituted. \square

Fast algorithms for performing the Cholesky decompositions, which exploit the structure in the matrices, appear in [9].

³The subscript indicates the dependence of this solution on Δ, δ , and \mathbf{P} .

IV. DISCUSSION OF SOME SPECIAL CASES

MMSE Linear Equalizer

We can derive the MMSE linear equalizer by setting $\mathbf{B}(n) = \mathbf{I}\delta(n)$. We can choose $\mathbf{P} = \mathbf{I}$ without any loss of generality.

$$\mathbf{B}(n) = \mathbf{I}\delta(n) \Rightarrow \bar{\mathbf{B}} = \mathbf{E}_\Delta \quad (20)$$

where $\mathbf{E}_\Delta = [\mathbf{0} \cdots \mathbf{0} \mathbf{I} \mathbf{0} \cdots \mathbf{0}]^T$, with the $M \times M$ identity matrix being positioned after Δ zero matrices, each of size $M \times M$. The optimum linear equalizer for a given Δ can then be written from (7) as

$$\bar{\mathbf{W}}_\Delta^H = (\mathbf{R}_{\bar{y}\bar{y}}^{-1} \mathbf{R}_{\bar{y}\bar{s}} \mathbf{E}_\Delta)^H \quad (21)$$

$$= \mathbf{E}_\Delta^H \mathbf{R}_{\bar{s}\bar{s}} \tilde{\mathbf{A}}^H (\tilde{\mathbf{A}} \mathbf{R}_{\bar{s}\bar{s}} \tilde{\mathbf{A}}^H + \mathbf{R}_{\bar{v}\bar{v}})^{-1} \quad (22)$$

and the corresponding MSE is

$$\eta_\Delta = \text{tr}(\mathbf{E}_\Delta^H \mathbf{R}_{\bar{s}\bar{s}} \mathbf{E}_\Delta - \bar{\mathbf{W}}^H \mathbf{R}_{\bar{y}\bar{y}} \bar{\mathbf{W}}) \quad (23)$$

When $\mathbf{s}(n)$ is white, this reduces to the expressions derived in [7].⁴

A Single FIR Filter Precoder ($M = 1$)

When $M = 1$, the transmitter is simply an FIR filter. Now the feed-forward filter $\mathbf{w}(n)$ and feedback filter $\mathbf{b}(n)$ are vectors. We will indicate this by using lower-case letters. Clearly, $\delta = 0$ and $\mathbf{P} = [1]$.

Let the $N_b \times N_b$ matrix \mathbf{R} and the $N_b \times 1$ vector \mathbf{r} be submatrices of the $(N_b + 1) \times (N_b + 1)$ matrix \mathbf{R}_Δ defined in (11),

$$\mathbf{R}_\Delta = \begin{pmatrix} R_{00} & \mathbf{r}^H \\ \mathbf{r} & \mathbf{R} \end{pmatrix}$$

From (15), the optimum feedback filter for a given Δ is

$$\mathbf{b}_\Delta = \begin{pmatrix} 1 \\ -\mathbf{R}^{-1} \mathbf{r} \end{pmatrix} \quad (24)$$

The optimum feed-forward filter, for a fixed Δ , is then given by (7)

$$\bar{\mathbf{w}}_\Delta^H = \mathbf{b}_\Delta^H \mathbf{R}_{\bar{s}\bar{s}} \tilde{\mathbf{A}}^H (\tilde{\mathbf{A}} \mathbf{R}_{\bar{s}\bar{s}} \tilde{\mathbf{A}}^H + \mathbf{R}_{\bar{v}\bar{v}})^{-1} \quad (25)$$

and the corresponding MSE by

$$\eta_\Delta = R_{00} - \mathbf{r}^H \mathbf{R}^{-1} \mathbf{r} \quad (26)$$

Using the identity

$$\mathbf{R}_\Delta^{-1} = \begin{pmatrix} (R_{00} - \mathbf{r}^H \mathbf{R}^{-1} \mathbf{r})^{-1} & (\mathbf{r}^H \mathbf{R}^{-1} \mathbf{r} - R_{00})^{-1} \mathbf{r}^H \mathbf{R}^{-1} \\ (\mathbf{r} \mathbf{R}_{00}^{-1} \mathbf{r}^H - R)^{-1} \mathbf{r} \mathbf{R}_{00}^{-1} & (R - \mathbf{r} \mathbf{R}_{00}^{-1} \mathbf{r}^H)^{-1} \end{pmatrix}$$

it can be shown that

$$\mathbf{b}_\Delta = \frac{\mathbf{R}_\Delta^{-1} \mathbf{e}_0}{\mathbf{e}_0^H \mathbf{R}_\Delta^{-1} \mathbf{e}_0} \quad (27)$$

and

$$\eta_\Delta = \frac{1}{\mathbf{e}_0^H \mathbf{R}_\Delta^{-1} \mathbf{e}_0} \quad (28)$$

This is precisely the solution obtained in [5].

⁴In [7], the expressions were derived for the optimum discrete-time fixed order FIR linear equalizer minimising the MSE in the presence of near- and far-end crosstalk.

Trailing-zeros Case

In [4], it is assumed that the precoder filters are such that $P = L_h + L_f - 1$ which gives $\nu = 0$. Then, the best choice of the feed-forward and feedback lengths is $N_f = 1$ and $N_b = 0$, and consequently the only choice of Δ and δ are zero. For simplicity we choose $\mathbf{P} = \mathbf{I}$.

Now from (11),

$$\mathbf{R}_\Delta = (\mathbf{R}_{\bar{s}\bar{s}}^{-1} + \tilde{\mathbf{A}}^H(0) \mathbf{R}_{\bar{v}\bar{v}}^{-1} \tilde{\mathbf{A}}(0))^{-1} \quad (29)$$

where $\mathbf{R}_{\bar{s}\bar{s}} = \mathbf{E}(\mathbf{s}(n) \mathbf{s}^H(n))$, and $\mathbf{R}_{\bar{v}\bar{v}} = \mathbf{E}(\mathbf{v}(n) \mathbf{v}^H(n))$. So, \mathbf{R}_Δ is $M \times M$ which implies that $\mathbf{R}_{00} = \mathbf{R}_\Delta$. Then (16) becomes

$$\mathbf{R}_{00} = (\mathbf{R}_{\bar{s}\bar{s}}^{-1} + \tilde{\mathbf{A}}^H(0) \mathbf{R}_{\bar{v}\bar{v}}^{-1} \tilde{\mathbf{A}}(0))^{-1} = \mathbf{L}_0^H \mathbf{D}_0 \mathbf{L}_0 \quad (30)$$

The optimum feedback and feed-forward filters from (15) and (7) respectively, are

$$\mathbf{B}_{opt} = \mathbf{L}_0^{-1} \quad (31)$$

$$\bar{\mathbf{W}}_{opt} = \mathbf{R}_{\bar{y}\bar{y}}^{-1} \mathbf{R}_{\bar{y}\bar{s}} \mathbf{B}_{opt} \quad (32)$$

and the MMSE is

$$\eta_{min} = \text{tr}(\mathbf{D}_0) \quad (33)$$

The expressions (31) to (33) match the results in [4].

V. CONCLUSIONS

We have considered the general case of decision feedback equalizers of finite complexity for block transmission systems and obtained the optimal solution. We have shown that several results derived elsewhere are special cases of our general solution.

ACKNOWLEDGMENTS

The authors would like to thank Prof. Thomas Kailath for a helpful discussion during the preparation of the paper.

REFERENCES

- [1] N. Al-Dhahir and J. M. Cioffi, "Block transmission over dispersive channels: Transmit filter optimization and realization, and MMSE-DFE receiver performance", *IEEE Trans. Inform. Theory*, vol. 42, pp. 137-160, Jan. 1996.
- [2] J. Yang and S. Roy, "Joint transmitter-receiver optimization for multi-input multi-output systems with decision feedback," *IEEE Trans. Inform. Theory* vol. 40, pp. 1334-1347, Sep. 1994.
- [3] G. K. Kaleh, "Channel equalization for block transmission systems," *IEEE J. Select. Areas Commun.* vol 13, pp. 110-121, Jan 1995.
- [4] A. Stamoulis, G.B. Giannakis, and A. Scaglione, "Self-recovering transceivers for block transmissions: filterbank precoders and decision-feedback equalizers," *Proc. of the 2nd IEEE-SP Workshop on Signal Proc. Advances in Wireless Communications*, Annapolis, MD, May 1999.
- [5] N. Al-Dhahir and J. M. Cioffi, "MMSE decision-feedback equalizers: Finite-length results," *IEEE Trans. Inform. Theory*, vol. 41, pp. 961-975, Jul. 1995.
- [6] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant filterbank precoders and equalizers Part I: Unification and optimal designs," *IEEE Trans. Sig. Proc.*, vol. 47, pp. 1988-2006, Jul. 1999.
- [7] M. L. Honig, P. Crespo, and K. Steiglitz, "Suppression of near- and far-end crosstalk by linear pre- and post-filtering," *IEEE J. Select. Areas Commun.*, vol. 10, pp. 614-629, Apr. 1992.
- [8] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*, Kluwer Academic Publishers, Boston, 1992
- [9] N. Al-Dhahir and A. H. Sayed, "The Finite-Length Multi-Input Multi-Output MMSE-DFE," *IEEE Trans. Sig. Proc.*, vol. 48, pp. 2921-2936, Oct. 2000.