Minimum-BER Linear-Combiner DFE

S. Chen †, *E.S. Chng* ¶, *B. Mulgrew* ‡ *and G. Gibson* §

† Department of Electrical and Electronic Engineering
University of Portsmouth, Anglesea Building, Portsmouth PO1 3DJ, U.K.

¶ Lab for ABS, RIKEN, 2–1 Hirosawa, Wako-shi, Saitama 351–01, Japan

‡ Department of Electrical Engineering, University of Edinburgh
King's Buildings, Edinburgh EH9 3JL, U.K.

§ Scottish Agricultural Statistics Service, University of Edinburgh
King's Buildings, Edinburgh EH9 3JZ, U.K.

Abstract

The paper derives a minimum bit error rate (BER) solution for the decision feedback equaliser (DFE) that employs a linear combination of the channel observations and the past decisions. We show that by using a geometric translation the DFE is reduced to a simpler linear equaliser. A BER expression for the linear equaliser is obtained under the assumption of linearly separable decision regions, and a method is proposed to optimally set the linear-combiner coefficients of the DFE. This minimum BER solution is superior to the usual minimum mean square error (MSE) solution.

1 Introduction

Decision feedback is a powerful technique for combating channel distortion. The conventional DFE [1] is based on a symbol-decision structure that employs a linear combination of the channel observations and the past decisions. We will call this DFE the linear-combiner DFE, in contrast to other DFE structures that use nonlinear combinations of the channel observations and the past decisions [2-7]. The Wiener or minimum MSE solution is often said to be the optimal solution for the coefficients of the linear-combiner DFE. It is well-known however that the minimum MSE solution does not correspond to the minimum BER solution, the BER being the ultimate performance criterion of equalisation.

The linear-combiner DFE realises a linear decision boundary. Because of decision feedback, the subsets of channel states corresponding to different values of the data symbol are usually linearly separable. We demonstrate that the optimal linear decision boundary can be very different from the decision boundary of the Wiener solution. This shows that significant BER reduction over the Wiener solution is possible.

To facilitate a deep understanding of the linear-combiner DFE, a simple geometric translation is introduced which reduces the DFE to an equivalent linear equaliser "without feedback". A BER expression is then derived. Using this BER estimator as the optimisation criterion, a method is proposed to optimally set the coefficients of the linear-combiner DFE. The decision boundary of this optimal linear-combiner DFE is the best linear approximation to the Bayesian decision boundary and its BER performance is close to that of the Bayesian DFE [2].

Throughout this study, the channel is modelled as a finite impulse response (FIR) filter with the transfer function

$$A(z) = \sum_{i=0}^{n_a-1} a_i z^{-i} \tag{1}$$

where n_a is the length of the channel impulse response and a_i are the channel tap weights. For notational simplicity, the symbol sequence $\{s(k)\}$ is assumed to be independently identically distributed (i.i.d.), taking values from the set $\{\pm 1\}$. The received signal is given by

$$r(k) = \hat{r}(k) + e(k) = \sum_{i=0}^{n_a-1} a_i s(k-i) + e(k) \tag{2}$$

0-7803-3250-4/96$5.00©1996 IEEE

where $\hat{r}(k)$ is the noiseless channel observation, $e(k)$ is an i.i.d. Gaussian noise source with zero mean and variance $\mathrm{E}[e^2(k)] = \sigma_e^2$ and is uncorrelated with $s(k)$. The signal to noise ratio (SNR) of the system is defined as

$$\mathrm{SNR} = \mathrm{E}[\hat{r}(k)]/\mathrm{E}[e^2(k)] = \sigma_s^2 \left(\sum_{i=0}^{n_a-1} a_i^2 \right) / \sigma_e^2 \qquad (3)$$

where $\sigma_s^2 = \mathrm{E}[s^2(k)]$ is the symbol variance.

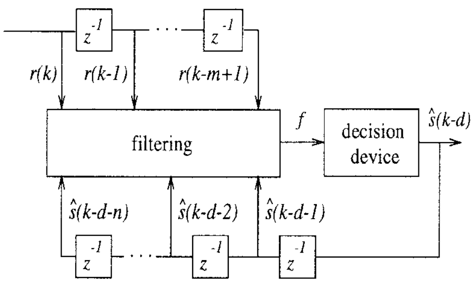

Figure 1: Schematic of DFE.

2 The linear-combiner DFE

The DFE, depicted in Fig.1, uses the information present in the channel output vector

$$\mathbf{r}(k) = [r(k) \cdots r(k-m+1)]^T \qquad (4)$$

and the past detected symbol vector

$$\hat{\mathbf{s}}_b(k) = [\hat{s}(k-d-1) \cdots \hat{s}(k-d-n)]^T \qquad (5)$$

to produce an estimate $\hat{s}(k-d)$ of $s(k-d)$. The integers d, m and n are known as the decision delay, feedforward order and feedback order respectively. Without the loss of generality, $d = n_a - 1$ is chosen to cover the entire channel dispersion, m is related to d by $m = d+1 = n_a$, and n is given by $n = n_a - 1$ (see [2]).

In the linear-combiner DFE, the decision is made by quantizing the filter output

$$f(\mathbf{r}(k), \hat{\mathbf{s}}_b(k)) = \mathbf{w}^T \mathbf{r}(k) + \mathbf{b}^T \hat{\mathbf{s}}_b(k) \qquad (6)$$

where $\mathbf{w} = [w_0 \cdots w_{m-1}]^T$ and $\mathbf{b} = [b_1 \cdots b_n]^T$ are the coefficients of the feedforward and feedback filters respectively. The Wiener solution is often said to provide the optimal $\mathbf{w}$ and $\mathbf{b}$. It is however optimal only in the sense of the minimum MSE solution and is generally not the minimum BER solution.

2.1 Space translation

We first show that the feedback term of the linear-combiner DFE performs a space translation. Define

$$\mathbf{r}'(k) = \mathbf{r}(k) - F'\hat{\mathbf{s}}_b(k) \qquad (7)$$

where F' is the $m \times n$ matrix given by

$$F' = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ a_{n_a-1} & 0 & \ddots & \vdots \\ a_{n_a-2} & a_{n_a-1} & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 \\ a_1 & \cdots & a_{n_a-2} & a_{n_a-1} \end{bmatrix} \qquad (8)$$

Previous research has realized that decision feedback corresponds to this translation [8,4]. We further provide the following new results.

Let $\hat{\mathbf{w}}$ and $\hat{\mathbf{b}}$ be the Wiener solution of $\mathbf{w}$ and $\mathbf{b}$. It can be proved that [9]

$$\hat{\mathbf{w}}^T \mathbf{r}(k) + \hat{\mathbf{b}}^T \hat{\mathbf{s}}_b(k) = \hat{\mathbf{w}}^T \mathbf{r}'(k) \qquad (9)$$

Thus by performing the translation (7), the linear-combiner DFE is reduced to an equivalent and simpler linear equaliser "without decision feedback"

$$f'(\mathbf{r}'(k)) = \mathbf{w}^T \mathbf{r}'(k) \qquad (10)$$

The decision boundary of this equivalent linear equaliser, $\{\mathbf{r}' : \mathbf{w}^T \mathbf{r}' = 0\}$, is a hyperplane passing through the origin of the $\mathbf{r}'(k)$ space.

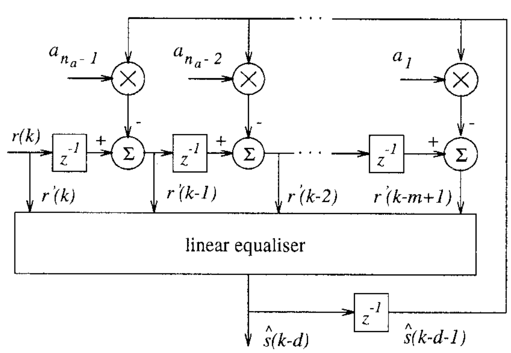

Figure 2: Schematic of translated DFE.

The translated vector $\mathbf{r}'(k)$ can be computed recursively according to

$$\left. \begin{array}{l} r'(k-i) = z^{-1} r'(k-i+1) - \\ \quad a_{n_a-i}\hat{s}(k-d-1),\ i = m-1, \cdots, 1 \\ r'(k) = r(k) \end{array} \right\} \qquad (11)$$

and an alternative implementation of the linear-combiner DFE is shown in Fig.2.

2.2 The minimum BER solution

The symbol vector $\mathbf{s}_f(k) = [s(k) \cdots s(k-d)]^T$ has $N_s = 2^{d+1} = 2^m$ combinations. Let these N_s sequences be $\mathbf{s}_{f,j}(k)$, $1 \leq j \leq N_s$. In the translated space, the noise-free channel states are given by

$$\mathbf{r}_j' = F''\mathbf{s}_{f,j}(k), \quad 1 \le j \le N_s \tag{12}$$

where the $m \times (d+1)$ matrix F'' has the form

$$F'' = \begin{bmatrix} a_0 & a_1 & \cdots & a_{n_a-1} \\ 0 & a_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_1 \\ 0 & \cdots & 0 & a_0 \end{bmatrix} \tag{13}$$

These states can be divided into two subsets, R^+ and R^-, corresponding to $s(k-d) = \pm 1$.

Let Z^+ and Z^- be the regions of $\mathbf{r}'(k)$ related to the decisions $\hat{s}(k-d) = \pm 1$ respectively. The BER of the linear equaliser (10) is given by

$$P_E(\mathbf{w}) = \sum_{\mathbf{r}_i' \in R^+} \beta_i \int_{\mathbf{r}' \in Z^-} p_{\mathbf{r}'}(\mathbf{r}'|\mathbf{r}_i')d\mathbf{r}' + \sum_{\mathbf{r}_j' \in R^-} \beta_j \int_{\mathbf{r}' \in Z^+} p_{\mathbf{r}'}(\mathbf{r}'|\mathbf{r}_j')d\mathbf{r}' \tag{14}$$

where $p_{\mathbf{r}'}(\mathbf{r}'(k)|\mathbf{r}_i')$ is the p.d.f. of $\mathbf{r}'(k)$ conditioned on the received channel state being $\mathbf{r}_i'$ and β_i is the a priori probability of $\mathbf{r}_i'$.

Since decision feedback usually makes R^+ and R^- linearly separable, (14) can be simplified to

$$P_E(\mathbf{w}) = \frac{2}{N_s} \sum_{i=1}^{N_s/2} Q\left(\frac{\rho_i}{\sigma_e}\right) \tag{15}$$

where

$$Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx \tag{16}$$

and

$$\rho_i = \frac{|(\mathbf{r}_i')^T \mathbf{w}|}{\|\mathbf{w}\|}, \quad \mathbf{r}_i' \in R^+ \tag{17}$$

The derivation of (15) can be found in [9]. It is obvious that the minimum MSE solution does not necessarily minimize $P_E(\mathbf{w})$. The minimum-BER linear-combiner DFE should minimize the BER expression (15). In practice, the optimal solution $\mathbf{w}_{opt}$ can be obtained using the gradient algorithm

$$\mathbf{w}(t+1) = \mathbf{w}(t) - \frac{\eta}{\left\|\frac{\partial P_E(\mathbf{w}(t))}{\partial \mathbf{w}}\right\|} \frac{\partial P_E(\mathbf{w}(t))}{\partial \mathbf{w}} \tag{18}$$

where η is an adaptive gain. The initial value $\mathbf{w}(0)$ can be chosen to be the Wiener solution. Notice that the algorithm (18) is an iterative procedure and does not involve any channel observation $\mathbf{r}'(k)$. The optimisation procedure is also independent of feedback states.

3 Examples

Two examples are used to compare the minimum-BER and minimum-MSE solutions of the linear-combiner DFE. The first example is a two-tap channel with the transfer function

$$A_1(z) = 0.5 + 1.0z^{-1} \tag{19}$$

The equaliser structure is chosen as $d = 1$, $m = 2$ and $n = 1$. Without feedback, there are 8 channel states listed in the Table 1.

No	$s(k)$	$s(k-1)$	$s(k-2)$	$\hat{r}(k)$	$\hat{r}(k-1)$
1	-1	-1	-1	-1.5	-1.5
2	1	-1	-1	-0.5	-1.5
3	-1	1	-1	0.5	-0.5
4	1	1	-1	1.5	-0.5
5	-1	-1	1	-1.5	0.5
6	1	-1	1	-0.5	0.5
7	-1	1	1	0.5	1.5
8	1	1	1	1.5	1.5

Table 1: Channel states for $A_1(z) = 0.5 + 1.0z^{-1}$.

The effect of decision feedback or space translation (11) is to translate two sets of the channel states related to the two values of $\hat{s}(k-d-1)$ into a single set as illustrated in Fig.3.

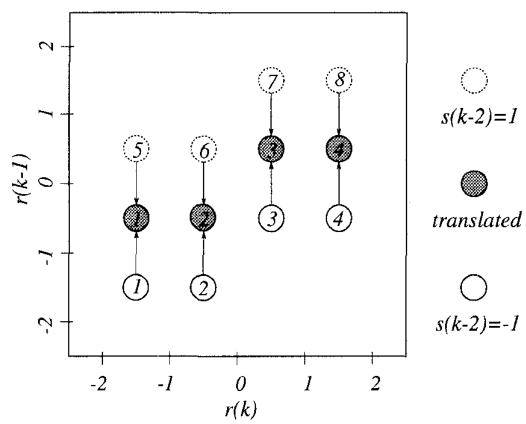

Figure 3: Illustration of effect of feedback.

The decision boundaries of the Bayesian DFE, the minimum-BER and minimum-MSE linear-combiner DFEs, plotted in the translated observation space, are shown in Fig.4. The decision boundary of the minimum-BER linear-combiner DFE is depicted in Fig.4 under the title "best linear approximation" to emphasize the fact that it is the best linear approximation to the optimal nonlinear Bayesian boundary. This example also demonstrates that the Wiener solution does not achieve the full performance potential of the linear-combiner DFE structure. Fig.5

compares the BERs as a function of SNR with detected symbols being fed back for these three DFEs.

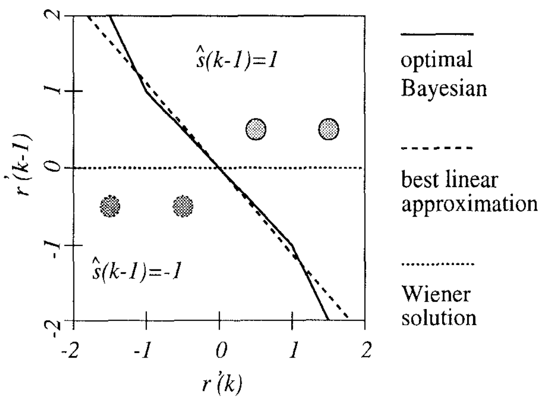

Figure 4: Asymptotic decision boundaries corresponding to large SNR for $A_1(z) = 0.5 + 1.0z^{-1}$.

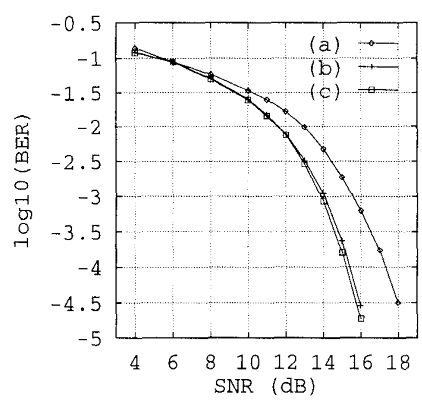

Figure 5: Performance comparison for $A_1(z) = 0.5 + 1.0z^{-1}$ with detected symbols being fed back. (a): minimum-MSE linear-combiner DFE, (b): minimum-BER linear-combiner DFE, (c): Bayesian DFE.

The second example is a 5-tap channel with the transfer function given by

$$A_2(z) = 0.227 + 0.466z^{-1} + 0.688z^{-2} + 0.466z^{-3} + 0.277z^{-4} \tag{20}$$

The structure of the DFE is chosen to be $d = 4$, $m = 5$ and $n = 4$. The BERs of the Bayesian DFE, the minimum-BER and minimum-MSE linear-combiner DFEs with detected symbols being fed back are plotted in Fig.6, where it can be seen that the performance of the minimum-BER linear-combiner DFE is significantly better than that of the Wiener solution. The performance gap between the Bayesian DFE and the minimum-BER linear-combiner DFE confirms the fact that the real optimal solution for the DFE structure is generally nonlinear. The "best linear solution" is suboptimal in nature. However the usual Wiener solution is inferior to this "best linear solution".

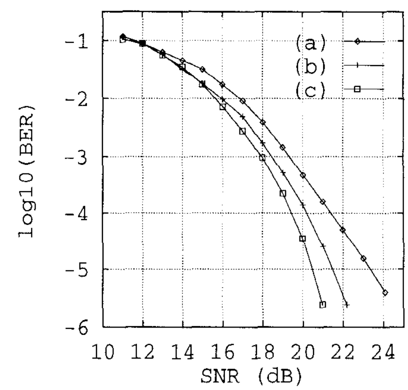

Figure 6: Performance comparison for $A_2(z) = 0.227 + 0.466z^{-1} + 0.688z^{-2} + 0.466z^{-3} + 0.227z^{-4}$ with detected symbols being fed back. (a): minimum-MSE linear-combiner DFE, (b): minimum-BER linear-combiner DFE, (c): Bayesian DFE.

4 Adaptive implementation

The algorithm (18) is an off-line optimisation procedure suitable for application to stationary channels. The following adaptive procedure is suggested for on-line implementation of the minimum-BER linear-combiner DFE. At the sample k, the procedure consists of:

(i) Using an adaptive algorithm such as the least mean square (LMS) algorithm to update the channel estimate and a noise variance estimator to estimate $\sigma_e^2(k)$;

(ii) Computing the subset of the channel states $R^+(k)$ and the gradient $\frac{\partial P_E(\mathbf{w}(k-1))}{\partial \mathbf{w}}$;

(iii) Using the feedback $\hat{s}(k-d-1)$ and the channel estimate to obtain the translated observation vector $\mathbf{r}'(k)$;

(iv) Updating the equaliser's weights using

$$\mathbf{w}(k) = \mathbf{w}(k-1) - \frac{\eta}{\left\| \frac{\partial P_E(\mathbf{w}(k-1))}{\partial \mathbf{w}} \right\|} \frac{\partial P_E(\mathbf{w}(k-1))}{\partial \mathbf{w}} \tag{21}$$

(v) Making the decision $\hat{s}(k-d)$ based on the filter output $\mathbf{w}^T(k)\mathbf{r}'(k)$.

Computational complexity of this adaptive linear-combiner DFE is considerably more than that of the standard adaptive linear-combiner DFE. However the performance gain justifies the increase in computation. Some of the channel states $\mathbf{r}_i' \in R^+$ are far away from the decision boundary and contribute little to the performance criterion (15). Computational requirements of the minimum-BER linear-combiner DFE can be reduced by neglecting these states from the optimisation procedure with little performance degradation. For example, consider the case of Fig.4. By just using the single state at (0.5, 0.5) in the optimisation, little performance degradation will occur, compared with using the full set R^+ of the two states.

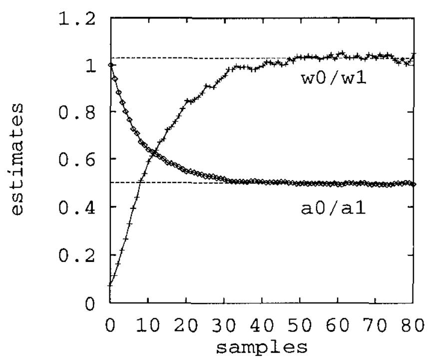

Figure 7: Trajectories of channel estimates and equaliser weights. Two lines indicate the respective optimal values.

The convergence behaviour of this adaptive procedure is tested using the following example. Initially, the channel has a transfer function $A_3(z) = 0.8 + 0.8z^{-1}$ with a SNR=15 dB. At the sample $k = 0$, the channel jumps to the transfer function $A_1(z) = 0.5+1.0z^{-1}$. The LMS algorithm is used to estimate the channel taps with an adaptive gain 0.1 and (21) is used to update the equaliser weights with $\eta = 0.1$. The trajectories of the channel estimates $a_0(k)/a_1(k)$ and the equaliser weights $w_0(k)/w_1(k)$, averaged over 50 different runs, are plotted in Fig.7. It can be seen from Fig.7 that the convergence speed of this adaptive procedure is reasonable.

5 Conclusions

The geometric translation property of the decision feedback in the linear-combiner DFE structure has been established in this paper. Basically, the decision feedback performs a space translation that maps the DFE onto an equivalent and simpler linear equaliser without feedback in the translated observation space. It has been shown that the Wiener solution can be far from the best possible performance of the linear-combiner DFE structure. Based on a BER expression, a novel minimum-BER linear-combiner DFE has been derived, which achieves the full performance potential of the linear-combiner DFE structure and offers the best linear approximation to the Bayesian solution.

6 References

[1] S.U.H. Qureshi, "Adaptive equalization," *Proc. IEEE*, Vol.73, No.9, pp.1349–1387, 1985.

[2] S. Chen, B. Mulgrew and S. McLaughlin, "Adaptive Bayesian equaliser with decision feedback," *IEEE Trans. Signal Processing*, Vol.41, No.9, pp.2918–2927, 1993.

[3] S. Siu, G.J. Gibson and C.F.N. Cowan, "Decision feedback equalisation using neural network structures and performance comparison with the standard architecture," *IEE Proc., Pt.I*, Vol.137, No.4, pp.221–225, 1990.

[4] D. Williamson, R.A. Kennedy and G.W. Pulford, "Block decision feedback equalization," *IEEE Trans. Communications*, Vol.40, No.2, pp.255-264, 1992.

[5] S. Chen, S. McLaughlin and B. Mulgrew, "Complex-valued radial basis function network, Part II: application to digital communications channel equalisation," *EURASIP Signal Processing J.*, Vol.36, pp.175–188, 1994.

[6] S. Chen, "Adaptive equalisation using neural networks," in *Applications of Neural Networks*, A.F. Murray (ed.), Kluwer, 1995, pp.241–265.

[7] I. Cha and S.A. Kassam, "Channel equalization using adaptive complex radial basis function networks," *IEEE J. Selected Areas in Communications*, Vol.13, No.1, pp.122–131, 1995.

[8] A.P. Clark, L.H. Lee and R.S. Marshall, "Developments of the conventional nonlinear equaliser," *IEE Proc., Pt.F*, Vol.129, No.2, pp.85–94, 1982.

[9] S. Chen, E.S. Chng, B. Mulgrew and G.J. Gibson, "On decision feedback equaliser," submitted to *IEEE Trans. Communications*, 1995.