# ADAPTIVE BAYESIAN DECISION FEEDBACK EQUALISER INCORPORATING CO-CHANNEL INTERFERENCE COMPENSATION

*S. Chen, S. McLaughlin, B. Mulgrew and P.M. Grant*

Department of Electrical Engineering, University of Edinburgh,
King's Buildings, Edinburgh EH9 3JL, UK

**Abstract**

The paper derives a Bayesian decision feedback equaliser (DFE) which incorporates co-channel interference (CCI) compensation. By exploiting the structure of CCI signals, the proposed Bayesian DFE can distinguish an interfering signal from white noise and uses this information to improve performance. Adaptive implementation of this Bayesian DFE includes first using the standard least mean square (LMS) algorithm to identify the channel model and then estimating the co-channel states by means of a simple unsupervised clustering algorithm. Simulation involving a binary signal constellation is used to compare both the theoretical and adaptive performance of this Bayesian DFE with those of the maximum likelihood sequence estimator (MLSE). The results obtained indicate that, by compensating CCI, the Bayesian DFE can outperform the MLSE for the CCI.

## 1. Introduction

Many communication systems, such as cellular radio and dual-polarised microwave radio channels, are impaired not only by channel intersymbol interference (ISI) but also by CCI. It is well-known that an adaptive equaliser can usually do better against CCI than it can against the same level of noise [1]. However, in doing so, most of the equalisers can only treat CCI as an additional noise source and do not fully exploit the differences between the interfering signals and the noise. For example, a linear equaliser only exploits the spectral characteristics of the interference through its autocorrelations [1],[2]. The MLSE [3], although very powerful in combating ISI with white noise, is less effective in dealing with CCI. This is because it is very difficult to write down the likelihood function which can explicitly distinguish the interfering signals from the noise. Therefore, the best that a standard MLSE can do is to treat these interfering signals as an additional coloured noise.

The probability density function (p.d.f.) of an interfering signal is quite different from that of the noise. An ideal equaliser should therefore be capable of distinguishing the interfering signal from the noise. In a previous study [2], a Bayesian transversal equaliser was derived which can effectively exploit the differences between the CCI and the noise and uses this information to improve performance. The present study extends this result to the decision feedback equaliser structure and develops a Bayesian DFE which incorporating CCI compensation. It is shown that, in the presence of severe CCI, this Bayesian DFE has superior performance over the MLSE. Adaptive implementation of this Bayesian DFE is also considered. In order to effectively compensate for the CCI, the set of co-channel states are required. These co-channel states can be estimated using a simple unsupervised clustering algorithm.

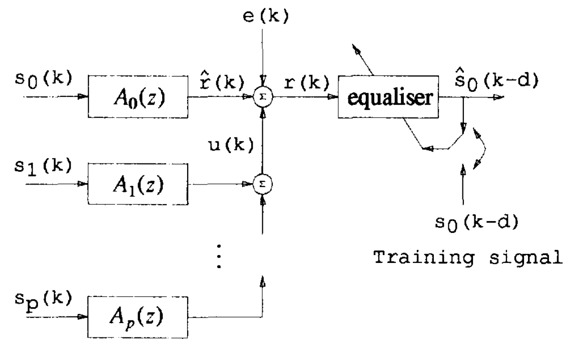


**Fig.1** Discrete-time model of communication system.

The system model considered in this study is depicted in Fig.1. This model is widely used (e.g. [4]). The channel $A_0(z)$ and the $p$ interfering co-channels $A_i(z)$, $1 \le i \le p$, are modelled by finite impulse response filters

$$A_i(z) = \sum_{j=0}^{n_i-1} a_{i,j} z^{-j}, \quad 0 \le i \le p. \tag{1}$$

The transmitted data $s_0(k)$ and the interfering data $s_i(k)$, $1 \le i \le p$, are independently identically distributed (i.i.d.) and they are mutually independent. The three components of the channel observation $r(k) = \hat{r}(k) + u(k) + e(k)$ will be called the desired signal, the interfering signal and the noise respectively. $e(k)$ is assumed to be Gaussian white noise with variance $\mathrm{E}[e^2(k)] = \sigma_e^2$. Let $\mathrm{E}[\hat{r}^2(k)] = \sigma_r^2$ and $\mathrm{E}[u^2(k)] = \sigma_u^2$. We define the signal to noise ratio (SNR) as $\mathrm{SNR} = \sigma_r^2/\sigma_e^2$, the signal to interference ratio (SIR) as $\mathrm{SIR} = \sigma_r^2/\sigma_u^2$, and the signal to interference and noise ratio (SINR) as $\mathrm{SINR} = \sigma_r^2/(\sigma_u^2 + \sigma_e^2)$ respectively. For notational simplicity and to highlight the basic concepts, $s_i(k)$, $0 \le i \le p$, are assumed to be binary and to take values from the symbol set $\{s^{(1)} = +1, s^{(2)} = -1\}$. The weights $a_{i,j}$ of $A_i(z)$ are therefore real-valued. Application to the complex-valued $A_i(z)$ and the 4-QAM symbol constellation is straightforward. The technique is equally applicable to higher-order symbol constellations but the computational complex will increase dramatically.

This work was supported by the UK Science and Engineering Research Council under award GR/G/72380.
SMcL gratefully acknowledges financial support from the Royal Society.

## 2. Bayesian DFE in the presence of CCI

The structure of a generic DFE is depicted in Fig.2. The equalisation process defined in Fig.2 uses the information present in the observed channel output vector

$$\mathbf{r}(k) = [r(k) \cdots r(k-m+1)]^T \qquad (2)$$

and the past detected symbol vector

$$\hat{\mathbf{s}}_b(k) = [\hat{s}_0(k-d-1) \cdots \hat{s}_0(k-d-n)]^T \qquad (3)$$

to produce a delayed estimate of the transmitted symbol. The three important structural parameters of the equaliser are the decision delay $d$, the feedforward order $m$ and the feedback order $n$ respectively. The feedforward order is usually related to the decision delay by $m = d+1$ and the feedback order is given by $n = n_0 + m - d - 2 = n_0 - 1$. In practice, $d = n_0 - 1$ is often chosen to cover the entire channel dispersion.

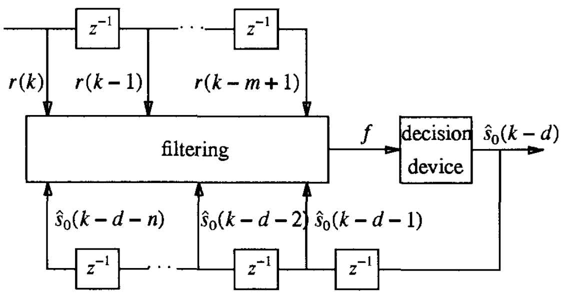


**Fig.2** Schematic of a generic decision feedback equaliser.

Given the channel $A_0(z)$, the value of the noiseless desired signal vector

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

is specified by the symbol sequence $\mathbf{s}(k) = [\mathbf{s}_f^T(k)\ \mathbf{s}_b^T(k)]^T$, where $\mathbf{s}_f(k) = [s_0(k) \cdots s_0(k-d)]^T$ and $\mathbf{s}_b(k) = [s_0(k-d-1) \cdots s_0(k-d-n)]^T$. Under the assumption that the given feedback vector is correct, that is, $\hat{\mathbf{s}}_b(k) = \mathbf{s}_b(k)$, the state of $\hat{\mathbf{r}}(k)$ is determined by $\mathbf{s}_f(k)$. For the binary constellation, $\mathbf{s}_f(k)$ has $N_r = 2^{d+1} = 2^m$ combinations and, therefore, the desired signal vector $\hat{\mathbf{r}}(k)$ has $N_r$ states. The states of $\hat{\mathbf{r}}(k)$ can be grouped into 2 sets according to the value of $s_0(k-d)$:

$$R_f^{(i)} = \{\hat{\mathbf{r}}(k) = \mathbf{r}_j^{(i)} | s_0(k-d) = s^{(i)}\},\ i = 1,2. \qquad (5)$$

Each $R_f^{(i)}$ contains $N_r^{(i)} = N_r/2 = 2^d$ states. Without loss of generality, we will assume that only one CCI ($p = 1$) is present. The interfering signal $u(k)$ has $N_{u,s} = 2^{n_1}$ scalar states $\{u_j,\ 1 \le j \le N_{u,s}\}$. Therefore, the interfering signal vector

$$\mathbf{u}(k) = [u(k) \cdots u(k-m+1)]^T \qquad (6)$$

has $N_u = 2^{m+n_1-1}$ states. The set of these co-channel vector states is denoted as $U = \{\mathbf{u}_j,\ 1 \le j \le N_u\}$.

The Bayesian DFE derived previously for combating ISI [5],[6] can now readily be extended to cover CCI. The p.d.f. of $\mathbf{r}(k)$ conditioned on $s_0(k-d) = s^{(i)}$ is

$$p_r(\mathbf{r}(k)|s_0(k-d) = s^{(i)}) = \sum_{j=1}^{N_r^{(i)}} \sum_{l=1}^{N_u} \alpha_{j,l}^{(i)} p_e(\mathbf{r}(k) - \mathbf{r}_j^{(i)} - \mathbf{u}_l),\quad i = 1,2, \qquad (7)$$

where $\mathbf{r}_j^{(i)} \in R_f^{(i)}$, $\mathbf{u}_l \in U$, $\alpha_{j,l}^{(i)}$ are the a-priori probabilities of $\mathbf{r}_j^{(i)} + \mathbf{u}_l$ and $p_e(\cdot)$ is the p.d.f. of the noise vector $\mathbf{e}(k) = [e(k) \cdots e(k-m+1)]^T$. Since all the $\mathbf{r}_j^{(i)} + \mathbf{u}_l$ can be assumed to be equiprobable and the noise p.d.f. is Gaussian, (7) leads to the 2 Bayesian decision variables

$$\eta_i(k) = \sum_{j=1}^{N_r^{(i)}} \sum_{l=1}^{N_u} \exp(-\| \mathbf{r}(k) - \mathbf{r}_j^{(i)} - \mathbf{u}_l \|^2 / 2\sigma_e^2),\ i = 1,2. \qquad (8)$$

The minimum-error-probability decision is defined by

$$\hat{s}_0(k-d) = \begin{cases} s^{(1)}, & \text{if } \eta_1(k) \ge \eta_2(k), \\ s^{(2)}, & \text{otherwise,} \end{cases} \qquad (9)$$

which provides the optimal solution for the equalisation structure of Fig.2.

The computational complexity of this Bayesian DFE with full CCI compensation obviously depends on the number of $\mathbf{r}_j^{(i)} + \mathbf{u}_l$, which is $N_r \times N_u$. To reduce the complexity, an approximation of this full Bayesian DFE can be adopted which only approximates co-channel states. The approximation can easily be achieved due to the symmetric structure of co-channel states, and this will be illustrated using an example. Another reason for adopting the approximation is due to practical considerations. The scalar co-channel states $u_l$ can only be estimated based on unsupervised learning. The resolution of unsupervised learning is limited, and it is not always possible to resolve all the co-channel states. In such a situation, it is natural to consider an approximation. Carrying out the approximation to an extreme and approximating the CCI as an additional noise, we obtained the Bayesian DFE with the decision variables

$$\eta_i(k) = \sum_{j=1}^{N_r^{(i)}} \exp(-\| \mathbf{r}(k) - \mathbf{r}_j^{(i)} \|^2 / 2\sigma^2),\ i = 1,2, \qquad (10)$$

where $\sigma^2 = \sigma_e^2 + \sigma_u^2$. This has the same form as the original Bayesian DFE [5],[6].

We now use an example to illustrate the above discussion and to compare the theoretical performance of the Bayesian DFE with that of the MLSE. The channel is given by

$$A_0(z) = 0.34 + 0.88z^{-1} + 0.34z^{-2}, \qquad (11)$$

and the interfering co-channel is

$$A_1(z) = \lambda(0.50 + 0.81z^{-1} + 0.31z^{-2}), \qquad (12)$$

where the value of the parameter $\lambda$ dictates the SIR requirement. For example, $\lambda = 0.32$ gives rise to a SIR=10 dB. The set of the scalar co-channel states is listed in Table 1.

| No. | $s_1(k)$ | $s_1(k-1)$ | $s_1(k-2)$ | $u_l$ |
|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 1.62 $\lambda$ |
| 2 | 1 | 1 | -1 | 1.00 $\lambda$ |
| 3 | 1 | -1 | 1 | 0.00 $\lambda$ |
| 4 | 1 | -1 | -1 | -0.62 $\lambda$ |
| 5 | -1 | 1 | 1 | 0.62 $\lambda$ |
| 6 | -1 | 1 | -1 | 0.00 $\lambda$ |
| 7 | -1 | -1 | 1 | -1.00 $\lambda$ |
| 8 | -1 | -1 | -1 | -1.62 $\lambda$ |

Table 1. Scalar co-channel states.

The symmetric structure of the co-channel states is apparent in Table 1. The set of the vector co-channel states, $U$, is obtained by expanding these scalar states. In this example, $U$ contains 32 vector states. Re-arrange the scalar states as

$$(1.62\lambda, 1.00\lambda, 0.62\lambda, 0.00\lambda,$$
$$-0.00\lambda, -0.62\lambda, -1.00\lambda, -1.62\lambda). \quad (13)$$

We may approximate $(1.62\lambda, 1.00\lambda)$ by its mean $1.31\lambda$ and $(0.62\lambda, 0.00\lambda)$ by $0.31\lambda$. This results in 4 approximated scalar co-channel states

$$(1.31\lambda, 0.31\lambda, -0.31\lambda, -1.31\lambda). \quad (14)$$

The number of the resulting approximated vector co-channel states is 16. This approximation may also be viewed from a different angle. The order of the co-channel is $n_1 = 3$. Suppose that we only have an approximated co-channel order $\hat{n}_1 = 2$. This will give us 4 scalar co-channel states, and each of these approximated states is the mean of a pair of two true states. The Bayesian DFE with decision variables described by (10) may then be viewed as the result of choosing $\hat{n}_1 = 0$.

For a SIR=10dB, Fig.3 plots the performance curves of the Bayesian DFE with no CCI compensation ($\hat{n}_1 = 0$), the MLSE, the Bayesian DFE with an approximated CCI compensation ($\hat{n}_1 = 2$) and the Bayesian DFE with full CCI compensation ($\hat{n}_1 = n_1 = 3$). For the range of SNRs from 0dB to 25dB, the corresponding SINRs are from -0.42dB to 9.76dB.

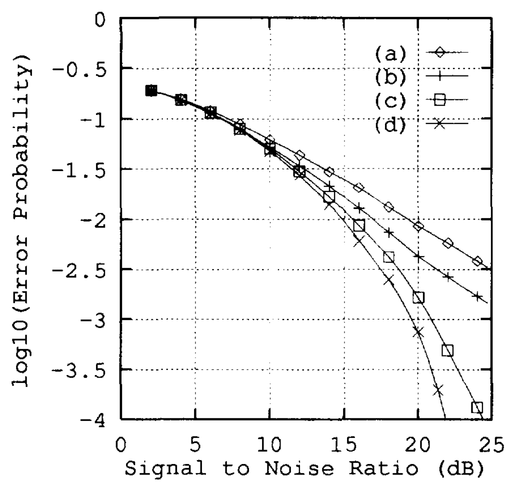


**Fig.3** Theoretical performance. (a) Bayesian DFE with no CCI compensation ($\hat{n}_1 = 0$); (b) MLSE; (c) Bayesian DFE with an approximated CCI compensation ($\hat{n}_1 = 2$); (d) Bayesian DFE with full CCI compensation ($\hat{n}_1 = n_1 = 3$).

### 3. Adaptive implementation

Since the equaliser has access to the transmitted data $\{s_0(k)\}$ during the training, the channel model can be identified using, for example, the LMS algorithm

$$\left.\begin{aligned} \varepsilon(k) &= r(k) - \sum_{j=0}^{n_0-1} \hat{a}_{0,j}(k-1)s_0(k-j), \\ \hat{a}_{0,j}(k) &= \hat{a}_{0,j}(k-1) + g_a\varepsilon(k)s_0(k-j), \ 0 \le j \le n_0 - 1, \end{aligned}\right\} \quad (15)$$

where $g_a$ is an adaptive gain. Given the channel estimate $\hat{\mathbf{a}}_0 = [\hat{a}_{0,0} \cdots \hat{a}_{0,n_0-1}]^T$, it is straightforward to calculate the channel states $\mathbf{r}_j^{(i)}$.

The equaliser does not have access to the interfering data $\{s_1(k)\}$, and the above supervised learning is not applicable for identifying the co-channel states. We propose the following unsupervised clustering algorithm for estimating the scalar co-channel states:

- o Compute the residual

$$\varepsilon(k) = r(k) - \sum_{j=0}^{n_0-1} \hat{a}_{0,j}s_0(k-j).$$

- o Compute the distances between $\varepsilon(k)$ and the scalar co-channel states $u_l(k-1)$, $1 \le l \le N_{u,s}$,

$$\rho_l(k) = (\varepsilon(k) - u_l(k-1))^2, \ 1 \le l \le N_{u,s}.$$

Find the minimum distance

$$\rho_{l^*}(k) = \min\{\rho_l(k), \ 1 \le l \le N_{u,s}\}.$$

- o Update the $l^*$-th co-channel state

$$u_{l^*}(k) = u_{l^*}(k-1) + g_u(\varepsilon(k) - u_{l^*}(k-1)),$$

where $g_u$ is an adaptive gain. Because of the symmetric structure of the co-channel states, $u_{N_{u,s}-l^*+1}(k)$ is set to $u_{N_{u,s}-l^*+1}(k) = -u_{l^*}(k)$.

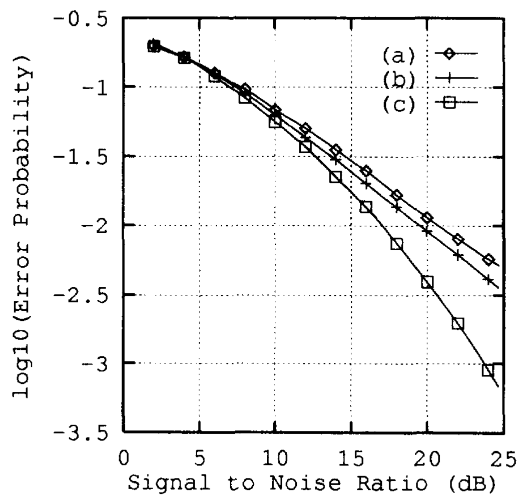


**Fig.4** Adaptive performance. (a) Bayesian DFE with no CCI compensation ($\hat{n}_1 = 0$); (b) MLSE; (c) Bayesian DFE with an approximated CCI compensation ($\hat{n}_1 = 2$).

For the example given in the previous section, 80 training samples were used to identify the channel model. The adaptive gain of the LMS algorithm was chosen to be $g_a = 0.03$ and the adaptive gain for the unsupervised clustering algorithm was $g_u = 0.01$. The performance curves of the adaptive Bayesian DFE with no CCI compensation ($\hat{n}_1 = 0$), the adaptive MLSE and the adaptive Bayesian DFE with an approximated CCI compensation ($\hat{n}_1 = 2$) are depicted in Fig.4.

A possible alternative for estimating co-channel states is to employ higher-order cumulants techniques. A co-channel model may be estimated blindly from the residuals $\{\varepsilon(k)\}$ using a higher-order cumulants algorithm [7]-[9]. Once a co-channel model is obtained, it is a straightforward task to calculate the co-channel states. This approach, however, requires futher investigation and will not be discussed in the present study.

## 4. Conclusions

Adaptive equalisation in the presence of ISI, additive Gaussian white noise and CCI has been investigated. It has been shown that, by exploiting the nature of interfering signals, the Bayesian DFE is capable of distinguishing an interfering signal from the noise. Simulation results have demonstrated that the Bayesian DFE which incorporates CCI compensation can outperform the MLSE in the presence of severe CCI. Adaptive implementation of this Bayesian DFE has been studied, and a simple unsupervised clustering algorithm has been suggested to learn the co-channel states.

## 5. References


[1] B.R. Petersen and D.D. Falconer, "Exploiting cyclostationary subscriber-loop interference by equalization," in *Proc. GLOBECOM'90* (San Diego), 1990, pp.1156-1160.

[2] S. Chen and B. Mulgrew, "Overcoming co-channel interference using an adaptive radial basis function equaliser," *EURASIP Signal Processing J.*, Vol.28, No.1, pp.91-107, 1992.

[3] G.D. Forney, "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Trans. Information Theory*, Vol.IT-18, No.3, pp.363-378, 1972.

[4] J.C. Campbell, A.J. Gibbs and B.M. Smith, "The cyclostationary nature of crosstalk interference from digital signals in multipair cable - Part I: fundamentals; Part II: applications and further results," *IEEE Trans. Communications*, Vol.COM-31, No.5, pp.629-649, 1983.

[5] S. Chen, B. Mulgrew and S. McLaughlin, "Adaptive Bayesian decision feedback equaliser based on a radial basis function network," in *Proc. ICC'92* (Chicago), 1992, Vol.3, pp.343.3.1-343.3.5.

[6] S. Chen, B. Mulgrew and S. McLaughlin, "Adaptive Bayesian equaliser with decision feedback," *IEEE Trans. Signal Processing*, to appear, September 1993.

[7] G.B. Giannakis and J.M. Mendel, "Identification of nonminimum phase system using higher order statistics," *IEEE Trans. Acoustics, Speech and Signal Processing*, Vol.ASSP-37, pp.360-377, 1989.

[8] H.-H. Chiang and C.L. Nikias, "Adaptive deconvolution and identification of nonminimum phase FIR systems based on cumulants," *IEEE Trans. Automatic Control*, Vol.AC-35, pp.36-47, 1990.

[9] F.-C. Zheng, S. McLaughlin and B. Mulgrew, "Cumulant-based deconvolution and identification: several new families of linear equations," *EURASIP Signal Processing J.*, Vol.30, No.2, pp.199-219, 1993.