Multiple Hyperplane Detector for Implementing the Asymptotic Bayesian Decision Feedback Equalizer

S. Chen†, L. Hanzo† and B. Mulgrew‡

† Department of Electronics and Computer Science
University of Southampton, Southampton SO17 1BJ, U.K.

‡ Department of Electronics and Electrical Engineering
University of Edinburgh, Edinburgh EH9 3JL, U.K.

Abstract— **A detector based on multiple-hyperplane partitioning of the signal space is derived for realizing the optimal Bayesian decision feedback equaliser (DFE). It is known that the optimal Bayesian decision boundary separating any two neighbouring signal classes is asymptotically piecewise linear and consists of several hyperplanes, when the signal to noise ratio (SNR) tends to infinity. The proposed technique determines these hyperplanes and uses them to partition the observation space. The resulting detector can closely approximate the optimal Bayesian detector, at an advantage of considerably reduced decision complexity.**

I. Introduction

For the class of DFEs that employ a symbol-decision finite-memory structure with a fixed decision delay, the optimal solution is the Bayesian detector [1]–[3]. The complexity of the optimal Bayesian DFE is determined by the factor of M^{n_a}, where M being the size of the symbol constellation and n_a the channel impulse response (CIR) length. As the complexity of this optimal detector increases exponentially with the size of symbol set M, the conventional or linear-combiner DFE [4]–[6] is often used in practice to provide a trade-off between performance and detector complexity.

For the 2-PAM case, the performance difference between the conventional and optimal Bayesian DFEs has a geometric explanation: a linear-combiner DFE can only partition the observation space with a hyperplane while the Bayesian detector can do so with a hypersurface [2]. Asymptotically, as the SNR tends to infinity, the Bayesian hypersurface becomes piecewise linear and is made up of a set of hyperplanes [7]. In practice, at large rather than infinite SNR, the Bayesian decision hypersurface can closely be approximated by a multiple-hyperplane form. This motivated our previous research on multiple-hyperplane detector [8].

Signal space partitioning techniques for binary channels have been developed from different motivations. Kim and Moon [9],[10] developed a novel partitioning design. Their technique determines a set of hyperplanes which separate clusters of noiseless channel states. The convex regions associated with individual states are constructed by intersecting hyperplanes. The overall decision region is then formed from these convex regions. The decision complexity and performance of the detector is controlled during design by a specified minimum separating distance. The main drawback of their design is that it involves extensive computational effort during the design process. In our previous work [8], we have proposed a much simpler alternative design to explicitly realize the asymptotic Bayesian decision boundary.

This paper extends this multiple-hyperplane detector design to M-PAM channels. Based on a geometric translation property for the M sets of noiseless channel states, the asymptotic Bayesian boundary for separating any two neighbouring signal classes can be deduced, and this allows us to extend the binary case design [8] to the general M-PAM case. Similar to the binary case, the design of our multiple-hyperplane detector for M-PAM channels is straightforward, and guarantees to realize the asymptotic Bayesian DFE detector. Furthermore, the reduction in detector complexity with signal space partitioning approach is more significant for $M > 2$.

II. The problem formulation

We will assume that the real-valued channel generates the received signal samples of:

$$y(k) = \sum_{i=0}^{n_a-1} a_i s(k-i) + e(k), \qquad (1)$$

where a_i are the CIR taps, the Gaussian white noise $\{e(k)\}$ has zero mean and variance σ_e^2, and the M-PAM symbol $s(k)$ takes the values from the set: $\mathcal{S} \triangleq \{s_i = 2i - M - 1, 1 \le i \le M\}$. The SNR is defined as $\left(\sum_{i=0}^{n_a-1} a_i^2\right) \sigma_s^2/\sigma_e^2$, where σ_s^2 is the symbol variance. The DFE uses the information present in the noisy observation vector $\mathbf{y}(k) = [y(k)\, y(k-1) \cdots y(k-m+1)]^T$ and the past detected symbol vector $\hat{\mathbf{s}}_b(k) = [\hat{s}(k-d-1) \cdots \hat{s}(k-d-n)]^T$ to produce an estimate $\hat{s}(k-d)$ of $s(k-d)$, where d, m and n are the decision delay, the feedforward and feedback orders, respectively. We will choose $d = n_a - 1$, $m = n_a$ and $n = n_a - 1$, as this

0-7803-7097-1/01/$10.00 ©2001 IEEE

choice is sufficient to guarantee a desired linear separability for different signal classes [5].

The observation vector $\mathbf{y}(k)$ can be expressed as [5]:

$$\mathbf{y}(k) = F_1\, \mathbf{s}_f(k) + F_2\, \mathbf{s}_b(k) + \mathbf{e}(k)\,, \qquad (2)$$

where $\mathbf{s}_f(k) = [s(k) \cdots s(k-d)]^T$, $\mathbf{s}_b(k) = [s(k-d-1) \cdots s(k-d-n)]^T$, $\mathbf{e}(k) = [e(k) \cdots e(k-m+1)]^T$, and

$$F_1 = \begin{bmatrix} a_0 & a_1 & \cdots & a_{n_a-1} \\ 0 & a_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_1 \\ 0 & \cdots & 0 & a_0 \end{bmatrix} \qquad (3)$$

$$F_2 = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ a_{n_a-1} & 0 & \ddots & \vdots \\ a_{n_a-2} & a_{n_a-1} & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 \\ a_1 & \cdots & a_{n_a-2} & a_{n_a-1} \end{bmatrix} \qquad (4)$$

are the $m \times (d+1)$ and $m \times n$ CIR matrices, respectively. Assuming correct past decisions, we have

$$\mathbf{y}(k) = F_1\, \mathbf{s}_f(k) + F_2\, \hat{\mathbf{s}}_b(k) + \mathbf{e}(k)\,. \qquad (5)$$

Thus the decision feedback translates the original space $\mathbf{y}(k)$ into a new space $\mathbf{r}(k)$:

$$\mathbf{r}(k) \triangleq \mathbf{y}(k) - F_2\, \hat{\mathbf{s}}_b(k)\,. \qquad (6)$$

Let the $N_f = M^{d+1}$ possible sequences of $\mathbf{s}_f(k)$ be $\mathbf{s}_{fj}$, $1 \le j \le N_f$. The set of the noiseless channel states in the translated signal space, namely,

$$R \triangleq \{\mathbf{r}_j = F_1\, \mathbf{s}_{fj},\ 1 \le j \le N_f\} \qquad (7)$$

can be partitioned into M conditional subsets:

$$R^{(i)} \triangleq \{\mathbf{r}_j \in R:\ s(k-d) = s_i\},\ 1 \le i \le M\,. \qquad (8)$$

The optimal Bayesian DFE [3] can now be summarized. The M decision variables are given by

$$\rho_i(\mathbf{r}(k)) = \sum_{\mathbf{r}_j \in R^{(i)}} e^{-\frac{\|\mathbf{r}(k)-\mathbf{r}_j\|^2}{2\sigma_e^2}}\,,\ 1 \le i \le M, \qquad (9)$$

and the minimum-error-rate decision is defined by

$$\hat{s}(k-d) = s_{i^*} \ \text{ with } \ i^* = \arg \max_{1 \le i \le M} \{\rho_i(\mathbf{r}(k))\} \qquad (10)$$

Table I gives the complexity of this optimal detector.

TABLE I

COMPARISON OF DECISION COMPLEXITY FOR THE FULL BAYESIAN AND MULTIPLE-HYPERPLANE DETECTORS.

	Full Bayesian	Multiple-hyperplane
Additions	$2n_a M^{n_a} - M$	$(n_a + M - 2)L$
Multiplications	$(n_a+1)M^{n_a}$	$n_a L$
e^x	M^{n_a}	—

III. MULTIPLE-HYPERPLANE DETECTOR

We first establish a geometric translation property for any two neighbouring subsets of channel states.

Lemma 1: For $1 \le i \le M-1$, the subset $R^{(i+1)}$ is a translation of $R^{(i)}$ by the amount $2\mathbf{a}_{\text{rev}}$

$$R^{(i+1)} = R^{(i)} + 2\mathbf{a}_{\text{rev}}\,, \qquad (11)$$

where $\mathbf{a}_{\text{rev}} = [a_{n_a-1} \cdots a_1\ a_0]^T$. Furthermore, $R^{(i)}$ and $R^{(i+1)}$ are linearly separable.

Proof: From the definitions of $R^{(i)}$ and F_1, for any $\mathbf{r}_l \in R^{(i)}$, there exists a $\mathbf{r}_j \in R^{(i+1)}$ such that $\mathbf{r}_j = \mathbf{r}_l + (s_{i+1} - s_i)\mathbf{a}_{\text{rev}} = \mathbf{r}_l + 2\mathbf{a}_{\text{rev}}$, which implies (11). To prove the linear separability, consider the hyperplane

$$H(\mathbf{r} + \mathbf{c}_i) \triangleq \hat{\mathbf{w}}^T \left(\mathbf{r} + 2\left(\frac{M}{2} - i\right)\mathbf{a}_{\text{rev}}\right) = 0 \qquad (12)$$

with $\hat{\mathbf{w}} = \left[0\,0 \cdots 0\ \frac{1}{a_0}\right]^T$. For any $\mathbf{r}_l \in R^{(i)}$ and any $\mathbf{r}_j \in R^{(i+1)}$, we have $H(\mathbf{r}_l + \mathbf{c}_i) = -1 < 0$ and $H(\mathbf{r}_j + \mathbf{c}_i) = 1 > 0$. Fig. 1 illustrates this lemma graphically.

A. Asymptotic optimal boundary for two neighbouring classes

Although it is always possible to construct a hyperplane to correctly separate $R^{(i)}$ from $R^{(i+1)}$, the optimal decision boundary $\mathcal{D}_i$ that separates $R^{(i)}$ from $R^{(i+1)}$ cannot generally be approximated by a single hyperplane. Without the loss of generality, consider $i = \frac{M}{2}$, the optimal decision boundary $\mathcal{D}_{\frac{M}{2}}$ for separating $R^{(\frac{M}{2})}$ and $R^{(\frac{M}{2}+1)}$. Because of lemma 1, when SNR$\rightarrow \infty$ (or $\sigma_e^2 \rightarrow 0$), the influence

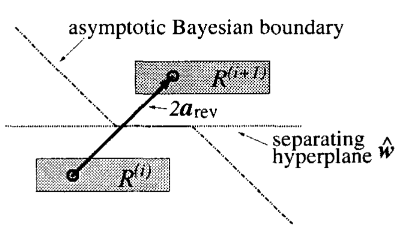

Fig. 1. Illustration of shift property.

from all the other $R^{(i)}$ for $i \neq \frac{M}{2}$ and $i \neq \frac{M}{2}+1$ vanishes much more quickly, and it effectively becomes a two-class problem. We have the following definition [8].

Definition 1: A pair of opposite-class channel states $(\mathbf{r}^{(+)} \in R^{(\frac{M}{2}+1)}, \mathbf{r}^{(-)} \in R^{(\frac{M}{2})})$ is said to be *dominant* if $\forall \mathbf{r}_j \in R^{(\frac{M}{2})} \bigcup R^{(\frac{M}{2}+1)}$, $\mathbf{r}_j \neq \mathbf{r}^{(+)}$ and $\mathbf{r}_j \neq \mathbf{r}^{(-)}$:

$$||\mathbf{r}_j - \mathbf{r}_0||^2 > ||\mathbf{r}^{(+)} - \mathbf{r}_0||^2, \qquad (13)$$

where $\bigcup$ denotes the union operator and

$$\mathbf{r}_0 = \frac{\mathbf{r}^{(+)} + \mathbf{r}^{(-)}}{2}. \qquad (14)$$

The following properties of $\mathcal{D}_{\frac{M}{2}}$ are useful in the derivation of a multiple-hyperplane detector (see [7]). A necessary condition for a point $\mathbf{r}_B \in \mathcal{D}_{\frac{M}{2}}$ is

$$\mathbf{r}_B = \frac{\mathbf{r}^{(+)} + \mathbf{r}^{(-)}}{2} + \left[\frac{\mathbf{r}^{(+)} - \mathbf{r}^{(-)}}{2}\right]^{\perp}, \qquad (15)$$

where $\mathbf{x}^{\perp}$ denotes an arbitrary vector in the subspace orthogonal to $\mathbf{x}$, $\mathbf{r}^{(+)}$ and $\mathbf{r}^{(-)}$ are a pair of dominant states; and the sufficient conditions for $\mathbf{r}_B \in \mathcal{D}_{\frac{M}{2}}$ are

$$||\mathbf{r}_B - \mathbf{r}^{(+)}||^2 < ||\mathbf{r}_B - \mathbf{r}_l||^2,\ \forall \mathbf{r}_l \in R^{(\frac{M}{2}+1)},\ \mathbf{r}_l \neq \mathbf{r}^{(+)}, \qquad (16)$$

$$||\mathbf{r}_B - \mathbf{r}^{(-)}||^2 < ||\mathbf{r}_B - \mathbf{r}_j||^2,\ \forall \mathbf{r}_j \in R^{(\frac{M}{2})},\ \mathbf{r}_j \neq \mathbf{r}^{(-)}, \qquad (17)$$

$$||\mathbf{r}_B - \mathbf{r}^{(+)}||^2 = ||\mathbf{r}_B - \mathbf{r}^{(-)}||^2. \qquad (18)$$

The following lemma describing $\mathcal{D}_{\frac{M}{2}}$ in the asymptotic case of $\sigma_e^2 \rightarrow 0$ is a direct consequence of the necessary and sufficient conditions (15)-(18).

Lemma 2: Asymptotically, the optimal decision boundary $\mathcal{D}_{\frac{M}{2}}$ separating $R^{(\frac{M}{2})}$ and $R^{(\frac{M}{2}+1)}$ is piecewise linear and made up of a set of L hyperplanes. Each of these hyperplanes is defined by a pair of dominant states, the hyperplane is orthogonal to the line connecting the pair of dominant states and passes through the midpoint of the line.

B. Multiple-hyperplane detector for two neighbouring classes

According to lemma 2, a multiple-hyperplane detector can be constructed to partition the signal space into the two regions of $\hat{s}(k-d) \leq -1$ and $\hat{s}(k-d) \geq 1$, respectively. The detector will consist of L linear discriminant functions and a many-to-one Boolean mapper, similar to the binary case given in [8]. For completeness, the design procedure for this multiple-hyperplane detector is produced here with the necessary modifications:

Step 1 Select all the L pairs of dominant channel states from the two subsets $R^{(\frac{M}{2})}$ and $R^{(\frac{M}{2}+1)}$. For each pair, compute a hyperplane that separates these two opposite-class states.
Step 2 A Boolean logic function is obtained to make a decision based on the location of the observation vector $\mathbf{r}(k)$ relative to each hyperplane. This is achieved by first defining a convex region associated with each state in a given class, e.g. the class $R^{(\frac{M}{2}+1)}$, and then forming a union of these regions.

From (15)–(18), it is seen that pairs of dominant states which define the asymptotic boundary can be selected using:

```
L = 0;
FOR r_q^(+) ∈ R^(M/2+1)
  FOR r_j^(-) ∈ R^(M/2)
    x = (r_q^(+) + r_j^(-))/2; η = ||r_q^(+) − x||^2;
    IF (||r_l^(+) − x||^2 > η, ∀r_l^(+) ∈ R^(M/2+1), l ≠ q) AND
       (||r_l^(-) − x||^2 > η, ∀r_l^(-) ∈ R^(M/2), l ≠ j)
         L += 1;
         R_Asym ← (r_L^(+), r_L^(-)) ≜ (r_q^(+), r_j^(-));
    END IF
  NEXT r_j^(-)
NEXT r_q^(+)
```

Each pair $(\mathbf{r}_l^{(+)}, \mathbf{r}_l^{(-)}) \in R_{\text{Asym}}$ determines a hyperplane

$$H_l(\mathbf{r}) = \mathbf{w}_l^T \mathbf{r} + b_l = 0 \qquad (19)$$

that is a part of the asymptotic optimal decision boundary. The weight vector $\mathbf{w}_l$ and bias b_l of the hyperplane can be computed straightforwardly as:

$$\mathbf{w}_l = \frac{2\left(\mathbf{r}_l^{(+)} - \mathbf{r}_l^{(-)}\right)}{||\mathbf{r}_l^{(+)} - \mathbf{r}_l^{(-)}||^2} \qquad (20)$$

and

$$b_l = -\frac{(\mathbf{r}_l^{(+)} - \mathbf{r}_l^{(-)})^T(\mathbf{r}_l^{(+)} + \mathbf{r}_l^{(-)})}{||\mathbf{r}_l^{(+)} - \mathbf{r}_l^{(-)}||^2}. \qquad (21)$$

The hyperplane defined by (20) and (21) is a *canonical* hyperplane with $(\mathbf{r}_l^{(+)}, \mathbf{r}_l^{(-)})$ as its two support vectors [6], and has the property that $H_l(\mathbf{r}_l^{(+)}) = 1$ and $H_l(\mathbf{r}_l^{(-)}) = -1$. The following definition is useful in the optimal multiple-hyperplane partitioning:

Definition 2: A state $\mathbf{r}_j \in R^{(\frac{M}{2})} \bigcup R^{(\frac{M}{2}+1)}$ is said to be *sufficiently separable* by H_l if H_l can separate $\mathbf{r}_j$ correctly with a "canonical distance" $|\mathbf{w}_l^T \mathbf{r}_j + b_l| \geq 1$.

Notice that $\mathbf{r}_j \in R^{(\frac{M}{2}+1)}$ is sufficiently separable by H_l iff $\mathbf{w}_l^T \mathbf{r}_j + b_l \geq 1$. Similarly, $\mathbf{r}_j \in R^{(\frac{M}{2})}$ is sufficiently separable by H_l iff $\mathbf{w}_l^T \mathbf{r}_j + b_l \leq -1$. Number the states in $R^{(\frac{M}{2})}$ as $\mathbf{r}_1^{(-)}$ to $\mathbf{r}_{N_s}^{(-)}$ and those in $R^{(\frac{M}{2}+1)}$ as $\mathbf{r}_1^{(+)}$ to $\mathbf{r}_{N_s}^{(+)}$, where $N_s = N_f/M$. All the states in $R^{(\frac{M}{2})} \bigcup R^{(\frac{M}{2}+1)}$ are tested to see if they can be separated sufficiently by H_l, and this generates the following "separability" matrix:

	$\mathbf{r}_1^{(-)}$	$\mathbf{r}_2^{(-)}$	$\cdots$	$\mathbf{r}_{N_s}^{(-)}$	$\mathbf{r}_1^{(+)}$	$\cdots$	$\mathbf{r}_{N_s}^{(+)}$
H_1	$h_{1,1}$	$h_{1,2}$	$\cdots$	h_{1,N_s}	h_{1,N_s+1}	$\cdots$	$h_{1,2N_s}$
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$
H_L	$h_{L,1}$	$h_{L,2}$	$\cdots$	h_{L,N_s}	h_{L,N_s+1}	$\cdots$	$h_{L,2N_s}$

where $h_{l,j} \in \{0,1\}$. The rule in generating this matrix is: if a state can *sufficiently* be separated by H_l, the corresponding binary index $h_{l,j} = 1$; otherwise $h_{l,j} = 0$.

Define the half-space $\mathcal{H}_l^{(+)} \triangleq \{\mathbf{r} : H_l(\mathbf{r}) \geq 0\}$. To construct a convex region $\mathcal{R}_q^{(+)}$ covering a state $\mathbf{r}_q^{(+)} \in R^{(\frac{M}{2}+1)}$, select those hyperplanes which can sufficiently separate $\mathbf{r}_q^{(+)}$ and denote $G_q^{(+)} \triangleq \{l : h_{l,q+N_s} = 1\}$. Then $\mathcal{R}_q^{(+)}$ is obtained by the intersection of all the $\mathcal{H}_j^{(+)}$ with $j \in G_q^{(+)}$

$$\mathcal{R}_q^{(+)} = \bigcap_{j \in G_q^{(+)}} \mathcal{H}_j^{(+)} . \tag{22}$$

In fact, a subset of the hyperplanes defined by $G_q^{(+)}$ is enough to construct $\mathcal{R}_q^{(+)}$, provided that every state in $R^{(\frac{M}{2})}$ can sufficiently be separated by at least one hyperplane in the subset. The overall decision region $\mathcal{R}^{(+)}$ associated with the decision $\hat{s}(k-d) \geq 1$ is simply

$$\mathcal{R}^{(+)} = \bigcup_{q=1}^{N_s} \mathcal{R}_q^{(+)} . \tag{23}$$

The Boolean logic function for the multiple-hyperplane detector is now completed. Define the threshold detector output $\beta_j(\mathbf{r}(k))$ for a linear discriminant function $H_j(\mathbf{r}(k))$:

$$\beta_j(\mathbf{r}(k)) \triangleq \begin{cases} 1, & H_j(\mathbf{r}(k)) \geq 0, \\ 0, & H_j(\mathbf{r}(k)) < 0. \end{cases} \tag{24}$$

A Boolean logic value $\theta_q(\mathbf{r}(k))$ indicating whether $\mathbf{r}(k) \in \mathcal{R}_q^{(+)}$ or not is obtained via a logic AND operation of $\{\beta_j(\mathbf{r}(k)) : j \in G_q^{(+)}\}$:

$$\theta_q(\mathbf{r}(k)) \triangleq \bigcap_{j \in G_q^{(+)}} \beta_j(\mathbf{r}(k)) . \tag{25}$$

A Boolean logic value $\alpha(\mathbf{r}(k))$ indicating whether $\mathbf{r}(k) \in \mathcal{R}^{(+)}$ (that is, $\hat{s}(k-d) \geq 1$) or not is obtained via a logic OR operation of $\{\theta_q(\mathbf{r}(k))\}$ for all q:

$$\alpha(\mathbf{r}(k)) \triangleq \bigcup_{q=1}^{N_s} \theta_q(\mathbf{r}(k)) . \tag{26}$$

C. Multiple-hyperplane detector for M classes

According to lemma 1, if $H_l(\mathbf{r})$ is a hyperplane that forms a part of the asymptotic decision boundary for separating $R^{(\frac{M}{2})}$ and $R^{(\frac{M}{2}+1)}$, $H_l(\mathbf{r} + \mathbf{c}_i)$ is a hyperplane that is a part of the asymptotic boundary for separating $R^{(i)}$ and $R^{(i+1)}$, where $\mathbf{c}_i = (M - 2i)\mathbf{a}_{\text{rev}}$. In fact, the asymptotic decision boundary for separating $R^{(i+1)}$ and $R^{(i+2)}$ is the translation of the asymptotic decision boundary for separating $R^{(i)}$ and $R^{(i+1)}$ by an amount $2\mathbf{a}_{\text{rev}}$. Note that

$$H_l(\mathbf{r}(k) + \mathbf{c}_i) \triangleq \mathbf{w}_l^T \mathbf{r}(k) + \bar{b}_{l,i} = \bar{H}_l(k) + \bar{b}_{l,i} , \tag{27}$$

where $\bar{b}_{l,i} = \mathbf{w}_l^T \mathbf{c}_i + b_l$. To indicate which asymptotic decision boundary, the index i, $1 \leq i \leq M-1$, is used. The half-space defined by the hyperplane $\mathbf{w}_l^T \mathbf{r} + \bar{b}_{l,i} = 0$ is $\mathcal{H}_l^{(+,i)} \triangleq \{\mathbf{r} : \mathbf{w}_l^T \mathbf{r} + \bar{b}_{l,i} \geq 0\}$, the convex region covering $\mathbf{r}_q^{(+,i)} \in R^{(i+1)}$ is $\mathcal{R}_q^{(+,i)}$, and the decision region for $\hat{s}(k-d) \geq s_{i+1}$ is $\mathcal{R}^{(+,i)}$. The corresponding Boolean logic value for the linear discriminant function $\bar{H}_l(k) + \bar{b}_{l,i}$ is denoted by $\beta_{l,i}(k) = \beta_l(\mathbf{r}(k) + \mathbf{c}_i)$, the Boolean logic value indicating whether $\mathbf{r}(k) \in \mathcal{R}_q^{(+,i)}$ or not is denoted by $\theta_{q,i}(k) = \theta_q(\mathbf{r}(k) + \mathbf{c}_i)$, and the Boolean logic value indicating whether $\mathbf{r}(k) \in \mathcal{R}^{(+,i)}$ or not is denoted by $\alpha_i(k) = \alpha(\mathbf{r}(k) + \mathbf{c}_i)$. The resulting multiple-hyperplane detector can now be summarized. At k:

```
FOR l = 1 to L
    COMPUTE H̄_l(k) = w_l^T r(k);
NEXT l
FOR i = 1 to M − 1
    FOR l = 1 to L
        COMPUTE H̄_l(k) + b̄_{l,i};
    NEXT l
    COMPUTE Boolean logic value α_i(k);
    IF ( NOT α_i(k) ) {
        ŝ(k − d) = s_i;
        BREAK;
    } ELSE IF ( (i == M − 1 ) {
        ŝ(k − d) = s_M;
        BREAK;
    }
NEXT i
```

As all the values of $\bar{b}_{l,i}$ are pre-computed at the design stage, the detector complexity is what is required to compute the L linear discriminant functions, as listed at Table I. Thus the complexity of this multiple-hyperplane detector is L times of the linear-combiner DFE. As long as $L < M^{n_a}$, this multiple-hyperplane detector requires less computation than the optimal Bayesian detector. The L pairs of dominant states are selected from two subsets, which have $2M^{n_a-1}$ states. Empirically, we have found usually $L < 2M^{n_a-1}$.

IV. A Simulation Example

An example was used to test the multiple-hyperplane detector, in which 4-PAM symbols were transmitted over a 3-tap channel specified by the CIR $\mathbf{a} = [0.4\ 1.0\ 0.6]^T$. The structure parameters of the DFE were accordingly set to $m = 3$, $d = 2$ and $n = 2$. The channel state set R had $N_f = 64$ states. Five pairs of dominant states were found from the subsets $R^{(2)}$ and $R^{(3)}$, giving rise to 5 separating hyperplanes. The separating matrix is listed in Table II, from which a required Boolean logic function was obtained. For instance, the states $\mathbf{r}_1^{(+,2)}$ to $\mathbf{r}_6^{(+,2)}$ in $R^{(3)}$ are separated from the opposite-class states $R^{(2)}$ by the two hyperplanes H_1 and H_2; $\mathbf{r}_7^{(+,2)}$ and $\mathbf{r}_8^{(+,2)}$ require H_3 and H_4 for separation from $R^{(2)}$; and $\mathbf{r}_9^{(+,2)}$ to $\mathbf{r}_{16}^{(+,2)}$ are separated from $R^{(2)}$ by H_5. The symbol error rate (SER) performance of this multiple-hyperplane detector is compared with those of the Bayesian and conventional minimum mean square error (MMSE) DFEs in Fig. 2 under different SNR conditions, where it can be seen that there is hardly any SER performance difference between the multiple-hyperplane and full Bayesian detectors. For this example, the full Bayesian detector requires 380 additions, 256 multiplications and 64 e^x function evaluations to detect a symbol. The multiple-hyperplane detector, however, needs only 25 additions and 15 multiplications to make a decision, which is less than 6% of the complexity required by the full Bayesian DFE.

V. Conclusions

We have extended a signal space partitioning technique, originally developed for binary channels, to M-PAM channels. A scheme is presented to construct a multiple-hyperplane partitioning that is asymptotically optimal. The resulting detector consists of a set of linear discriminant functions and associated Boolean logic values, and it has much lower decision complexity compared with the Bayesian detector. Although this multiple-hyperplane detector achieves the optimal Bayesian performance only at the asymptotic case of infinite SNR, in practice, it can close approximate the optimal performance under finite SNR conditions.

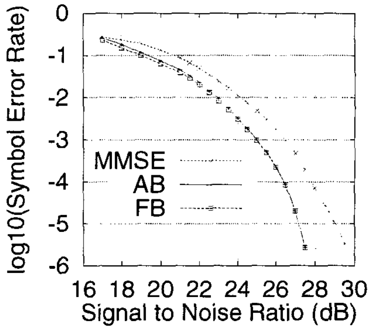

Fig. 2. Performance comparison of the classical MMSE DFE (MMSE), the multiple-hyperplane detector (AB) and the full Bayesian DFE (FB) with detected symbols being fed back.

References

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

[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. Chen, S. McLaughlin, B. Mulgrew and P.M. Grant, "Bayesian decision feedback equaliser for overcoming co-channel interference," *IEE Proc. Communications*, Vol.143, No.4, pp.219–225, 1996.

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

[5] S. Chen, B. Mulgrew, E.S. Chng and G. Gibson, "Space translation properties and the minimum-BER linear-combiner DFE," *IEE Proc. Communications*, Vol.145, No.5, pp.316–322, 1998.

[6] S. Chen, S. Gunn and C.J. Harris, "Decision feedback equalizer design using support vector machines," *IEE Proc. Vision, Image and Signal Processing*, Vol.147, No.3, pp.213–219, 2000.

[7] R.A. Iltis, "A randomized bias technique for the importance sampling simulation of Bayesian equalizers," *IEEE Trans. Communications*, Vol.43, No.2/3/4, pp.1107–1115, 1995.

[8] S. Chen, B. Mulgrew and L. Hanzo, "Asymptotic Bayesian decision feedback equalizer using a set of hyperplanes," *IEEE Trans. Signal Processing*, Vol.48, No.12, pp.3493–3500, 2000.

[9] Y. Kim and J. Moon, "Delay-constrained asymptotically optimal detection using signal-space partitioning," in *Proc. ICC'98* (Atlanta, USA), 1998.

[10] Y. Kim and J. Moon, "Multi-dimensional signal space partitioning using a minimal set of hyperplanes for detecting ISI-corrupted symbols," *IEEE Trans. Communications*, to appear, 2000.

TABLE II

Separability Matrix for the Two Subsets of Channel States.

$R^{(2)}$																$R^{(3)}$															
1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0	0	0	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1
0	0	0	0	0	0	0	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	0	0
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1

0-7803-7097-1/01/$10.00 ©2001 IEEE