# Decision Feedback Equalization Using Multiple-Hyperplane Partitioning for Detecting ISI-Corrupted $M$-ary PAM Signals

S. Chen, L. Hanzo and B. Mulgrew

*Abstract*— **A decision feedback equalizer (DFE) scheme is derived based on multiple-hyperplane partitioning of signal space for detecting $M$-ary PAM symbols transmitted through a noisy intersymbol interference (ISI) channel. The proposed scheme is based on the fact that the optimal Bayesian decision boundary separating two neighbouring signal classes is asymptotically piecewise linear and consists of several hyperplanes, when the signal to noise ratio (SNR) tends to infinity. An algorithm is developed to determine these hyperplanes, which are then used to partition the observation signal space. The resulting detector can closely approximate the optimal Bayesian detector, at an advantage of considerably reduced detector complexity.**

## I. Introduction

The DFE is a powerful technique for combating distortion and interference in communication links and high-density data storage systems. For the class of DFE structures that employs symbol-by-symbol detection, the optimal solution is well understood to be the maximum *a posteriori* probability or Bayesian detector [1]–[5]. 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 the optimal detector increases exponentially with the size of symbol set $M$, the conventional or linear-combiner DFE [6]–[10] is often used in practice to provide a trade-off between performance and detector complexity.

For the binary case, the difference in performance between the conventional and optimal Bayesian DFEs can be explained geometrically: a linear-combiner DFE can only partition the observation space with a hyperplane while the optimal Bayesian detector can do so with a hypersurface [4]. Asymptotically, the optimal Bayesian hypersurface becomes piecewise linear and is made up of a set of hyperplanes [11]. In practice, at large SNR (usually 10 to 20 dB), the Bayesian decision hypersurface can closely be approximated by a multiple-hyperplane form. Thus, a detector that partitions the observation space with multiple hyperplanes can approximate the optimal Bayesian detector more closely than the conventional DFE.

Signal space partitioning techniques for binary ISI channels have been developed. Notably, Moon and Jeon [12] developed a detector using the Voronoi diagram. Kim and Moon [13],[14] improved on the design and developed a new signal space partitioning technique. 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 a combinatorial search and optimization process is carried out to find those hyperplanes, which requires extensive computational efforts. In our previous work [15], we have shown that this design complexity can be avoided and we have proposed a much simpler alternative design to explicitly realize the asymptotic Bayesian decision boundary for binary ISI channels.

In this paper, we extend this multiple-hyperplane detector design to $M$-ary 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 algorithm for the binary case [15] to the general $M$-ary PAM case. Similar to the binary case, the design of our multiple-hyperplane detector for $M$-ary 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 $M > 2$. A simulation example is included to compare the performance and detection complexity of the resulting multiple-hyperplane detector with those of the optimal Bayesian DFE.

## 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 $\sigma_e^2$, and the $M$-ary PAM symbol $s(k)$ takes the value from the symbol set: $\mathcal{S} \triangleq \{s_i = 2i - M - 1, 1 \le i \le M\}$. The SNR of the system is defined as $\text{SNR} \triangleq \left(\sum_{i=0}^{n_a-1} a_i^2\right) \sigma_s^2/\sigma_e^2$, where $\sigma_s^2$ is the symbol variance. The generic DFE studied in this paper

S. Chen and L. Hanzo are with Department of Electronics and Computer Science, University of Southampton, Highfield, Southampton SO17 1BJ, U.K. (Contact author: S. Chen, Tel./Fax: +44 (0)23 8095 6660/4508; E-mail: sqc@ecs.soton.ac.uk.)

B. Mulgrew is with Department of Electronics and Electrical Engineering, University of Edinburgh, King's Buildings, Edinburgh EH9 3JL, U.K.

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 choice is sufficient to guarantee a desired linear separability for different signal classes [8],[9].

The observation vector $\mathbf{y}(k)$ can be expressed as [5],[8]:

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

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

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

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

are the $m \times (d+1)$ and $m \times n$ CIR matrices, respectively. Assuming correct past decisions, we have $\hat{\mathbf{s}}_b(k) = \mathbf{s}_b(k)$ and

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

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

$$\mathbf{r}(k) \stackrel{\triangle}{=} \mathbf{y}(k) - F_2\ \hat{\mathbf{s}}_b(k) \tag{6}$$

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,

$$R \stackrel{\triangle}{=} \{\mathbf{r}_j = F_1\ \mathbf{s}_{fj},\ 1 \le j \le N_f\} \tag{7}$$

can be partitioned into $M$ subsets conditioned on $s(k-d)$:

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

The optimal Bayesian DFE [5] can now be summarized. The $M$ decision variables are given by

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

and the minimum-error-probability decision is defined by

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

The complexity of this optimal detector is given in Table I.

TABLE I

Comparison of decision complexity for the full Bayesian and multiple-hyperplane detectors.

| | Full Bayesian detector | Multiple-hyperplane detector |
|---|---|---|
| Additions | $2n_a M^{n_a} - M$ | $(n_a + M - 2)L$ |
| Multiplications | $(n_a+1)M^{n_a}$ | $n_a L$ |
| $\exp(\cdot)$ | $M^{n_a}$ | — |

## 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 $2\mathbf{a}_{\text{rev}}$

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

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

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

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.

*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$\to\infty$ (or $\sigma_e^2 \to 0$), the influence from all the other $R^{(i)}$ for $i \ne \frac{M}{2}$ and $i \ne \frac{M}{2}+1$ vanishes much more quickly, and it effectively becomes a two-class problem. We have the following definition, similar to the one given in [15].

*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

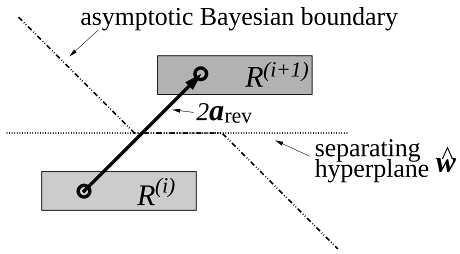


Fig. 1. Illustration of shift property for subsets of channel states.

$\forall \mathbf{r}_j \in R^{(\frac{M}{2})} \bigcup R^{(\frac{M}{2}+1)}$, $\mathbf{r}_j \neq \mathbf{r}^{(+)}$ and $\mathbf{r}_j \neq \mathbf{r}^{(-)}$:

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

where $\bigcup$ denotes the union operator and

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

The following properties of $\mathcal{D}_{\frac{M}{2}}$ are useful in the derivation of a multiple-hyperplane detector (see [11]). A necessary condition for a point $\mathbf{r}_B \in \mathcal{D}_{\frac{M}{2}}$ is

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

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

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

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

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

The following lemma describing $\mathcal{D}_{\frac{M}{2}}$ in the asymptotic case of $\sigma_e^2 \to 0$ 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.

*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 [15]. For the 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 easily seen that pairs of dominant states which define the asymptotic boundary can be selected using the following algorithm:

$L = 0$;
FOR $\mathbf{r}_q^{(+)} \in R^{(\frac{M}{2}+1)}$
  FOR $\mathbf{r}_j^{(-)} \in R^{(\frac{M}{2})}$
    $\mathbf{x} = \frac{\mathbf{r}_q^{(+)}+\mathbf{r}_j^{(-)}}{2}$; $\eta = \|\mathbf{r}_q^{(+)} - \mathbf{x}\|^2$;
    IF $\left(\|\mathbf{r}_l^{(+)} - \mathbf{x}\|^2 > \eta, \forall \mathbf{r}_l^{(+)} \in R^{(\frac{M}{2}+1)}, l \neq q\right)$ AND $\left(\|\mathbf{r}_l^{(-)} - \mathbf{x}\|^2 > \eta, \forall \mathbf{r}_l^{(-)} \in R^{(\frac{M}{2})}, l \neq j\right)$
      $L$ += 1;
      $R_{\text{Asym}} \leftarrow (\mathbf{r}_L^{(+)}, \mathbf{r}_L^{(-)}) \triangleq (\mathbf{r}_q^{(+)}, \mathbf{r}_j^{(-)})$;
    END IF
  NEXT $\mathbf{r}_j^{(-)}$
NEXT $\mathbf{r}_q^{(+)}$

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

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

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:

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

and

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

The hyperplane defined by (20) and (21) is a *canonical* hyperplane with $(\mathbf{r}_l^{(+)}, \mathbf{r}_l^{(-)})$ as its two support vectors [16], 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 the hyperplane $H_l$ 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 $H_l$ if and only if $\mathbf{w}_l^T \mathbf{r}_j + b_l \geq 1$. Similarly, $\mathbf{r}_j \in R^{(\frac{M}{2})}$ is sufficiently separable by $H_l$ if and only if $\mathbf{w}_l^T \mathbf{r}_j + b_l \leq -1$. 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 $H_l$, $1 \leq l \leq L$. This generates the following "separability" matrix:

| | $\mathbf{r}_1^{(-)}$ | $\mathbf{r}_2^{(-)}$ | $\cdots$ | $\mathbf{r}_{N_s}^{(-)}$ | $\mathbf{r}_1^{(+)}$ | $\cdots$ | $\mathbf{r}_{N_s}^{(+)}$ |
|---|---|---|---|---|---|---|---|
| $H_1$ | $h_{1,1}$ | $h_{1,2}$ | $\cdots$ | $h_{1,N_s}$ | $h_{1,N_s+1}$ | $\cdots$ | $h_{1,2N_s}$ |
| $H_2$ | $h_{2,1}$ | $h_{2,2}$ | $\cdots$ | $h_{2,N_s}$ | $h_{2,N_s+1}$ | $\cdots$ | $h_{2,2N_s}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\cdots$ | $\vdots$ | $\vdots$ | $\cdots$ | $\vdots$ |
| $H_L$ | $h_{L,1}$ | $h_{L,2}$ | $\cdots$ | $h_{L,N_s}$ | $h_{L,N_s+1}$ | $\cdots$ | $h_{L,2N_s}$ |

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 $h_{l,j} = 0$. Notice that in every row there are at least two nonzero elements (associated with a dominant pair), and in every column there is at least one nonzero element.

The half-space defined by a hyperplane $H_l$ is given by $\mathcal{H}_l^{(+)} \triangleq \{\mathbf{r} : H_l(\mathbf{r}) \geq 0\}$. 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 formed as the union of all the $\mathcal{R}_q^{(+)}$

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

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

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

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 $q$:

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

*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 $R^{(\frac{M}{2}+1)}$, $H_l(\mathbf{r} + (M-2i)\mathbf{a}_{\text{rev}})$ is a hyperplane that is a part of the asymptotic boundary for separating $R^{(i)}$ and $R^{(i+1)}$. 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 $2\mathbf{a}_{\text{rev}}$. Notice that

$$H_l(\mathbf{r}(k)+(M-2i)\mathbf{a}_{\text{rev}}) \triangleq \mathbf{w}_l^T\mathbf{r}(k)+\bar{b}_{l,i} = \bar{H}_l(k)+\bar{b}_{l,i} \tag{27}$$

where $\bar{b}_{l,i} = \mathbf{w}_l^T((M-2i)\mathbf{a}_{\text{rev}}) + b_l$. 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 $\beta_{l,i}(k) = \beta_l(\mathbf{r}(k) + (M-2i)\mathbf{a}_{\text{rev}})$, the Boolean logic value indicating whether $\mathbf{r}(k) \in \mathcal{R}_q^{(+,i)}$ or not is denoted by $\theta_{q,i}(k) = \theta_q(\mathbf{r}(k)+(M-2i)\mathbf{a}_{\text{rev}})$, and the Boolean logic value indicating whether $\mathbf{r}(k) \in \mathcal{R}^{(+,i)}$ or not is denoted by $\alpha_i(k) = \alpha(\mathbf{r}(k) + (M-2i)\mathbf{a}_{\text{rev}})$.

The resulting multiple-hyperplane detector can now be summarized. At sample $k$:

```
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 proposed multiple-hyperplane detector, in which 4-ary 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 for this example is listed in Table II, from which a required Boolean logic function can be obtained. For instance, the states $\mathbf{r}_1^{(+,2)}$ to $\mathbf{r}_6^{(+,2)}$ in $R^{(3)}$ require the two hyperplanes $H_1$ and $H_2$ for separation from $R^{(2)}$; $\mathbf{r}_7^{(+,2)}$ and $\mathbf{r}_8^{(+,2)}$ are separated from $R^{(2)}$ by $H_3$ and $H_4$; and, finally, $H_5$ separates $\mathbf{r}_9^{(+,2)}$ to $\mathbf{r}_{16}^{(+,2)}$ from the opposite-class $R^{(2)}$.

Thus, given the 5 values of linear discriminant functions $\bar{H}_l(k) + \bar{b}_{l,i}$, $1 \leq l \leq 5$, the detector requires 2 logic AND and 2 logic OR operations to complete the Boolean logic value $\alpha_i(k)$.

The symbol error rate (SER) performance of this multiple-hyperplane detector is compared with those of the full Bayesian and conventional minimum mean square error (MMSE) DFEs in Fig. 2, 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 DFE requires 380 additions, 256 multiplications and 64 $\exp(\cdot)$ 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. The linear MMSE DFE requires only 5 additions and 3 multiplications to make a decision, as it contains a single hyperplane, but its performance is much inferior.

## V. Conclusions

We have extended a signal space partitioning technique, originally developed for binary channels, to $M$-ary PAM channels. A design scheme is presented to automatically 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 optimal 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.

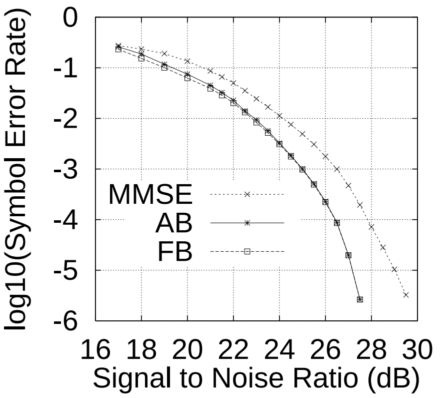


Fig. 2. Performance comparison of the classical MMSE DFE (MMSE), the multiple-hyperplane detector (AB) and the full Bayesian DFE (FB) with detected symbols being fed back.

## References


[1] K. Abend and B.D. Fritchman, "Statistical detection for communication channels with intersymbol interference," *Proc. IEEE*, Vol.58, No.5, pp.779–785, 1970.

[2] J.F. Hayes, T.M. Cover and J.B. Riera, "Optimal sequence detection and optimal symbol-by-symbol detection: similar algorithms," *IEEE Trans. Communications*, Vol.30, No.1, pp.152-157, 1982.

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

[4] 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.

[5] 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.

[6] E. Shamash and K. Yao, "On the structure and performance of a linear decision feedback equalizer based on the minimum error probability criterion," in *Proc. ICC'74*, 1974, pp.25F1–25F5.

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

[8] 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.

[9] S. Chen and B. Mulgrew, "The minimum-SER linear-combiner decision feedback equaliser," *IEE Proc. Communications*, Vol.146, No.6, pp.347–353, 1999.

[10] 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.

[11] 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.

[12] J. Moon and T. Jeon, "Sequence detection for binary ISI channels using signal-space partitioning," *IEEE Trans. Communications*, Vol.47, No.7, pp.891–901, 1998.

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

[14] 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.

[15] S. Chen, B. Mulgrew and L. Hanzo, "Asymptotic Bayesian decision feedback equalizer using a set of hyperplanes," *IEEE Trans. Signal Processing*, to appear, November 2000.

[16] V. Vapnik, *The Nature of Statistical Learning Theory.* New York: Springer-Verlag, 1995.


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 |