# REDUCING BAYES EQUALIZER COMPLEXITY: A NEW APPROACH FOR CLUSTERS DETERMINATION


**J. R. MONTALVÃO F.[1,2], J. C. M.MOTA[2,3], B. DORIZZI[2], C. C. CAVALCANTE[2,4]**

(1) Univ. Tiradentes R.Lagarto, 264 49010-390, Aracaju-SE, Brazil jugurta.montalvao@int-evry.fr
(2) Inst.Nat.Télécom./EPH 9, Rue Charles Fourier 91011 Evry France Bernadette.Dorizzi@int-evry.fr
(3) Univ.Fed.Ceará/CT/DEE C.P.6001, Fortaleza-CE Brazil 60455-760 mota@dee.ufc.br
(4) Univ.Fed.Ceará/PET-Elétrica/CAPES C.P.6001, Fortaleza-CE Brazil 60455-760 charlesc@dee.ufc.br



*Abstract— A new strategy to channel equalization in digital communication is presented. In this approach,* the clustering problem is treated analytically. We propose a systematic bayesian classification using a gaussian approximation of the probability density function for each cluster. The quality of the approximation depends on the number of clusters considered. We show analytically that we can obtain the Bayes equalizer performance, if we use the maximum number of clusters, and the Wiener one, if we use only two clusters (binary signal case). Some computational simulations illustrate power of the presented strategy.*


## 1. Introduction

Equalization in digital communication systems is used in order to reconstruct the transmitted symbols and combat the intersymbol interference (ISI) effect. When channels are characterized by a finite impulse response (FIR) filter and an additive white gaussian noise source, it is shown in some recent works [1][2][7] that it is possible to use a Radial Basis Function network (RBF) to perform the optimum bayesian equalizer.

However, despite its desirable performance, this approach is strongly limited by the inherent equalizer complexity and the high convergence time of the most common learning techniques.

On the other hand, linear equalizers have been used for long time. Their importance is associated to their low complexity and theoretical tractability. However, it has been shown [7] that the optimum equalizer is nonlinear in all realistic cases where noise is present and the channel is non-minimum phase. Furthermore, the considered error function minimized by the Wiener linear equalizer, the Mean Square Error (MSE) [4], is not equivalent to the symbol error rate (SER) [5] which is the normally used criterion in the digital communication context.

It is well known [9] that a nonlinear block detection equalization based on the principle of Maximum Likelihood Sequence Estimator will provide the best equalization performance when the channel is completely known. Its high implementation complexity is one of the main reasons for using other nonlinear symbol decision class equalizers with simpler implementations but poorer performance. In this context, the communication community has recognized the bayesian symbol-decision class equalizers as optimal solutions which deals with the equalization problem as a classification one [10].

Several recent works have been done to reduce complexity using clustering methods [2][3] over the channel's output. The most common way to reduce complexity is to find an approximate optimum Bayes decision boundary. When a RBF neural network is considered, complexity can be highly reduced by using variable selection algorithms as in [8].

In this work, we propose a new analytical strategy to reduce the bayesian equalizer complexity using an estimation of the channel model parameters. Our approach confirms analytically that we can obtain the Wiener SER performance, if we use the less complex bayesian equalizer structure, or the optimal Bayes equalizer, if we use the more complex one.

The main interest of this approach is thus to render possible a full range of gradual choices between complexity and performance.

In the section 2, we present the theoretical basis to this approach and in section 3 our simulation results are presented and exposed compared with the performance of others classical equalizers. The conclusions are presented in the last part.

## 2. Equalization and Classification

In Fig.1, we depict a classical digital communication system model. The message source emits one symbol $a(n)$ every $T$ seconds, with the symbol belonging to a finite alphabet. In this work we will consider the bipolar case, where $a(n)$ is taken from the set $\{\pm 1\}$, forming an i.i.d. sequence. The noise $b(n)$ is an additive gaussian noise with zero mean and variance $\sigma_b^2$, the causal channel impulse response $f(n)$ has a finite length $N$, and $\hat{a}(n-d)$ is a decided symbol with delay decision

0-7803-5030-8/98/$10.00 ©1998 IEEE

$d \geq 0$.

For simplicity of representation, we use column vectors defined as:

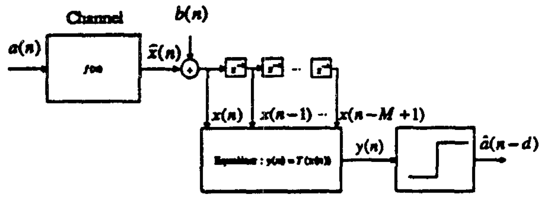


Fig. 1: Model of a data transmission system with equalizer.

$\mathbf{a}(n) = [a(n) \quad a(n-1) \quad \cdots \quad a(n-d) \quad \cdots \quad a(n-N-M+1)]^T$

$\mathbf{f} = [f_0 \quad f_1 \quad \cdots \quad f_{N-1}]^T$,

$\mathbf{b}(n) = [b(n) \quad b(n-1) \quad \cdots \quad b(n-M+1)]^T$ and

$\mathbf{x}(n) = [x(n) \quad x(n-1) \quad \cdots \quad x(n-M+1)]^T$. We also define the matrix

$$\mathbf{F} = \begin{bmatrix} f_0 & 0 & \cdots & 0 \\ f_1 & f_0 & \cdots & \vdots \\ f_2 & f_1 & \cdots & 0 \\ \vdots & f_2 & \cdots & f_0 \\ f_{N-1} & \vdots & \cdots & f_1 \\ 0 & f_{N-1} & \cdots & f_2 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & f_{N-1} \end{bmatrix}$$, where $\mathbf{F}$ has $N+M-1$ rows and $M$ columns. Then, we can write:

$$\mathbf{x}(n)^T = \hat{\mathbf{x}}(n)^T + \mathbf{b}(n)^T = \mathbf{a}(n)^T \cdot \mathbf{F} + \mathbf{b}(n)^T \quad (1)$$

where $\hat{\mathbf{x}}(n)$ is the channel state vector and $\mathbf{b}(n)$ is the noise vector.

An interesting equalizer that we will consider in the first part of subsection 2.1 is a linear mapping $y(n) = T_l(\mathbf{x}(n)) = \mathbf{x}(n)^T \cdot \mathbf{h}$, where $\mathbf{h}$ is the vector $\mathbf{h} = [h_0 \quad h_1 \quad \cdots \quad h_{M-1}]^T$. This also corresponds to find the best projection of the samples x(n) onto a line in the direction of **h**. In the classification sense [6], this corresponds to find the Fisher's linear discriminant **h**. The Fig. 2 illustrates the effect of choosing two different directions for a two-dimensional example.

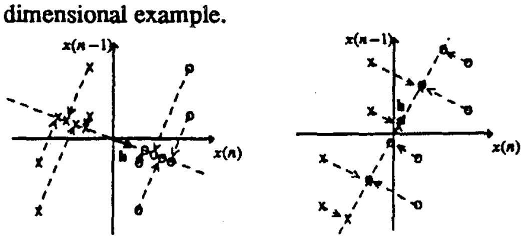


Fig. 2: Projection of the samples on 2 different lines (in 2-D case).

In this illustration, class 1 ('o') corresponds to the transmitted symbol $a(n-d) = 1$ and class $-1$ ('x') corresponds to the transmitted symbol $a(n-d) = -1$. Clearly, we look for an orientation for which the projected samples have maximum *between-class* separation and, at the same time, the minimum *within-class* dispersion [6]. This corresponds to maximize the functional $J(\mathbf{h}) = \dfrac{\mathbf{h}^T \mathbf{R}_c \mathbf{h}}{\mathbf{h}^T \mathbf{R}_{x-c} \mathbf{h}}$, where $\mathbf{R}_c$ is the *between-class scatter matrix* and $\mathbf{R}_{x-c}$ is the *within-class scatter matrix*. In the stochastic context of digital signal processing, it is interesting to interpret the *total scatter matrix* $\mathbf{R}_x = \mathbf{R}_c + \mathbf{R}_{x-c}$ as the correlation matrix of the random vector **x**(n).

### 2.1. Calculations for Two Clusters

To perform a simple equalizer device, the first possible approach is to project the corrupted samples $\mathbf{x}(n)$ onto a line in the direction of **h**. In this sense, the best direction of **h** to cluster the projected samples can be investigated. Considering the bipolar case, we can calculate two centers related to the symbol *a(n-d)*:

$$\mathbf{c}_1 = E\{\mathbf{x}(n)\} \text{ when } a(n\text{-}d)\text{=}1 \quad (2)$$

$$\mathbf{c}_{-1} = E\{\mathbf{x}(n)\} \text{ when } a(n\text{-}d)\text{=-}1. \quad (3)$$

Then, applying (1) to (2) and (3),

$\mathbf{c}_1^T = E\{\mathbf{a}(n)^T \cdot \mathbf{F} + \mathbf{b}(n)^T\}$ when $a(n\text{-}d)\text{=}1$, or

$\mathbf{c}_1^T = [0 \quad 0 \quad \ldots \quad 1 \quad \ldots \quad 0] \cdot \mathbf{F}$, where 1 is at rank *d*,

and $\mathbf{c}_{-1}^T = E\{\mathbf{a}(n)^T \cdot \mathbf{F} + \mathbf{b}(n)^T\}$ when $a(n\text{-}d)\text{=-}1$, or

$\mathbf{c}_{-1}^T = [0 \quad 0 \quad \ldots \quad -1 \quad \ldots \quad 0] \cdot \mathbf{F}$. So, defining an helpful auxiliary matrix as :

$$\mathbf{F}_* = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 \\ f_d & f_{d-1} & \cdots & f_{d-M+1} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \leftarrow \begin{array}{c}\text{delay position} \\ \text{row}: d\end{array}$$

where *d* can be greater than *N*-1, we can write:

$$\mathbf{c}^T = \mathbf{a}(n)^T \cdot \mathbf{F}_* \quad (4)$$

Then, $\mathcal{C} = \{\mathbf{c}_{-1}, \mathbf{c}_1\}$ is a set of centers and **C** is a binary random variable with $P(C\text{=}\mathbf{c}_1)\text{=}0.5$ and $P(C\text{=}\mathbf{c}_{-1})\text{=}0.5$.

In a similar way, we define two random variables around each center:

$\mathbf{x}_{-1}(n)^T = \mathbf{a}(n)^T \cdot \mathbf{F} + \mathbf{b}(n)^T$, when $a(n-d) = -1$
and
$\mathbf{x}_1(n)^T = \mathbf{a}(n)^T \cdot \mathbf{F} + \mathbf{b}(n)^T$, when $a(n-d) = 1$

To calculate the between-class scatter matrix $\mathbf{R}_c$, we need the mean vector $\mathbf{m} = \mathrm{P}(\mathrm{C} = \mathbf{c}_1)\mathbf{c}_1 + \mathrm{P}(\mathrm{C} = \mathbf{c}_{-1})\mathbf{c}_{-1}$ $= \begin{bmatrix} 0 & \cdots & 0 \end{bmatrix}^T$. Then from equation (4),

$$\mathbf{R}_c = E\left\{(\mathbf{c} - \mathbf{m})(\mathbf{c} - \mathbf{m})^T\right\} = E\left\{\mathbf{F}_*^T \cdot \mathbf{a}(n)\left(\mathbf{a}(n)^T \cdot \mathbf{F}_*\right)\right\}$$

and $\mathbf{R}_c = \mathbf{F}_*^T E\left\{\mathbf{a}(n)\mathbf{a}(n)^T\right\}\mathbf{F}_*$

$$\mathbf{R}_c = \sigma_a^2 \mathbf{F}_*^T \mathbf{F}_* \qquad (5)$$

In a similar way, we can calculate the within-class scatter matrix $\mathbf{R}_{x-c}$: ($i$ denotes cluster $i$, $i$=1 or $i$=-1)

$$\mathbf{R}_{x_i - c_i} = E\left\{(\mathbf{x}_i - \mathbf{c}_i)(\mathbf{x}_i - \mathbf{c}_i)^T\right\} =$$
$$= E\left\{\left((\mathbf{F}^T - \mathbf{F}_*^T)\mathbf{a}(n) + \mathbf{b}(n)\right)\left(\mathbf{a}(n)^T(\mathbf{F} - \mathbf{F}_*) + \mathbf{b}(n)^T\right)\right\}$$

$$\mathbf{R}_{x_i - c_i} = \sigma_a^2(\mathbf{F}^T - \mathbf{F}_*^T)(\mathbf{F} - \mathbf{F}_*) + \sigma_b^2 \mathbf{I}.$$

One can remark that $\mathbf{R}_{x_i - c_i}$ doesn't depend of $i$. In other words, the within-class scatter matrix is equal to the scatter matrix of each cluster:

$$\mathbf{R}_{x-c} = (\mathbf{R}_{x_{-1} - c_{-1}} + \mathbf{R}_{x_1 - c_1})/2 =$$
$$\sigma_a^2(\mathbf{F}^T - \mathbf{F}_*^T)(\mathbf{F} - \mathbf{F}_*) + \sigma_b^2 \mathbf{I}.$$

Expanding this expression and taking into account that
$\mathbf{F}_*^T \mathbf{F} = \mathbf{F}_*^T \mathbf{F}_* = \mathbf{F}^T \mathbf{F}_*$, we have:

$$\mathbf{R}_{x-c} = \sigma_a^2(\mathbf{F}^T\mathbf{F} - \mathbf{F}_*^T\mathbf{F}_*) + \sigma_b^2 \mathbf{I} \qquad (6)$$

Now, we can apply these results to $J(\mathbf{h})$. The vector $\mathbf{h}_o$ which maximizes this functional satisfies the eigenvalue problem $\mathbf{R}_c \mathbf{h}_o = \lambda \mathbf{R}_{x-c} \mathbf{h}_o$. Then, to investigate the direction of this vector, we can apply (5) and (6) in this last equation:

$$\sigma_a^2 (\mathbf{F}_*^T \mathbf{F}_*)\mathbf{h}_o = \lambda\left(\sigma_a^2(\mathbf{F}^T\mathbf{F} - \mathbf{F}_*^T\mathbf{F}_*) + \sigma_b^2 \mathbf{I}\right)\mathbf{h}_o$$

, we have thus:

$$\sigma_a^2(1+\lambda)(\mathbf{F}_*^T \mathbf{F}_*)\mathbf{h}_o = \lambda\left(\sigma_a^2(\mathbf{F}^T\mathbf{F}) + \sigma_b^2 \mathbf{I}\right)\mathbf{h}_o.$$

Considering equality $\mathbf{F}_*^T \mathbf{F} = \mathbf{F}_*^T \mathbf{F}_* = \mathbf{F}^T \mathbf{F}_*$, we can obtain: $(1+\lambda)k\mathbf{p} = \lambda \mathbf{R}_x \mathbf{h}_o$, where

$$\mathbf{p} = \sigma_a^2 \mathbf{c}_1 = \sigma_a^2 \left[f_d \quad f_{d-1} \quad \cdots \quad f_{d-M+1}\right]^T =$$
$$= E\{\mathbf{x}(n)a(n-d)\} \qquad (7)$$

and $k = \mathbf{c}_1^T \mathbf{h}_o$ and $\mathbf{R}_x = E\left\{\mathbf{x}(n)\mathbf{x}(n)^T\right\}$. Finally, we have:

$$\mathbf{h}_o = \left(\frac{1+\lambda}{\lambda}\right)k \cdot \mathbf{R}_x^{-1}\mathbf{p} \qquad (8)$$

If we consider only the vector $\mathbf{h}_o$ direction, we will show that, if $\left(\frac{1+\lambda}{\lambda}\right)k \neq 0$, $\mathbf{h}_o$ and the Wiener equalizer solution ($\mathbf{h}_W = \mathbf{R}_x^{-1}\mathbf{p}$) [4] have the same direction.

In the decision error sense, the decision boundary has more importance than the equalization mapping itself. For this linear equalizer, the decision boundary corresponds to the values of $\mathbf{x}$ satisfying:

$$\mathbf{x}^T \mathbf{h}_o = 0 \qquad (9)$$

Applying (8) in (9) gives:
$\left(\frac{1+\lambda}{\lambda}\right)k \cdot \mathbf{x}^T \mathbf{R}_x^{-1}\mathbf{p} = 0$, or

$$\mathbf{x}^T \mathbf{R}_x^{-1}\mathbf{p} = 0 \qquad (10)$$

This direction is thus the same found by the Wiener equalizer.

In another hand, we can implement a bayesian equalizer, with a RBF structure, considering only two gaussians centered at $\mathbf{c}_{-1}$ and $\mathbf{c}_1$, respectively. The mapping is now a nonlinear one:

$$y(n) = T_{nl}(\mathbf{x}(n)) =$$
$$= \exp\left(\frac{-r_1(\mathbf{x}(n))^2}{2}\right) - \exp\left(\frac{-r_{-1}(\mathbf{x}(n))^2}{2}\right)$$

where $r_1(\mathbf{x}(n))^2 = \left(\mathbf{x}(n) - \mathbf{c}_1\right)^T \mathbf{R}^{-1}\left(\mathbf{x}(n) - \mathbf{c}_1\right)$
and
$r_{-1}(\mathbf{x}(n))^2 = \left(\mathbf{x}(n) - \mathbf{c}_{-1}\right)^T \mathbf{R}^{-1}\left(\mathbf{x}(n) - \mathbf{c}_{-1}\right)$ are the Mahalanobis distance [6] between $\mathbf{x}(n)$ and each center and $\mathbf{R}$ is the metric used to compensate for the nonradial dispersion of x(n). Fig.3 illustrates the scheme of this bayesian equalizer.

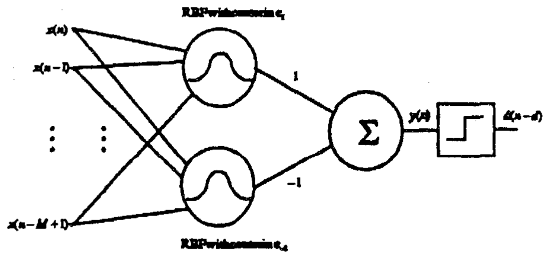


Fig. 3: Bayesian equalizer with two centers.

In this case, the decision boundary corresponds to the values of x(n) satisfying the equation :

$$\left(\mathbf{x}(n)-\mathbf{c}_1\right)^T \mathbf{R}^{-1}\left(\mathbf{x}(n)-\mathbf{c}_1\right)=$$
$$=\left(\mathbf{x}(n)-\mathbf{c}_{-1}\right)^T \mathbf{R}^{-1}\left(\mathbf{x}(n)-\mathbf{c}_{-1}\right)$$

To find this boundary, we know that $\mathbf{c}_1=-\mathbf{c}_{-1}$, so

$$\left(\mathbf{x}(n)-\mathbf{c}_1\right)^T \mathbf{R}^{-1}\left(\mathbf{x}(n)-\mathbf{c}_1\right)=$$
$$=\left(\mathbf{x}(n)+\mathbf{c}_1\right)^T \mathbf{R}^{-1}\left(\mathbf{x}(n)+\mathbf{c}_1\right)$$

After some simplifications, we have

$$\mathbf{c}_1{}^T \mathbf{R}^{-1}\mathbf{x}(n)=-\mathbf{x}(n)^T \mathbf{R}^{-1}\mathbf{c}_1$$

Assuming that $\mathbf{R}$ is symmetric and positive definite, then $\mathbf{R}^{-1}$ exists and is symmetric too. The last equation can be rewritten as

$(\mathbf{x}(n)^T(\mathbf{R}^{-1})^T\mathbf{c}_1)^T=-\mathbf{x}(n)^T\mathbf{R}^{-1}\mathbf{c}_1$ or

$(\mathbf{x}(n)^T\mathbf{R}^{-1}\mathbf{c}_1)^T=-\mathbf{x}(n)^T\mathbf{R}^{-1}\mathbf{c}_1$, and the

decision boundary satisfies $\mathbf{x}(n)^T\mathbf{R}^{-1}\mathbf{c}_1=0$.

Then, using (7) we have

$$\mathbf{x}(n)^T\mathbf{R}^{-1}\mathbf{p}=0 \tag{11}$$

Now, considering that $\mathbf{R}_{x-c}$ and $\mathbf{R}_x$ are both symmetrical and positive definite, we can take $\mathbf{R}=\mathbf{R}_{x-c}$ or $\mathbf{R}=\mathbf{R}_x$. In the second case, if we compare (11) and (10), it is evident that we will obtain a decision boundary equal to that of Wiener. Hence, this two clusters bayesian equalizer has the same SER as the Wiener one. However, in order to approximating the optimum Bayes decision boundary, the best choice is $\mathbf{R}=\mathbf{R}_{x-c}$ because this metric takes into account only the dispersion of each cluster.

For both choices of R, the decision boundaries are equivalent to that implemented by a transversal linear filter. Fig.4 illustrates this equivalence. Moreover, as we can see in the experimental results, both choices of R result in a similar SER performance, despite the best MSE performance of the Wiener equalizer.

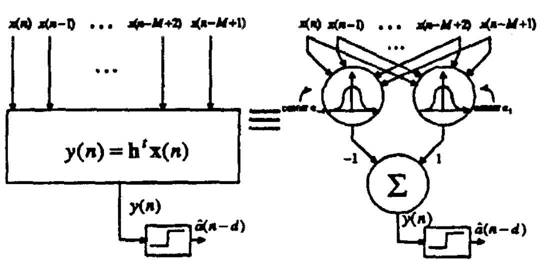


Fig. 4: Equivalence between the two equalizers.

### 2.2. Calculations for Several Clusters

To improve the equalizer performance, we can find several clusters by similar considerations. The first step is to take the auxiliary matrix as :

$$\mathbf{F}_* = \begin{array}{cl} \begin{bmatrix} 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 \\ f_{d-P} & f_{d-P-1} & \cdots & f_{d-P-M+1} \\ \vdots & \vdots & \cdots & \vdots \\ f_d & f_{d-1} & \cdots & f_{d-M+1} \\ \vdots & \vdots & \cdots & \vdots \\ f_{d+Q} & f_{d+Q-1} & \cdots & f_{d+Q-M+1} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} & \leftarrow \text{delay position row} \end{array}$$

where $P$ and $Q$ are constants. We can write

$$\mathbf{c}^T=\mathbf{a}(n)^T\cdot\mathbf{F}_* \tag{12}$$

Then, $C=\{\mathbf{c}_{-1,-1,\ldots,-1},\mathbf{c}_{-1,-1,\ldots,1},\ldots,\mathbf{c}_{1,1,\ldots,1}\}$ is now a set of $2^{P+Q+1}$ vectors. Each center is associated to the label 1 or −1, according to the correspondent symbol at the delay position, a($n$-$d$).

The mean vector is always the null vector:

$$\mathbf{m}=\sum_{i,j,\ldots,m}\mathrm{P}(\mathrm{C}=\mathbf{c}_{i,j,\ldots,m})\cdot\mathbf{c}_{i,j,\ldots,m}$$

$=\begin{bmatrix}0 & \cdots & 0\end{bmatrix}^T$, and the matrices $\mathbf{R}_c$ and $\mathbf{R}_{c-x}$ are calculated as in (5) and (6).

The resulting equalizer assumes that each hyperellipsoidal cluster of equal size and shape, corresponding to each center, has a gaussian distribution of probability. Such approximation makes possible an equalizer implementation as shown in Fig.5, with 4 clusters in this example. Each weight in the last layer is equal to the label of its corresponding cluster, i. e. $a(n-d)$.

Furthermore, if we take $P+Q=N+M-2$, we have the maximum number of clusters and $F_*=F$. Consequently, $\mathbf{R}_{x-c}=\sigma_b^2\mathbf{I}$, and this equalizer will perform the optimal Bayes one.

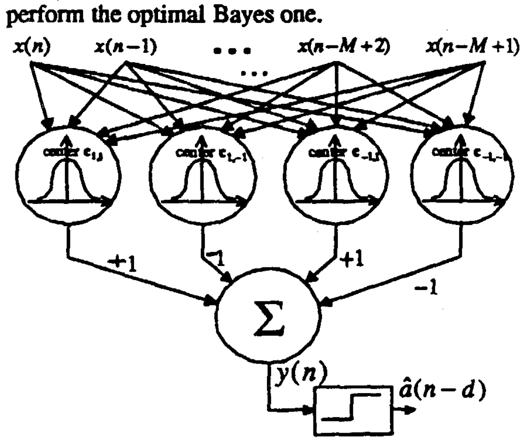


Fig. 5: Bayesian equalizer with four centers.

### 2.3 A Good Choice of $P$ and $Q$

This approach clearly supposes that the channel is known or well identified. In this way, we can choose the centers considering the rows of **F** where the highest (in module) channel coefficient appears. The resulting bayesian equalizer has $2^M$ clusters and $P+Q+1=M$. For example, considering the simple channel $\mathbf{f}=[-0.2 \;\; 0.5 \;\; 1 \;\; -0.6 \;\; 0.3]^T$, we have

$$\mathbf{F}=\begin{bmatrix} -0.2 & 0 \\ 0.5 & -0.2 \\ 1 & 0.5 \\ -0.6 & 1 \\ 0.3 & -0.6 \\ 0 & 0.3 \end{bmatrix}$$

The greatest coefficient is equal to 1 and it appears at the third and fourth rows of **F**. Then, the chosen centers are the four vectors $[\pm(1-0.6) \;\; \pm(0.5+1)]^T$. The labels of these centers depend on the delay, which can be 2 or 3 in this example. Choosing $d$=3, the two centers associated to the symbol $a(n\text{-}3)$=1 are $\mathbf{c}_{1,1}=[(1-0.6) \;\; (0.5+1)]^T$ and $\mathbf{c}_{-1,1}=[(-1-0.6) \;\; (-0.5+1)]^T$, and those associated to the symbol $a(n\text{-}3)$=-1 are $\mathbf{c}_{1,-1}=[(1+0.6) \;\; (0.5-1)]^T$ and $\mathbf{c}_{-1,-1}=[(-1+0.6) \;\; (-0.5-1)]^T$. Fig.6 shows these centers ('*'), all the possible states ('x' and 'o'), and indicates their dispersion around each corresponding center. The resulting equalizer is also shown in Fig. 5.

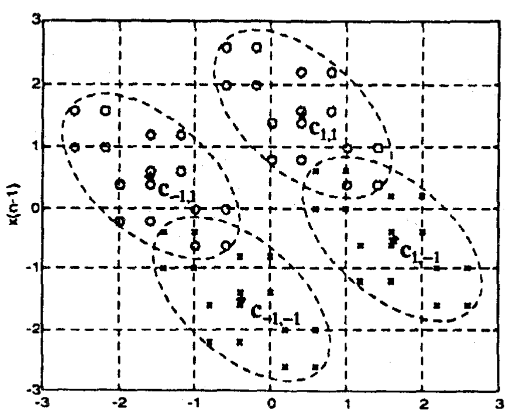


Fig. 6: Centers and clusters illustration in a two-dimensional example.

### 3. Experimental Results

Taking into account the justifications in [7] to use nonlinear equalizers, we know that the linear transversal equalizer order can be increased but, in a highly noisy situation, this increases also the total power of the noise at the equalizer input. So, it has been shown that there is little to be gained in terms of SER by increasing the order $M$ beyond a certain limit.

Fig.7 illustrates this linear equalizer limitation. In this example, the channel is $\mathbf{f}=[1.0 \;\; 0.8 \;\; 0.5]$, the delay decision is $d$= 0 and two different noise levels have been considered: $\sigma_b^2=0.2$ and $\sigma_b^2=0.1$.

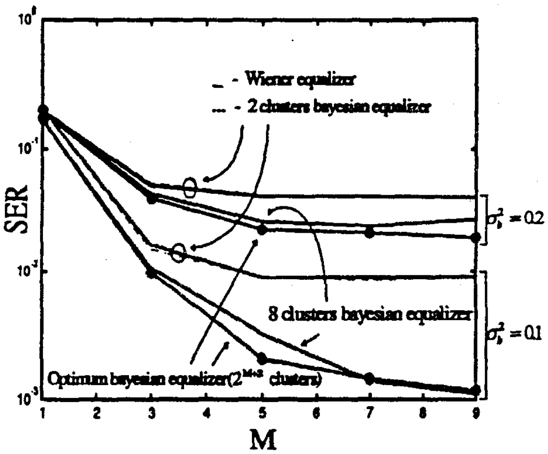


Fig. 7 : Performance versus number of taps.

The same Fig.7,shows the simulation results of the proposed bayesian equalizer with 2, 8 and $2^{M+N-1}$ clusters. The first one and the Wiener equalizer have a similar performance, as expected. The last equalizer corresponds to the optimum Bayes equalizer. Comparing these results we can see that the 8-clusters bayesian equalizer ($P$+$Q$+1=3) outperforms the Wiener equalizer even beyond its minimum SER. Moreover, this 8-clusters bayesian equalizer has not a « prohibitive » complexity, mainly if compared to the optimum Bayes equalizer to this channel ($2^{M+2}$ clusters).

We present also some other simulation results in Fig.8. In this case the channel used is $\mathbf{f}=[-0.21 \;\; -0.50 \;\; 0.72 \;\; 0.36 \;\; 0.21]$, the equalizer order is $M$=5 and the decision delay is d=4. In this figure we can see the gradual relationships that exists between equalizer complexities and their respective performance. The 32-clusters equalizer corresponds to that proposed in the section 2.3. The complexity reduction results are compatible to that in [8].

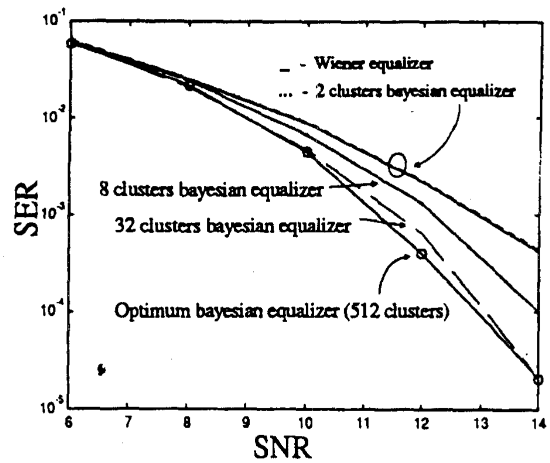


Figure 8 : Performance versus SNR.

## 4. Conclusions

We have investigated in this paper the implementation of a bayesian equalizer, after an identification of the channel coefficients. The presented approach makes possible a gradual compromise between complexity and performance depending of the number of states chosen for the gaussian modelisation of each class. The analytical formulation shows and the experimental results confirm that the less complex equalizer implementation provides equivalent Wiener SER performance. Moreover, it is also shown that, from this lower born, we can increase the equalizer complexity (number of centers) enhancing its performance. At the upper born we recover the optimal bayesian equalizer.

However, in this approach we need to know the channel impulse response or its estimation. Then, we see as a natural continuity work a clustering based channel identification strategy which take in consideration this presented strategy to accelerate the clustering algorithm.

Another point to be investigated is the performance loss of the proposed bayesian equalizer related to the channel identification error.


## 5. Acknowledgments

The authors would like to thank Mr. F. Jacquet by his suggestions and remarks. We also thank the INT-Evry/France and the CNPq/Brazil by the partial financial support.


## 6. References:


[1] S. CHENG, B. MULGREW, P. M. GRANT (1993) "A Clustering Technique for Digital Communications Channel Equalization Using Radial Basis Function Networks", *IEEE Transaction on Neural Network*, vol. 4, NO. 4, pp. 571-579.

[2]E. CHNG, H. YANG, W. SKARBEK (1996) "Reduced Complexity Implementation of Bayesian Equalizer Using Local RBF Networks for Channel Equalization Problem", *Electronics Letters*, vol. 32, NO. 1, pp. 17-19.

[3] S. THEODORIDIS, C. M. S. SEE, C. F. N. COWAN (1992) "Nonlinear Channel Equalization Using Clustering Technique", ICC/IEEE, pp. 1277-1279.

[4] S. HAYKIN (1996). *Adaptive Filter Theory*, 3rd ed., Prentice-Hall, Englewood Cliffs, NJ.

[5] Jean-Marc BROSSIER (1997). *Signal et Communication Numérique : égalisation et synchronisation*, Hermes, Paris.

[6] DUDA, R. O. and P. E. HART (1973). *Pattern Classification and Scene Analysis*, Wiley, New York.

[7] G. J. GIBSON, S. SIU, S. CHEN, C. F. N. COWAN and P. M. GRANT (1990). "The Application of Nonlinear Architectures to Adaptive Channel Equalization", Proc. IEEE, pp. 0649-0653.

[8] B. DORIZZI, J. C. MOTA, F. ALBU (1997). "A Step Towards Equalization for Radiomobile Channel: Neural Networks and Variable Selection", COST 254/IEEE.

[9] J. G. PROAKIS (1984). *Digital Communications*, McGraw-Hill.

[10] B. MULGREW (1996). "Applying Radial Basis Functions", IEEE Signal Processing Magazine.