A novel cluster based MLSE equalizer for 2-PAM signaling scheme

Yannis Kopsinis , Sergios Theodoridis
Dept. of Informatics and Telecommunications, University of Athens
Panepistimioupolis, Ilisia 15784, Athens, Greece
e-mail: kopsinis@di.uoa.gr, stheodor@di.uoa.gr

ABSTRACT

In this paper a new cluster based Maximum Likelihood Sequence Equalizer is presented. The novelty of the algorithm consists of a novel technique for the estimation of all the centers around which the received observations are clustered. For a channel of order L and a 2-PAM signaling scheme, only L of the cluster centers need to be estimated, and the rest $2^L - L$ clusters are subsequently obtained via simple operations. This has a two fold advantage compared to previously proposed cluster based algorithms. It reduces dramatically the computational complexity as well as the required length of the training sequence. The method is compared with the standard LMS based MLSE and the Bayesian RBF equalizer. The results are very favorable for the new technique, both from a computational as well as performance point of view.

1 INTRODUCTION

One of the major problems encountered in the receiver design of any communication system is that of Intersymbol Interference (ISI). The part of the receiver used to mitigate ISI is the equalizer, and the literature related to the task is very rich [1].

An optimal sequence equalizer is based on the Maximum Likelihood Sequence Estimation (MLSE) scheme [2] and is efficiently implemented via the Viterbi algorithm. The resulting equalizer is known as the MLSE-VA equalizer or simply MLSE equalizer and requires the channel impulse response (CIR) to be known. In practice, the LMS algorithm and its variants, are very attractive for the channel estimation due to their structural simplicity and their low computational requirements.

In the current paper, a novel MLSE equalizer is presented that circumvents the problem of explicit CIR parametric modeling and at the same time leads to substantial computational savings. The proposed equalizer belongs to the family of Cluster-Based Sequence Equalizers (CBSE) [3]-[6]. These equalizers utilize the clusters formed by the received observations at the receiver front end. The current paper addresses the two basic drawbacks associated with the cluster-based equalizers, i.e., the high computational requirements and the need of long training sequences. The first problem is solved by a new CBSE equalizer, called 1-D CBSE, that operates in the one dimensional space leading to a complexity lower than that of the standard MLSE equalizer, without any loss in performance. The second one is addressed by a novel cluster center estimation method, that exploits the symmetries that underly the generation mechanism of the clusters formed by the received observations.

2 DESCRIPTION OF THE COMMUNICATION SYSTEM AND CHANNEL MODEL

Figure 1 illustrates the adopted communication system, where x_k is the kth transmitted symbol taking values from the data set 1, -1 , y_k is the kth received observation corrupted by noise n_k, and $\bar{y}_k$ denotes the corresponding noiseless received sample.

The transmitted symbol sequence has been assumed to be independently and identically distributed (i.i.d.) and the communication channel can be modeled as a finite impulse response filter spanning over L consecutive transmitted symbols, with transfer function $H(z)$. Thus the received signal sampled at $t = kT$, with T being the transmission period of the symbols, is given by

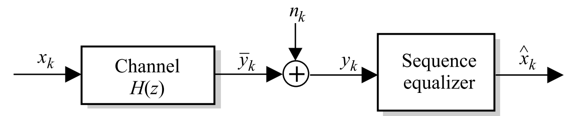

Figure 1: Communication system model

$$y_k = \sum_{j=0}^{L-1} x_{k-j} h_j + n_k = \boldsymbol{h}^T \boldsymbol{x}_k + n_k \equiv \bar{y}_k + n_k \quad (1)$$

where n_k is the white noise, h_i are the coefficients of the discrete equivalent of the channel impulse response (CIR) and $\boldsymbol{h} = [h_0, h_1, \cdots, h_{L-1}]^T$ is the corresponding vector. Moreover the vector $\boldsymbol{x}_k = [x_k,\ x_{k-1},\ \cdots\ , x_{k-L+1}]^T$ of L successively transmitted symbols, is associated with the noiseless observations $\bar{y}_k$ at any time instant k.

3 SUMMARY OF THE MLSE

The task of the MLSE equalizer is to "choose" that sequence of symbols (out of 2^N) $\hat{X} = \{\hat{x}_1, \hat{x}_2, \cdots, \hat{x}_k, \cdots, \hat{x}_N\}$ that maximizes the likelihood of the received sequence of observations $Y = \{y_1, y_2, \cdots, y_k, \cdots, y_N\}$, i.e, maximizes the joint probability $P(Y|X)$. The MLSE equalizer, which is adopted here [2], comprises a whitened matched filter, followed by the Viterbi algorithm. The Viterbi algorithm is efficiently implemented by utilizing a trellis diagram.

The states at any stage k of the trellis diagram are related to the $L-1$ most recent transmitted symbols, i.e.,

$$s_k \rightarrow (x_{k-1}, x_{k-2}, \cdots, x_{k-L+1}) \quad (2)$$

The number of all possible $(L-1)$-symbol length sequences is 2^{L-1}. Thus each state corresponds to one of these 2^{L-1} possible vectors that can be formed from $L-1$ symbols. There are 2 allowable transitions that emerge from a state s_k and terminate at 2 different states s_{k+1}, leading to a total of 2^L transitions B_i, $1 \leq i \leq 2^L$. Thus each transition B_i at every stage k, is associated with a sequence of L symbols, determined by the two successive states associated with the specific transition.

$$B_i : (s_{k+1}, s_k) \rightarrow (x_k, x_{k-1}, x_{k-2}, \cdots, x_{k-L+1}) \quad (3)$$

Each transition is associated with a cost, contributing to the total cost of a path along the states. The cost of the ith transition, at any stage k is given by the Euclidean distance metric

$$D_k^i = |y_k - \boldsymbol{h}^T \boldsymbol{x}_i|^2 \quad (4)$$

where $\boldsymbol{x}_i$ is the $L-$element vector of the sequence of symbols $[x_k, x_{k-1}, \cdots, x_{k-L+1}]^T$, which is related to the ith transition B_i.

4 THE NEW CLUSTER BASED SEQUENCE EQUALIZER 1-D CBSE

Considering that each transition is associated with a symbol vector $\boldsymbol{x}_i$ and taking into account equations (1) and (4) we can infer that each one of the transitions in the trellis diagram corresponds to one of the 2^L possible noiseless observations $\bar{y}^i = \boldsymbol{h}^T\boldsymbol{x}_i$, which are uniquely determined by the vector $\boldsymbol{x}_i$, $1 \leq i \leq 2^L$ and the channel impulse response $\boldsymbol{h}$. Hence, the possible values that $\bar{y}_i$ can take are nothing else than the points (centers) around which the received samples (observations) y_k are clustered, due to the presence of the noise. Figure 2 shows the 1-dimensional plot of

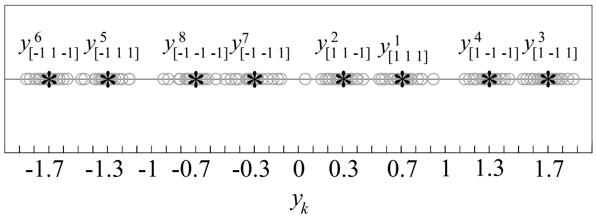

Figure 2: Plot of the clusters formed by the received observations. The stars denote the cluster centers, and the gray circles are the corrupted by noise received observations.

the received observations for a 3-tap channel with transfer function $H(z) = 1 - 0.5z^{-1} + 0.2z^{-2}$, when a white Gaussian noise, corresponding to an $SNR = 20dB$, is also present. The notation $\bar{y}^i_{[x_k\ x_{k-1}\ x_{k-2}]}$ denotes the ith cluster center $\bar{y}^i$, which is associated with the transmitted symbol sequence $\boldsymbol{x}_i = [x_k\ x_{k-1}\ x_{k-2}]$. The number and the position of the clusters are determined by the length of the CIR and the spread by the power of the noise. Figure 3 illustrates the connection between transitions and cluster

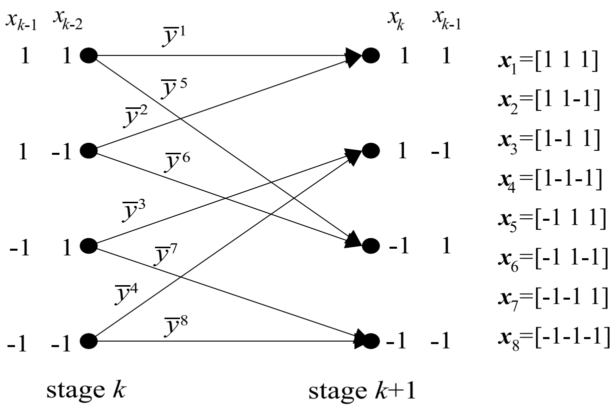

Figure 3: Trellis diagram for a two-tap channel case.

centers in the trellis diagram for a two tap channel. Indeed, each transition defines the vector $\boldsymbol{x}_i$ related to a specific cluster. This observation frees us from the need to know the explicit channel estimate $\boldsymbol{h}$. $\boldsymbol{h}^T\boldsymbol{x}_i$ in eq. (4), is nothing else but the corresponding cluster center $\bar{y}^i$, and the distance metric becomes:

$$D_k^i = |y_k - \bar{y}^i|^2 \tag{5}$$

The resulting technique is equivalent to the MLSE with respect to performance, but it is computationally more efficient due to the fact that the computation of the convolutions (eq. 4) are not required.

Any supervised clustering technique [7] can be used in order to detect the centers (e.g., a simple averaging), based on the known training sequence of symbols. We call the resulting sequence equalizer as the One-Dimensional Clustering Based Sequence Equalizer (1-D CBSE).

5 A NOVEL CENTER ESTIMATION TECHNIQUE

The major drawback of the Cluster Based Sequence Equalizers, as well as of symbol by symbol equalizers which also require cluster center estimation, e.g., [9], [8] is that each cluster has to be represented with a sufficient number of observations in order to achieve accurate estimates of the centers. Consequently, the required training sequence has to be relatively long. Sometimes, in order to alleviate this problem, channel estimation is first perfomed [9].

In this section we propose a novel method for the center detection, which does not require the direct estimation of all the clusters but obtains the estimates of the 2^L cluster centers utilizing the direct estimates *of only L properly selected centers.*

Let us assume a general L-tap channel with impulse response vector $\boldsymbol{h} = [h_0, h_1, \cdots, h_m, \cdots, h_{L-1}]^T$. We define as the tap *contribution* c_x^m, associated with the mth tap h_m, the quantity

$$c_x^m = xh_m, \qquad x \in S = \{+1, \ -1\} \tag{6}$$

In other words, this is the contribution of the h_m tap in the convolution sum in eq. (1). We can observe that c_x^m can take either of 2 different values, depending on the value of the symbol x. We denote these values as c_1^m, c_{-1}^m and is trivial to see that $c_{-1}^m = -c_1^m$. Using this notation, equation eq. (1) can be rewritten as

$$\bar{y}_k = \sum_{m=0}^{L-1} c^m_{x_{k-m}} \equiv \bar{y}_{[x_k, x_{k-1}, \cdots, x_{k-L+1}]} \tag{7}$$

where the notation $\bar{y}_{[x_k, x_{k-1}, \cdots, x_{k-L+1}]}$ is adopted to stress the dependence of the respective cluster center on the transmitted $L-$tuple $[x_k, x_{k-1}, \cdots, x_{k-L+1}]$. Therefore the computation of all the cluster centers requires the estimation of the L tap contributions. Next, we will show, via a series of examples, how we can detect the tap contributions utilizing the estimates of only L properly selected centers.

Example 1: $H(z) = 1$, $(L = 1)$. In this extreme case of a single tap channel, the number of clusters that are formed is $2^1 = 2$. The pair of the cluster centers is illustrated in figure 4. Actually, these

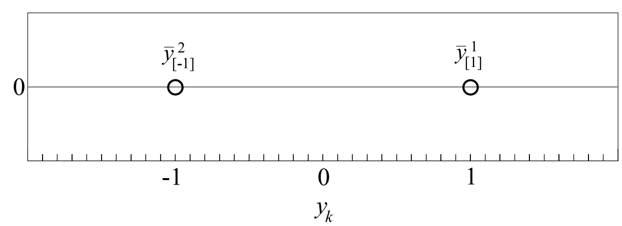

Figure 4: The circles define the cluster centers which correspond to the 1-tap channel $H(z) = 1$

centers coincide with the two possible values that the contribution of the single tap $h_0 = 1$ can take. Therefore, in this case the contribution c_1^0 coincides with the observed center $\bar{y}^1$. Once this has been obtained, the other cluster center is obtained as: $\bar{y}^2 = -c_1^0$. Thus, for $L = 1$, it suffices to estimate only one cluster center.

Example 2: $H(z) = 1 - 0.5z^{-1}$, $(L = 2)$. In this example, a second tap, $h_1 = -0.5$, has been added to the 1-tap channel of the first example. In this case, each one of the centers corresponds to one of the possible 2-symbols combination $[x_k, x_{k-1}]$ leading to the formation of $2^2 = 4$ centers, illustrated in figure 5 by stars. The circles in the figure correspond to the "single tap" centers, i.e. $\bar{y}_{[x]}$, $x = 1$ or $x = -1$. Observe that the clusters for the $L = 2$ system are located symmetrically on either side of $\bar{y}_{[x]}$. That is, one pair is located around $\bar{y}_{[1]}$ and the other around $\bar{y}_{[-1]}$. The distance of each one of the four cluster centers from the respective $\bar{y}_{[1]}$ or $\bar{y}_{[-1]}$ is equal to $\mid c_x^1 \mid$. Exploring this structure of the cluster centers one can find various ways to compute the tap contributions using

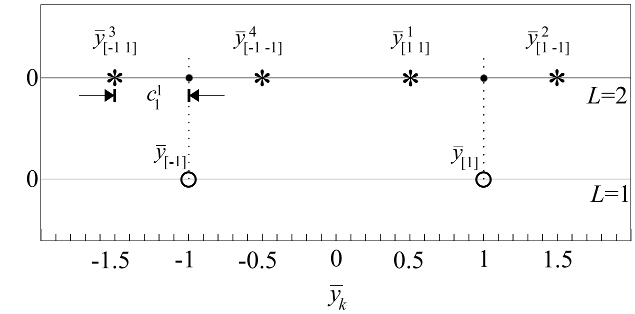

Figure 5: The stars define the cluster centers which correspond to the 2-tap channel $H(z) = 1 - 0.5z^{-1}$

only two ($L = 2$) properly selected cluster centers. For example, it is easy to see that

$$c_1^0 = \frac{\bar{y}^1 + \bar{y}^2}{2}, \quad c_1^1 = \frac{\bar{y}^1 - \bar{y}^2}{2}$$

After the computation of the two contributions the detection of any cluster center is straightforward, e.g., $\bar{y}^3_{[-1\ 1]} = c^0_{-1} + c^1_1 = -c^0_1 + c^1_1$.

Example 3: $H(z) = 1 - 0.5z^{-1} + 0.2z^{-2}$, ($L = 3$). In the same manner, if a third tap is added, say, $h_2 = 0.2$, $2^3 = 8$ clusters are formed with centers grouped in 4 pairs. Each pair is symmetrically located around the 4 centers of the $L = 2$ system. The final structure for this 3-tap channel example is shown in figure 6. Following similar arguments as before, one can see that it suffices to estimate only 3 ($L = 3$) centers in order to compute the tap contributions.

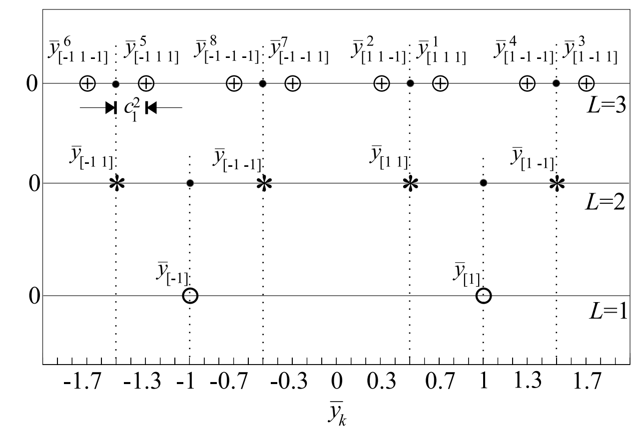

Figure 6: The $\oplus$ definecluster centers which correspond to the 3-tap channel $H(z) = 1 - 0.5z^{-1} + 0.2z^{-2}$

It turns out that in the general case of L taps, it suffices to estimate only L (out of 2^L) properly selected centers.

5.1 COMPUTATION OF THE L CONTRIBUTIONS

Let us define as *basic center* the cluster center $C_{basic} = \bar{y}_{[1,1,\cdots,1]}$, and the associated symbol sequence basic sequence $x_{basic} = [1, 1, \cdots, 1]$. In order to estimate the L tap contributions c_1^m, $0 \leq m \leq L - 1$ the estimation of the following centers are needed.

$$C_0 = \bar{y}_{[-1,1,\cdots,1]},\ C_1 = \bar{y}_{[1,-1,\cdots,1]}, \cdots, C_{L-1} = \bar{y}_{[1,1,\cdots,-1]}$$

The simplest way to estimate a center C_m is by averaging the corresponding observations i.e.,

$$C_m = \frac{1}{N_{C_m}} \sum_{k=1}^{N_{C_m}} y_k^{(C_m)}, \quad 0 \leq m \leq L - 1 \tag{8}$$

where $y_k^{(C_m)}$ is the $k-$th observation belonging to C_m and N_{C_m} is the number of observations associated with C_m.

The C_{basic} can be computed based on the estimations of the L centers C_m following the expression

$$C_{basic} = \frac{\sum_{m=0}^{L-1} C_m}{L - 2}, \quad L > 2. \tag{9}$$

However, for best performance, is better not only to compute C_{basic} via equation 9, but also to directly estimate it like the other L centers applying eq. 8. Then

$$C_{basic} = \left(\frac{\sum_{m=0}^{L-1} C_m}{L - 2} + \hat{y}_{[1,1,\cdots,1]} \right) /2. \tag{10}$$

where $\hat{y}_{[1,1,\cdots,1]}$ is the direct estimation of the basic center.

In order to estimate the required cluster centers directly we need to construct a training sequence in such a way in order to represent the specific clusters with as many observations as possible. It turns out, that a proper training sequence results by repeating the following sequence of symbols successively:

$$[\underbrace{1, 1, \cdots, 1}_{L}, -1] \tag{11}$$

The number of repetitions is constrained by the available training sequence length N_{tr}. For example for a 3-tap channel and $N_{tr} = 8$ symbols, the adopted training sequence is the $[1, 1, 1, -1, 1, 1, 1, -1]$. This sequence corresponds to observations, which are cyclically distributed among the cluster centers C_m and C_{basic}, and leads to two observations per cluster.

It can now be shown that the tap-contributions can be obtained by the $L + 1$ centers

$$c_{1+j}^m = (C_{basic} - C_m)/2, \quad 0 \leq m \leq L - 1 \tag{12}$$

6 COMPUTATIONAL COMPLEXITY REQUIREMENTS AND PERFORMANCE RESULTS

The computational complexity of the new cluster center estimation method is shown in the upper part of the table 1, together with that of the LMS algorithm, in terms of real multiplications and additions. N_{tr} is the number of training symbols. It is very important to point out that the number of multiplications and divisions required by the new method, *is independent of the amount of training symbols* and it is much lower than that required by the LMS algorithm. *Divisions and multiplications are performed once per training block.*

The lower part of the table refers to the complexity of the Viterbi algorithm. The more consuming part of the 1-D CBSE equalizer is the computation of the Euclidean distance metric which is given by equation eq. (5), which can be computed based on equation eq. (7). The overall complexity of the equalizer in terms of real operations is shown in table 1 together with the complexity of the Bayesian DFE, and the complexity required by the standard form of the MLSE equalizer based on explicit modeling. The delay used for the Bayesian DFE is the optimum $d = L - 1$ [8].

In the first set of experiments, the convergence speed of the new center estimation method was studied (in terms of the required number of training symbols) and compared with that of the LMS algorithm in various noise levels. In order to have a statistically more representative result, we performed the experiments using 1000 different 5-tap channels and the reported results are the obtained mean values.

Figure 7 summarize the results for two SNR levels, 30 and 10 dB. The plotted quantity is the mean total deviation between the 2^5 true and the corresponding estimated cluster centers.

We can easily see that the new method exhibits faster convergence compared to the LMS algorithm. This is clearer for the case of high signal to noise ratio (SNR=30dB), where the new method has converged in less than ten received samples, in contrast to the LMS which needs approximately 20 ($5L$) samples. The training

	Method	Mul / Div	Add / Sub	$(\cdot)^2$	$exp(\cdot)$
Center estimation method	LMS	$N_{tr}\cdot(2\cdot L+1)$	$N_{tr}\cdot 2\cdot L$	0	0
	New CE	$2L+1$	$N_{tr}+L$	0	0
Equalization	Bayesian DFE	2^L	$2\cdot L\cdot 2^L-2$	$L\cdot 2^L$	2^L
	MLSE	$L\cdot 2^L$	$2^L(L+1)$	2^L	0
	1-D CBSE	0	$2^L(L+1)$	2^L	0

Table 1: The computational complexity of the Bayesian-DFE, the standard MLSE and the 1-D CBSE.

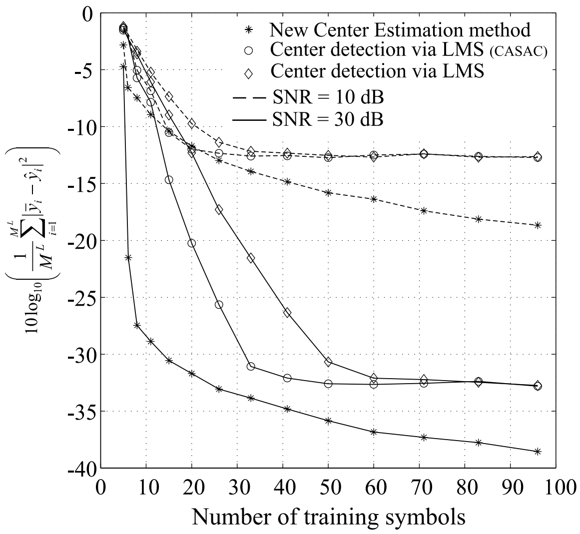

Figure 7: MSE between real and estimated cluster centers for various training sequence lengths averaged over 1000 5-tap channels.

sequence for the LMS algorithm was a repeated CAZAC (Constant Amplitude Zero Autocorrelation) sequence [11]. For other training sequences, as it can be seen from figure 7, (curves indicated by $\diamond$), the performance of the LMS algorithm was degraded substantially.

In the sequel, the symbol error rate (SER) performance of the 1-D CBSE, employing the new center estimation (CE) technique, is treated and compared to the Bayesian-RBF employing both the new CE method and the LMS algorithm for CIR estimation. The transmission was realized in blocks, where each block comprised 200 data symbols together with a number of training symbols placed in the front of each block. Figure 8 shows the performance curves of the tested equalizers for the case of a 5-tap channel with transfer function $H(z) = 0.227 + 0.466z^{-1} + 0.688z^{-1} + 0.466z^{-1} + 0.227$ when the training sequence is 10 and 30 symbols long. It is readily seen than the proposed equalizer out-performs the Bayesian-DFE not only when the center estimates are obtained via the channel estimation, using the LMS algorithm, but also when utilizes the cluster-centers obtained by the new (CE) method.

7 CONCLUSIONS

A novel cluster based MLSE technique has been presented. The new method offers substantial computational savings compared to previously proposed cluster based methods. The enhanced performance of the new method compared with standard LMS based MLSE and Bayesian RBF equalizers has been established via simulations.

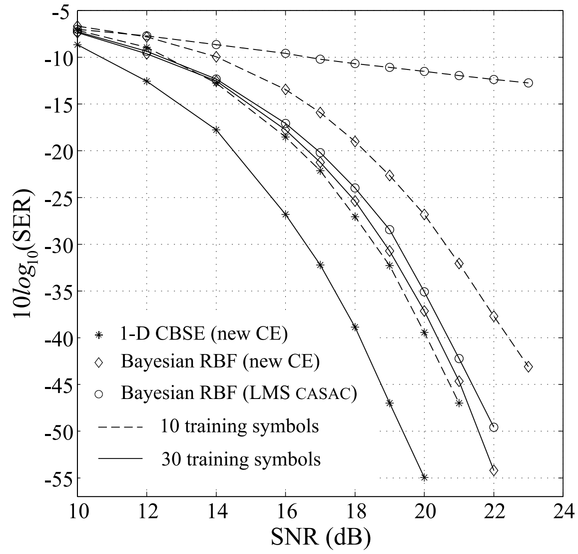

Figure 8: Symbol error rate performance for 10 and 30 training symbols per data block. The transfer function of the channel is $H(z) = 0.227 + 0.466z^{-1} + 0.688z^{-2} + 0.466z^{-3} + 0.227z^{-4}$

References

[1] John G. Proakis, "Digital communications, 3rd ed.". McGraw-Hill International Editions, 1995.

[2] G. D. Forney, "Maximum-Likelihood sequence estimation of digital sequences in the presence of intersymbol interference", *IEEE Trans. on Information Theory*, vol. IT-18, no. 3, pp. 363-378, May 1972.

[3] S. Theodoridis, C.F.N. Cowan, C.P. Callender, C.M.S. See "Schemes for Equalization of Communications Channels with Nonlinear Impairments", *IEE Proc. Communications*, vol. 142, no. 3, pp. 165-171, Jun. 1995.

[4] K. Georgoulakis, S. Theodoridis "Efficient Clustering Techniques for Channel Equalization in Hostile Environments" *Signal Processing*, Vol. 58, pp. 153-164, 1997.

[5] K. Georgoulakis, S. Theodoridis "Channel Equalization for Coded Signals in Hostile Environments" *IEEE Trans. on Signal Processing*, Vol. 47, pp. 1783-1787, June 1999.

[6] Y. Kopsinis, S. Theodoridis "Reduced-Complexity clustering techniques for nonlinear channel equalization", *Proceedings of WCC 2000(ICSP)*, Beijing, China.

[7] S. Theodoridis, K. Koutroumbas "Pattern Recognition" *Academic Press*, 1998.

[8] S. Chen, B. Mulgrew, S. McLaughlin "Adaptive Bayesian Equalizer with Decision Feedback" *IEEE Trans. on Signal Processing*, vol.41, no. 9, Sep. 1993.

[9] S. Chen, B. Mulgrew, P. M. Grant "A Clustering Technique for Digital Communications Channel Equalization Using Radial Basis Function Networks", *IEEE Trans. on Neural Networks*, vol. 4, no. 4, Jul. 1993.

[10] S. Haykin, "Adaptive Filter Theory, 3rd ed.". Prentice-Hall, 1996.

[11] IEEE 802.16.1 standart for fixed wireless access LMDS (local multipoint distribution systems).