AN OPTIMAL MMSE FUZZY PREDICTOR FOR SISO AND MIMO BLIND EQUALIZATION

Rafael Ferrari, Ricardo Suyama, Renato R. Lopes, Romis R. F. Attux, João M. T. Romano

FEEC – State University of Campinas – CP 6101 – CEP: 13083-970 – Campinas – SP – Brazil
E-mails: {rferrari, rsuyama, rlopes, romano}@decom.fee.unicamp.br, attux@dca.fee.unicamp.br

ABSTRACT

The present work deals with the research of optimal solutions in unsupervised and nonlinear signal processing. The proposed framework is based on nonlinear prediction, to be implemented by a fuzzy filter structure. Our main result consists in establishing the optimality of the approach by showing the equivalence between the minimum mean squared error estimator and the fuzzy predictor. The result is then applied in the contexts of SISO equalization and convolutive source separation (MIMO equalization). We also propose a strategy for the updating of the unsupervised nonlinear equalizer. Simulation results confirm the effectiveness of the proposal in SISO and MIMO scenarios.

Index Terms— blind equalization, blind source separation, fuzzy filter, prediction-error filter.

1. INTRODUCTION

Over the last decades, classical approaches in statistical signal processing have encountered a significant success in several applications, especially due to their well-established concepts as well as the efficiency and simplicity of a number of algorithms. On the other hand, it must be noted that the solution of some recent and challenging problems is rather limited, even unfeasible, if classical assumptions are considered. As a consequence, advanced researches in signal processing have been characterized by the adoption of some new theoretical frameworks. In particular we can mention a sort of tripod formed by the growing interest on nonlinear processing, the disuse of the classical Gaussian hypothesis and the consequent use of higher order statistics [1] [2].

Both classical and advanced approaches seek a fundamental purpose in signal processing techniques, that is, to perform optimally. About sixty years ago, such objective was attained by Kolmogorov and Wiener, by using the classical linearity and Gaussianity hypotheses together with linear algebra and functional analysis tools. However, this fundamental task faces increasing difficulties if classical assumptions are discarded.

The present work deals with the research of optimal solutions in unsupervised and nonlinear signal processing. Such problem is present in important applications, like channel equalization and identification, source separation, data clustering and others. Our proposed framework is based on nonlinear prediction, to be implemented by a fuzzy filter structure. Our main result consists in establishing the optimality of the approach by showing the equivalence between the minimum mean squared error (MMSE) estimator and the fuzzy predictor. The result is then applied in the contexts of single-input/single-output (SISO) equalization and convolutive source separation (multiple-input/multiple-output (MIMO) equalization). Also, a technique for both SISO and MIMO equalizers adaptation is proposed.

This paper is organized as follows. In section 2, we present a discussion on unsupervised and nonlinear equalization. The purpose of this section is to render evident that the fundamental results on unsupervised signal processing are not applicable when nonlinear devices are used; so we present an alternative approach based on nonlinear prediction. The optimal MMSE SISO predictor, according to the estimation theory, is derived in section 3. Then, in section 4, we introduce the fuzzy prediction- error filter and prove its optimality by demonstrating its equivalence with the MMSE solution. As a final contribution, an updating strategy is proposed in section 5; it combines an iterated local search (ILS) clustering algorithm with a classical recursive least squares (RLS) algorithm. Section 6 brings some simulation results to evaluate the proposed method in the particular case of SISO equalization; comparisons with the Bayesian optimal criterion are provided. In section 7, the proposed approach is extended to the case of convolutive mixtures, or rather MIMO equalization, and simulation results in that context are provided. Finally, our conclusions are stated in section 8.

2. UNSUPERVISED AND NONLINEAR EQUALIZATION

Fig. 1 illustrates a convolutive mixture scheme followed by a separation and deconvolution device.

In the present work we assume the input signals $s_i(k)$ to have the same statistical distributions as well as to belong to a finite alphabet. Under these assumptions, the system depicted in Fig. 1 corresponds to a MIMO transmission and

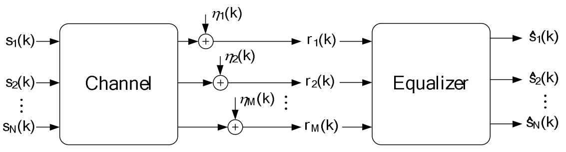


Figure 1: Convolutive mixture and equalization.

equalization one. The channel is considered to be linear, so that the received signals can be expressed as:

$$\mathbf{r}(k)=\sum_{i=0}^{n_c-1}\mathbf{H}_i\mathbf{s}(k-i)+\boldsymbol{\eta}(k), \cdot \qquad (1)$$

where $\mathbf{H}_i$ denotes the mixing matrix, with dimensions MxN, associated with time instant i, $\mathbf{s}(k)$ represents the vector with the sources, $\boldsymbol{\eta}(k)$ the noise vector and n_c is the channel memory. The simpler case of SISO equalization holds when $M=N=1$.

The use of unsupervised (or blind) strategies is motivated by the well-known systemic advantages in avoiding supervision, while the interest in nonlinear equalizers is due to their abilities in dealing with severe channels and preventing noise enhancement.

However, there is a twofold arduousness in handling with unsupervised and nonlinear optimization. Firstly, the general theoretical framework of unsupervised signal processing assumes linear mixture (channels) and deconvolution (equalizers). Secondly, for nonlinear devices, optimality is provided by the Bayesian solution, the practical implementation of which assumes a supervised training period. That is, besides the classical Wiener-based approach, it can be noted that the most usual and analytically tractable solutions concern either linear filters in unsupervised mode or else nonlinear structures adapted with the aid of a pilot signal. Such reasoning will be detailed in the sequel.

2.1. Fundamental results on unsupervised signal processing

A seldom discussed but rather clear statement is that blind source separations (BSS) of convolutive mixtures can be viewed as the most general problem in unsupervised signal processing. Such problem can be particularized to MIMO equalization, to multiuser processing [3] and even to SISO equalization, where the channel performs only convolution and no mixtures take place.

It is well known that a solid theoretical framework for BSS is provided by independent component analysis (ICA) [4]. In addition, Taleb and Jutten [5] have shown that ICA is valid when the whole scheme in Fig. 1 constitutes a linear system, but not necessarily in the general nonlinear case. In this case, independence between the output signals does not guarantee a correct retrieval of the original sources.

It is curious to show that an analogous result can be obtained for the particular case of SISO blind equalization. In this case, the theoretical framework comes from the well-known Benveniste-Goursat-Rouget (BGR) [6] and Shalvi-Weinstein (SW) [7] theorems. Now, these results assume that the channel and the equalizer constitute a linear system. However, nonlinear equalizers can provide equality between the distribution of the transmitted and received signals, even if the recovered symbols do not correspond to the transmitted ones. This is simply demonstrated by the following example.

Example 1: Let us consider the equalizer in Fig. 1 as a nonlinear device, the input vector of which is given by $\mathbf{r}_f(k)=[\mathbf{r}^T(k)\ \ldots\ \mathbf{r}^T(k-m-1)]^T$. The channel states are defined as the possible values that $\mathbf{r}_f(k)$ can assume in the absence of noise. Mathematically, the channel states vector is given by:

$$\mathbf{c}_j=E\left[\mathbf{r}_f(k)\,|\,\tilde{\mathbf{s}}_j(k)\right] \ . \qquad (2),$$

where $\tilde{\mathbf{s}}_j(k)=\left[\mathbf{s}^T(k)\ \vdots\ \cdots\ \vdots\ \mathbf{s}^T(k-n_c-m-1)\right]^T$ denotes a possible combination of transmitted symbols. Let us now consider the particular case of a SISO channel with transfer function $H(z)=0.5+z^{-1}$ and a nonlinear equalizer with input vector given by $\mathbf{r}_f(k)=[r(k)\ r(k\text{-}1)]^T$. Assuming binary transmission, there are 8 possible channel states, each one corresponding to a possible triple $[s(k)\ s(k\text{-}1)\ s(k\text{-}2)]^T$. The role of a nonlinear equalizer is mapping each channel state into a recovered symbol $f(k)$. The Bayesian equalizer provides optimal mapping in that it maximizes the probability of the correct recovering of a transmitted symbol $s(k)$. Concerning the present example, Table 1 shows all the channel states, and two possible mappings, $f_1(k)$ and $f_2(k)$, provided by nonlinear structures.

Table 1: Channel states, transmitted signals, equalizer inputs and recovered symbols of Example 1.

$\mathbf{c}_j$	$s(k)$	$s(k\text{-}1)$	$s(k\text{-}2)$	$\mathbf{r}_f^T(k)$		$f_1(k)$	$f_2(k\text{-}1)$
				$r(k)$	$r(k\text{-}1)$		
$\mathbf{c}_1$	1	1	1	1.5	1.5	1	-1
$\mathbf{c}_2$	1	1	-1	1.5	-0.5	1	1
$\mathbf{c}_3$	1	-1	1	-0.5	0.5	1	1
$\mathbf{c}_4$	1	-1	-1	-0.5	-1.5	1	-1
$\mathbf{c}_5$	-1	1	1	0.5	1.5	-1	1
$\mathbf{c}_6$	-1	1	-1	0.5	-0.5	-1	-1
$\mathbf{c}_7$	-1	-1	1	-1.5	0.5	-1	-1
$\mathbf{c}_8$	-1	-1	-1	-1.5	-1.5	-1	1

Since the channel states are equiprobable, both $f_1(k)$ and $f_2(k)$ have the same statistical distribution, which is also the same of the original transmitted signal, $s(k)$. However, the first one corresponds to the output of the Bayesian equalizer, which recovers the transmitted signal, while the second equalizer does not provide correct retrieval for any considered delay.

Hence, the theoretical foundations in BSS and blind equalization, as well as their resulting algorithms, can not be generally applied to provide optimal solutions in the context of nonlinear processing. Then, our next task is the search of other classes of methods.

2.2 - Nonlinear prediction and equalization

The concept of prediction is present in many applications of signal processing techniques [1]. It consists in estimating a given sample $x(k)$, based on the information provided by a set of available samples. This idea can be expressed by:

$$\hat{x}(k) = P[\boldsymbol{\chi}(k)]. \qquad (2)$$

where $\hat{x}(k)$ is the predicted signal. If the vector $\boldsymbol{\chi}(k)$ contains past samples $x(k\text{-}i)$, i = 1, 2, ..., K, the operation is called forward prediction. Backward prediction corresponds to estimate the sample $x(k)$ from a set of posterior samples, while using past and posterior available data to estimate a given sample correspond to the problem of interpolation. In all cases, P[.] express the mapping performed by the predictor and the prediction error is defined by:

$$e_P(k) = x(k) - \hat{x}(k) = x(k) - P[\boldsymbol{\chi}(k)] \ . \qquad (3)$$

The design of the optimal predictor can be carried out by means of the minimization of the mean square prediction error. In the particular case of linear prediction, P[.] is implemented by a linear FIR filter and its coefficients can be calculated by the Wiener procedure.

A number of previous works has established interesting relationships between prediction and equalization. In the SISO context, it is known that a linear forward prediction-error filter is able to equalize minimum-phase channels [1], while linear backward prediction-error filters are effective only for maximum-phase ones [1]. These properties motivated a series of works dealing with cascaded prediction-error equalizers, with linear [8] and nonlinear [9] configurations. The basic idea is illustrated in Fig. 2.

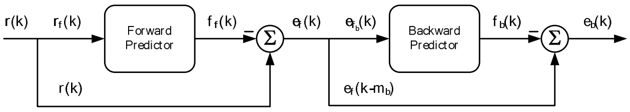


Figure 2: Cascaded prediction-error equalizer.

Linear prediction principles have also been investigated in many important works on SIMO and MIMO equalization, as [10] and [11] to mention a few. In this case, the minimum-phase restriction is weakened to one about the absence of common zeros between the subchannels.

However, Cavalcante et al. [12] showed that the restrictions over the phase response of the channel can be overcome by using nonlinear prediction. Moreover, when considering a cascade configuration, the nonlinear predictor can further improve the efficacy of the equalizer. This result provided motivation to the present work in order to look for an optimal solution for unsupervised nonlinear equalization based on the prediction-error principle.

3. MINIMUM MEAN SQUARED ERROR PREDICTOR

Let us consider a SISO predictor in the present section. According to estimation theory, the MMSE estimator for $r(k)$ given the vector of past samples $\mathbf{r}_f(k) = [r(k\text{-}d_f\text{-}1) \ldots r(k\text{-}m\text{-}1)]^T$, i.e., the (d_f+1)-step forward predictor, is given by the conditional mean:

$$f_{MMSE}(\mathbf{r}_f(k)) = E[r(k) \mid \mathbf{r}_f(k)] \ . \qquad (5)$$

Hence, in order to determine the MMSE estimator, it is necessary to obtain the conditional distribution $p(r(k)|\mathbf{r}_f(k))$, or equivalently, using Bayes theorem, the joint distribution of $p(r(k), \mathbf{r}_f(k))$ as well as $p(\mathbf{r}_f(k))$.

The distribution of $\mathbf{r}_f(k)$ can be expressed by:

$$p(\mathbf{r}_f) = \sum_{j=1}^{S^{m+n_c-d_f-2}} p(\mathbf{r}_f \mid \mathbf{c}_{a_j}) P(\mathbf{c}_{a_j}), \qquad (6)$$

where S is the cadinality of the transmitted signal alphabet. The time index is ommited to simplify the notation. In (6), $p(\mathbf{r}_f|\mathbf{c}_{aj})$ denotes the conditional distribution of $\mathbf{r}_f$ given the channel state $\mathbf{c}_{aj}$ with dimension $(m\text{-}d_f\text{-}1)$. Due to the presence of additive gaussian noise, $p(\mathbf{r}_f|\mathbf{c}_{aj})$ will be a normal distribution centered at $\mathbf{c}_{aj}$, i.e.,

$$p(\mathbf{r}_f \mid \mathbf{c}_{a_j}) = (2\pi\sigma_\eta^2)^{\frac{-(m-d_f-1)}{2}} \exp\left(-\frac{\left\|\mathbf{r}_f - \mathbf{c}_{a_j}\right\|^2}{2\sigma_\eta^2}\right). \qquad (7)$$

Since $\mathbf{c}_{aj}$ are equiprobable vectors, with probablity $S^{-(m+n_c-d_f-2)}$, substituting (6) in (7) yields:

$$p(\mathbf{r}_f) = \frac{(2\pi\sigma_\eta^2)^{\frac{-(m-d_f-1)}{2}}}{S^{m+n_c-d_f-2}} \sum_{j=1}^{(S^{m+n_c-d_f-2})} \exp\left(-\frac{\left\|\mathbf{r}_f - \mathbf{c}_{a_j}\right\|^2}{2\sigma_\eta^2}\right). \qquad (8)$$

Assume that $d_f<(n_c\text{-}1)$ so that r and $\mathbf{r}_f$ are correlated and hence r can be predicted from $\mathbf{r}_f$. Now, using total probability and the fact that r and $\mathbf{r}_f$ are conditionally independent given $\mathbf{c}_j$, $p(r, \mathbf{r}_f)$ is given by

$$\begin{aligned} p(r, \mathbf{r}_f) &= \sum_{j=1}^{(S^{m+n_c-1})} p(r, \mathbf{r}_f \mid \mathbf{c}_j) \cdot P(\mathbf{c}_j) \\ &= \sum_{j=1}^{(S^{m+n_c-1})} p(r \mid \mathbf{c}_j) \cdot p(\mathbf{r}_f \mid \mathbf{c}_j) \cdot P(\mathbf{c}_j) \end{aligned} , \qquad (9)$$

where $\mathbf{c}_j$ is a channel estate vector of dimension m (note that $\mathbf{c}_{aj}$ is only a part of $\mathbf{c}_j$). Using (7), it is possible to notice that

$$p(r \mid \mathbf{c}_j) = (2\pi\sigma_\eta^2)^{-\frac{1}{2}} \cdot \exp\left(-\frac{\left\|r - c_{j,0}\right\|^2}{2\sigma_\eta^2}\right), \qquad (10)$$

and

$$p(\mathbf{r}_f \mid \mathbf{c}_j) = (2\pi\sigma_\eta^2)^{-\frac{(m-d_f-1)}{2}} \cdot \exp\left(-\frac{\|\mathbf{r}_f - \hat{\mathbf{c}}_j\|^2}{2\sigma_\eta^2}\right), \quad (11)$$

where $\hat{\mathbf{c}}_j = [c_{j,d_f+1}\ c_{j,d_f+2} \cdots c_{j,m-1}]^T$. Again, due to the fact that the states are equiprobable, i.e., $P(\mathbf{c}_j)=S^{-(m+n_c-1)}$, substituting (10) and (11) in (9), yields

$$p(r,\mathbf{r}_f) = \frac{(2\pi\sigma_\eta^2)^{-\frac{(m-d_f)}{2}}}{S^{m+n_c-1}} \sum_{j=1}^{(S^{m+n_c-1})} \exp\left(-\frac{|r-c_{j,0}|^2}{2\sigma_\eta^2}\right) \cdot \exp\left(-\frac{\|\mathbf{r}_f - \hat{\mathbf{c}}_j\|^2}{2\sigma_\eta^2}\right). \quad (12)$$

Dividing (12) by (8) we finally obtain:

$$p(r \mid \mathbf{r}_f) = \frac{(2\pi\sigma_\eta^2)^{-\frac{1}{2}}}{S^{d_f+1}} \cdot \frac{\sum_{j=1}^{(S^{m+n_c-1})} \exp\left(-\frac{|r-c_{j,0}|^2}{2\sigma_\eta^2}\right) \cdot \exp\left(-\frac{\|\mathbf{r}_f - \hat{\mathbf{c}}_j\|^2}{2\sigma_\eta^2}\right)}{\sum_{j=1}^{(S^{m+n_c-d_f-2})} \exp\left(-\frac{\|\mathbf{r}_f - \mathbf{c}_{a_j}\|^2}{2\sigma_\eta^2}\right)}, \quad (13)$$

and using this result in (5), it is possible to obtain the following expression for the MMSE estimator:

$$f_{MMSE}(\mathbf{r}_f) = \int_{-\infty}^{\infty} r \cdot p(r \mid \mathbf{r}_f)\,dr = \frac{\sum_{j=1}^{(S^{m+n_c-d_f-2})} w_j \cdot \exp\left(-\frac{\|\mathbf{r}_f(k) - \mathbf{c}_{a_j}\|^2}{2\sigma_\eta^2}\right)}{\sum_{j=1}^{(S^{m+n_c-d_f-2})} \exp\left(-\frac{\|\mathbf{r}_f(k) - \mathbf{c}_{a_j}\|^2}{2\sigma_\eta^2}\right)} \quad (14)$$

where

$$w_j = \frac{1}{S^{d_f+1}} \sum_{n=1}^{S^{d_f+1}} c_{n,0}\,, \quad (15)$$

with $\hat{\mathbf{c}}_n = \mathbf{c}_{a_j}$.

The derivation of the MMSE backward predictor follows an analogous proof. Now we look for a suitable nonlinear structure to be associated to the optimal forward predictor in (14), and then to the analogous backward predictor, so that a nonlinear equalizer can be implemented in accordance with the idea depicted in Fig. 2.

Our investigations have been oriented to the fuzzy structures, which were successfully applied in previous works [13] [14] in different equalization problems. The aim of the following section is to use the fuzzy filter structure to implement a predictor-based equalizer and demonstrate the optimality of this approach.

4. THE MMSE FUZZY PREDICTOR

A fuzzy filter is a nonlinear filtering structure endowed with universal approximation capability. It is capable of processing information in conformity with a basis of logical rules that employ non-binary membership functions (fuzzy sets) [15]. If these membership functions are Gaussian, product is the inference operator, and centroid defuzzification is utilized, it is possible to obtain the following input-output mapping:

$$y = \frac{\sum_{l=1}^{N_r} w_l \prod_{j=1}^{m} \exp\left(-\frac{|x_j - c_{j,l}|^2}{2\sigma_{j,l}^2}\right)}{\sum_{l=1}^{N_r} \prod_{j=1}^{m} \exp\left(-\frac{|x_j - c_{j,l}|^2}{2\sigma_{j,l}^2}\right)} \quad (16),$$

where N_r and m are respectively the numbers of rules and inputs of the fuzzy system, $c_{j,l}$ and $\sigma^2_{j,l}$ are respectively the centers and variances of the Gaussian membership functions, and w_l are the output weights.

By comparing equations (14) and (16) and expanding the norm in (14) in a product of exponentials, it can be observed that the input-output mappings of the MMSE predictor and the fuzzy filter have a structural equivalence. Therefore, it is possible to obtain the MMSE predictor using a fuzzy filter by properly setting its parameters.

To obtain the MMSE (d_f+1)-step forward predictor, the parameters of the fuzzy filter should be set according to (14), so that the input-output mapping of the fuzzy filter is

$$f_f(k) = \frac{\sum_{j=1}^{N_a} w_j \prod_{l=0}^{m-d_f-2} \exp\left(-\frac{(r(k-d_f-l) - c_{a_{j,l}})^2}{2\sigma_\eta^2}\right)}{\sum_{j=1}^{N_a} \prod_{l=0}^{m-d_f-2} \exp\left(-\frac{(r(k-d_f-l) - c_{a_{j,l}})^2}{2\sigma_\eta^2}\right)} \quad (17)$$

where the weights w_j are given by (15).

It can be shown that the one step backward MMSE predictor is equivalent to the fuzzy filter of mapping

$$f_b(k) = \frac{\sum_{j=1}^{N_a} w_j \prod_{l=0}^{m_b-1} \exp\left(-\frac{(r(k-l) - c_{a_{j,l}})^2}{2\sigma_\eta^2}\right)}{\sum_{j=1}^{N_a} \prod_{l=0}^{m_b-1} \exp\left(-\frac{(r(k-l) - c_{a_{j,l}})^2}{2\sigma_\eta^2}\right)} \quad (18)$$

where m_b is the number of inputs of the filter and the weights are given by

$$w_j = \frac{1}{S} \sum_{n=1}^{S} c_{n,m_b}\,, \quad (19)$$

with $[c_{n,0}\ \ c_{n,1} \ldots c_{n,mb-1}] = \mathbf{c}_{aj}$.

The cascaded prediction-error equalizer can be obtained by combining a fuzzy forward predictor with a backward one, so that the output of the forward prediction error equalizer be used as the input of the backward prediction error equalizer.

5. TRAINING ALGORITHM

In this section, we discuss the training procedure used to obtain the parameters of the fuzzy predictors. In (17) and (18), it is noticeable that the free parameters are the channel states, $\mathbf{c}_{aj}$, which are the centers of the Gaussian membership functions; the noise variance, σ_η^2; and the output weights, w_j. The training process may be divided into two distinct stages: estimating the channel states and noise variance, and finding the adequate set of output weights.

The problem of estimating the channel states from the received data is fundamentally an unsupervised clustering problem. To perform this task, we used the ILS algorithm [16]. This algorithm combines an evolutionary-based global search and a local search strategy, performed in our implementation by the k-means algorithm [16]. The conjunction of these features confers to the ILS a good balance between exploration and exploitation of the search space, which constitutes an essential feature for the correct estimation of the channel states.

Once we have the channel states, it is straightforward to evaluate the dispersion of the data around them, and therefore obtain the noise variance.

Given the centers and noise variance, the problem of determining the weights of the fuzzy predictor becomes linear in the parameters, which allows us to resort to a vast amount of tools and results belonging to the classical adaptive filtering framework. In our implementation, we use an RLS algorithm [14] to adapt the weights of the fuzzy predictor.

The training procedure of the fuzzy predictor can therefore be summarized in three steps: i) channel states estimation using ILS; ii) noise variance estimation by evaluating the data dispersion around the states; iii) adaptation of the output weights using the RLS algorithm. For the cascaded prediction-error equalizer, the training is performed first for the forward predictor and then repeated for the backward predictor using the forward prediction error signal.

6. SISO CHANNEL EQUALIZATION

In this section, we present simulation results to assess the performance of the proposed prediction-error equalizers for SISO equalization. The transmitted signal was supposed to belong to a binary alphabet, {-1;+1}, and the transfer function of the channel is $H(z) = 0.5 + 0.7z^{-1} + 0.5z^{-2}$. Since the second coefficient is the most significant, a cascaded prediction-error equalizer with $m = 5$, $m_b = 4$ and $d_f = 1$ was used. To train the fuzzy predictors, 8000 samples of the received signal were used. We compare the results of the proposal with three other solutions: the MMSE FIR linear equalizer, the Bayesian equalizer and the linear cascaded-prediction error equalizer. We used an FIR filter with 8 coefficients to implement the linear MMSE equalizer, since this is the number of samples of the received signal necessary to obtain one output of the cascaded prediction-error equalizer. Its equalization delay was set to one to obtain the best possible performance. The number of inputs of the Bayesian equalizer was set to 4 so that its overall computational complexity is equivalent to that of the proposed cascaded prediction-error equalizer; its equalization delay was set to one for best performance. The linear FIR predictors used in the linear cascaded-prediction error equalizer represent the MMSE linear solution.

Fig. 3 shows the average bit error rate (BER) curves of the equalizers. It can be observed that the performance of the cascaded-prediction error equalizer, using the MMSE predictors, is close to that of the Bayesian equalizer. The same occurs for the trained cascaded prediction-error equalizer using fuzzy predictors, which confirms the good performance of the proposed training strategy. The BER of the linear cascaded prediction-error equalizer exceeds 0.1, even for high values of SNR, which renders evident the unsuitableness of this structure in the present case. The linear equalizer performs slightly better than the nonlinear cascaded prediction-error equalizer for low values of SNR, but it decreases significantly for SNR values above 15dB.

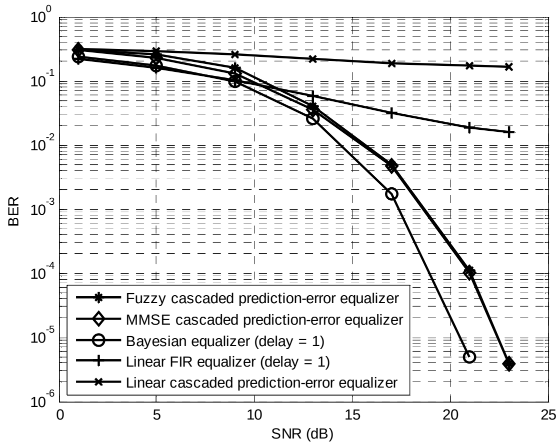


Figure 3: BER curves for different equalization approaches and channel $H(z) = 0.5 + 0.7z^{-1} + 0.5z^{-2}$.

7. EXTENSION TO SOURCE SEPARATION (MIMO EQUALIZATION)

The proposed approach can be extended to the problem of MIMO channel equalization, with M=N>1. The idea is depicted in Fig. 4, where a set of forward prediction-error filters is employed.

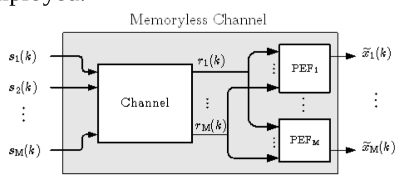


Figure 4: Prediction-error filter approach for MIMO channels.

Let the input of the prediction filter be composed of past samples of all received signals, i.e.,

$$\mathbf{r}_f = [r_1(k), \cdots, r_1(k-m) \mid \cdots \mid r_M(k), \cdots, r_M(k-m)]^T . \quad (20)$$

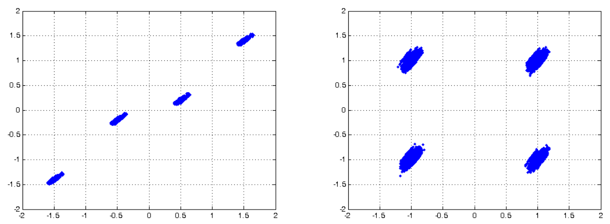
The MMSE estimator is also given by (5), and its derivation follows the steps in (6)-(15), with a few differences regarding the number of possible channel states, which depends on the number of sources and can be considerably higher than in the SISO scenario. Therefore, the same training method explained in Section 5 is also effective for the MIMO case.

However, the main difference is that the nonlinear filters will only be able to eliminate the convolutive aspect of the channel [17]. In other words, the outputs will be equivalent to those of a memoryless system, also known as an instantaneous mixing system. In this context, to recover the original signals, one can rely on blind source separation tools, as the independent component analysis [4]. For this task, several algorithms have been proposed in the literature. In this work, we employ the FastICA algorithm [18].

Fig. 5 illustrates the results obtained with the nonlinear prediction approach. For this example, a system determined by two square matrices ($\mathbf{A}_0$, $\mathbf{A}_1$) was considered, and the two sources were drawn from a BSPK constellation.

Fig. 5(a) shows the joint distribution of the prediction errors. It is possible to notice that the points are gathered around the vertices of a parallelogram, which corresponds to the action of an instantaneous mixture on the original distribution of the sources. This indicates that the prediction-error filters were able to eliminate the convolutive aspect of the channel.

The next step was to apply the FastICA algorithm on the error signals. The distribution of the resulting signals is depicted in Fig. 5(b), where it can be seen that the original distribution was recovered.



(a) (b)

Figure 5: Joint distribution of (a) the prediction errors, representing a channel without memory, and (b) the signals obtained after applying FastICA.

8. CONCLUSIONS

Many recent researches in signal processing have been characterized by the adoption of some new theoretical frameworks. In this context, unsupervised and nonlinear signal processing becomes one of the major focuses of interest. The present paper aims to contribute with the field by proposing the MMSE fuzzy predictor as an optimal solution for blind nonlinear equalization in both SISO and MIMO scenarios. An updating strategy was also proposed. To perform optimally in an unsupervised and nonlinear context may be an important acquisition for different applications in modern signal processing. This work intends to show a first set of results; some theoretical aspects as well as other potential applications can still be explored.

9. REFERENCES

[1] S. Haykin, *Adaptive Filter Theory*, 3rd Ed., Prentice Hall, 1996.

[2] D. Erdogmus, J. C. Principe, "From Linear Adaptive Filtering to Nonlinear Information Processing," *IEEE Signal Processing Magazine*, vol. 23, no. 6, pp. 14–33, Nov 2006.

[3] C. C. Cavalcante, D. Zanatta Filho and J. M. T. Romano. "Multiuser Processing Using Blind Source Separation Methods," *13th European Wireless Conference* (EW2007), Paris, April 2007.

[4] P. Comon, "Independent component analysis, a new concept?," *Signal Processing*, vol. 36 no. 3, pp. 287–314, April 1994.

[5] A. Taleb and C. Jutten, "Source separation in post-nonlinear mixtures," *IEEE Transactions on Signal Processing*, vol. 47, no. 10, pp. 2807–2820, 1999.

[6] A. Benveniste, M. Goursat and G. Ruget, "Robust identification of a nonminimum phase system: Blind adjustment of a linear equalizer in data communications," *IEEE Trans. on Automatic Control*, vol. 25, no. 3, pp. 385–399, June 1980.

[7] O. Shalvi and E. Weinstein, "New Criteria for Blind Deconvolution of Nonminimum Phase Systems (Channels)," *IEEE Trans. on Information Theory*, vol. 36, no. 2, pp. 312–321, 1990.

[8] O. Macchi and A. Hachicha, "Self-Adaptive Equalization Based on a Prediction Principle," *Proc. IEEE Globecom'86*, pp. 1641–1645, 1986.

[9] C. A. F Da Rocha, O. Macchi and J. M. T. Romano, "An adaptive nonlinear IIR filter for self-learning equalization," *ITC 94*, Rio de Janeiro, pp. 6–10, Brazil, 1994.

[10] J. K. Tugnait, "On linear predictors for MIMO channels and related blind identification and equalization," *IEEE Signal Processing Letters*, vol. 5, no. 11, pp. 289–291, 1998.

[11] D. T. M. Slock, A. Meraim, P. Duhamel, D. Gesbert, P. Loubaton, S. Mayrargue and E. Moulines, "Prediction error methods for time-domain blind identification of multichannel FIR filters," ICASSP'95, pp 1968–1971, Detroit, USA, 1995.

[12] C. C. Cavalcante, J. R. Montalvão Filho, B. Dorizzi and J. C. M. Mota, "A Neural Predictor for Blind Equalization in Digital Communication: Is It Plausible?," *Proceedings of IEEE Neural Networks for Signal Processing*, Sydney, Australia, 2000.

[13] L. Wang, J. M. Mendel, "Fuzzy Adaptive Filters, with Application to Nonlinear Channel Equalization", *IEEE Trans. on Fuzzy Systems*, vol. 1, no. 3, pp. 161–170, 1993.

[14] S. K. Patra, *Development of fuzzy system based channel equalizers*, Ph.D. Thesis, Department of Electronics and Electrical Engineering, Edinburgh University, UK, 1998.

[15] J. M. Mendel, "Fuzzy Logic Systems for Engineering: A Tutorial," *Proceedings of the IEEE*, vol. 83, pp. 345–377, 1995.

[16] R. Ferrari et al., "A Hybrid Unsupervised Clustering Algorithm for Channel Equalization," *Proceedings of IEEE Machine Learning for Signal Processing*, Maynooth, Ireland, 2006.

[17] R. Suyama et al., "A Nonlinear Prediction Approach to the Blind Separation of Convolutive Mixtures," *EURASIP Journal on Applied Signal Processing*, vol. 2007, Article ID 4386, 2007.

[18] A. Hyvärinen, "Fast and Robust Fixed-Point Algorithms for Independent Component Analysis," *IEEE Transactions on Neural Networks*, vol. 10, no. 3, pp. 626–634, 1999.

ACKNOWLEDGEMENT

The authors would like to thank CAPES, CNPq and FAPESP for the financial support.