PREFILTERED BLIND EQUALIZATION: HOW TO FULLY BENEFIT FROM SPATIO-TEMPORAL DIVERSITY?

A. de Baynast†*, I. Fijalkow*

Equipe de Traitement du Signal et des Images,
ENSEA - Univ. de Cergy-Pontoise, France
`e-mail: debaynas, fijalkow@ensea.fr`

ABSTRACT

We propose a new structure for blind equalization with spatio-temporal diversity at the receiver (SIMO system). This structure is derived from the Wiener filter. It is composed by a prefilter deduced from the data spectral density (no blind criterion is needed) and a FIR filter, independent from the noise level, to be obtained blindly. The new prefilter is robust to a total lack of effective diversity and benefits from any available diversity. Simulations show that this new blind equalizer performs similarly to the Wiener filter even with a very small amount of data.

1. INTRODUCTION

It is crucial for many applications of digital communication (such as eavesdropping) to bypass the training period. Blind techniques (i.e. without training) have also received an increasing interest in both last decades. In this paper, we consider a filter at the receiver front end to improve the blind equalization performance. The reference algorithm in this context is the Constant Modulus Algorithm (CMA) [1]. Unfortunately, this algorithm requires a large number of samples to give a good estimation for its fourth order statistics elements. However in high-rate wireless applications, the channel may vary rapidly. One challenge of blind equalization is also to achieve good performance even with a small number of samples (less than 500). To speed up the convergence, a possible choice consists in whitening the received data before CMA as in [2]. Its convergence analysis is studied in [3].

To improve the blind equalization performance, we consider equalization structures using training. Optimal performance to reduce the intersymbol interferences (ISI) introduced by the frequency selective channel is achieved by the Maximum Likelihood Sequence Estimation receiver (MLSE). Unfortunately, the MLSE has a computational complexity that grows exponentially with the channel time dispersion. In most channels of practical interest, such a large computational complexity is prohibitively expensive to implement. A suboptimum channel equalization approach to compensate for the ISI employs the Wiener filter minimizing the Mean Square Error (MSE). For a Finite Impulse Response (FIR) channel even with spatio-temporal diversity, the optimal MMSE filter is recursive [4]. In this paper, we also design a prefilter directly derived from the structure of the Wiener filter in the context of spatio-temporal diversity at the receiver (oversampling or multisensor). The recursive filter being optimized from second order statistics only, it requires few samples to be estimated. Moreover, we show it is able to fully benefit from spatio-temporal diversity.

Section 2 contains a description of the system model. Section 3 shows how to adapt the structure of the Wiener receiver in a blind context with diversity. Simulations and analysis are described respectively in sections 4 and 5.

†A. de Baynast is supported by a DGA (French Defense) PhD grant.

2. PROBLEM FORMULATION

2.1. Channel modelization

We are interested in the multichannel communication system obtained from spatio-temporal diversity at the receiver. A data stream, $s[n]$ is transmitted at the rate $1/T_s$ over a frequency selective channel corrupted by additive noise. At the m^{th} sensor , the received continuous time baseband signal can be written as: $y^{(m)}(t) = \sum_i s[i]e^{(m)}(t - iT_s) + w(t)$ where $e^{(m)}(t)$ is the overall channel response (propagation channel to the m^{th} antenna + transmit filters). Oversampling the signal at the receiver at N times the symbol rate $1/T_s$ leads to a multivariate discrete-time model at the symbol rate that can be expressed as:

$$\mathbf{y}[n] = \sum_i \mathbf{e}[i]s[n-i] + \mathbf{w}[n] \tag{1}$$

where
$\mathbf{y}[n] = \big(y^{(0)}(nT_s)y^{(0)}(nT_s + T_c)\dots y^{(0)}(nT_s + (N-1)T_c)$ $\dots y^{(M-1)}(nT_s)\dots y^{(M-1)}(nT_s + (N-1)T_c)\big)$ and $\mathbf{e}[n]$ of degree L_e is defined accordingly. (1) can be viewed as a

convolutive Single Input/Multiple Outputs (SIMO) system with $L = NM$ outputs.

Let $\mathbf{y}(z)$ be the representation of $\mathbf{y}[n]$ using the z-transform. Similarly, we define $\mathbf{e}(z)$ and $\mathbf{w}(z)$. With these definitions, we may write:

$$\mathbf{y}(z) = \mathbf{e}(z)s(z) + \mathbf{w}(z) \tag{2}$$

2.2. Linear equalizer

We are looking for an equalizer as a linear filter $\mathbf{g}(z)$. The equalized data are:

$$x(z) = \hat{s}(z) = \mathbf{g}^{\dagger}(z)\mathbf{y}(z).$$

In the context of spatio-temporal diversity, a FIR Zero-Forcing equalizer exists if the convolution matrix of the channel $\mathcal{T}(\mathbf{e})$ is left-invertible, i.e. full rank which implies in particular no common zero to the components of $\mathbf{e}(z)$. However, in the presence of additive noise, the Minimum MSE (MMSE) is achieved by an IIR filter as explained below.

3. A PREFILTERED BLIND STRUCTURE

3.1. Structure motivation: MMSE equalizer

Assuming $s[n]$ is white, unit variance and uncorrelated with the white noise $w[n]$, the filter optimizing the MSE criterion $J(g) = \mathbb{E}[|x[n] - s[n-\nu]|^2]$ is described by:

$$\mathbf{g}_{\text{MMSE}}(z) = z^{-\nu}(\sigma^2 I_L + \mathbf{e}(z^{-1})\mathbf{e}^{\dagger}(z))^{-1}\mathbf{e}(z^{-1}) \tag{3}$$

where σ^2 denotes the noise variance [4].
Using the matrix inversion lemma[1], it can be rewritten as,

$$\begin{aligned} \mathbf{g}_{\text{MMSE}}(z) &= \frac{z^{-\nu}}{\sigma^2}\left\{ I_L - \frac{\mathbf{e}(z^{-1})\mathbf{e}^{\dagger}(z)}{\sigma^2 + \mathbf{e}^{\dagger}(z)\mathbf{e}(z^{-1})} \right\} \mathbf{e}(z^{-1}) \\ &= z^{-\nu}\frac{1}{\sigma^2 + \mathbf{e}^{\dagger}(z)\mathbf{e}(z^{-1})}\mathbf{e}(z^{-1}) \end{aligned} \tag{4}$$

$\mathbf{g}_{\text{MMSE}}(z)$ can be decomposed in a scalar recursive filter $g_{\text{AR}}(z) = \frac{1}{\sigma^2 + \mathbf{e}^{\dagger}(z)\mathbf{e}(z^{-1})} = \frac{1}{\sigma^2 + \sum_{l=0}^{L-1} e^{(l)*}(z)e^{(l)}(z^{-1})}$ and $\mathbf{g}_{\text{MA}}(z)$ matched to the channel (see fig.1).

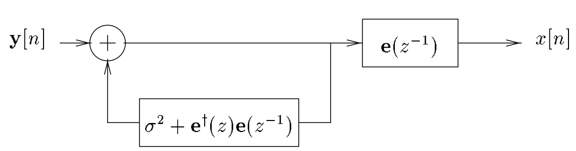

Fig. 1. Wiener receiver for a SIMO system.

The interest of this structure is its robustness to a lack of diversity. Indeed, as in the mono-variate case, the variations of the function $d(\sigma, \omega) = ||\mathbf{e}(e^{j\omega})||^2 + \sigma^2 = \sum_{l=0}^{L-1} |e^{(l)}(e^{j\omega})|^2 + \sigma^2$ determine the difficulty to equalize the channel. This function is all the more flat that $\mathbf{e}(z)$ has no common zeros to all its components and that close to common zeros are far away from the unit circle. This is exactly the concept of effective diversity, see [5]. Consequently, increasing the diversity factor L can only be beneficial, even if zeros of the different components are close of each other (the gain is small). Note that in the case of a total lack of diversity, all components are equal, $g_{\text{AR}}(z) = \frac{1}{L}\frac{1}{\frac{\sigma^2}{L} + e^{\dagger}(z)e(z^{-1})}$ reduces to the case of having no diversity at all. Anyhow, it achieves still MMSE equalization.

Concerning the transversal part, $\mathbf{g}_{\text{MA}}(z)$ cannot suffer from a lack of diversity (the optimal filter for $\mathbf{g}_{\text{MA}}(z)$ is the filter matched to the channel). In addition, the filter $\mathbf{g}_{\text{MA}}(z)$ is independent of the noise level.

3.2. Proposed structure

The goal of this section is to calculate blindly $\mathbf{g}(z)$ using the structure of (4). First, we show that $g_{\text{AR}}(z)$ can be deduced from the spectral density of $\mathbf{y}^{\dagger}$. Indeed, we use:

$$s_{\mathbf{y}^{\dagger}\mathbf{y}}(z) = L\sigma^2 + \mathbf{e}^{\dagger}(z)\mathbf{e}(z^{-1}) \tag{5}$$

from which we subtract $(L-1)\sigma^2$,where σ^2 is the smallest eigenvalue of the covariance matrix related to $\mathbf{y}$. Applying spectral factorization (Féyer-Riesz theorem), we obtain the minimum phase factor defined as a scalar, causal and stable transfer function $a(z)$ satisfying:
$\left(\sigma^2 + \mathbf{e}^{\dagger}(z^{-1})\mathbf{e}(z)\right)^{-1} = a(z)a(z^{-1}) = g_{\text{AR}}(z)$. We suggest also to do a forth and back recursive filtering with $a(z)$.

Remark 1 *In the mono-variate case, $a(z)$ is exactly a whitening filter, since the spectral density is exactly equal to $s_y(z) = \sigma^2 + e^*(z)e(z^{-1})$, see [6].*

The remaining processing, $z^{-\nu}\mathbf{e}(z^{-1})$, can only be performed blindly (since the channel $\mathbf{e}(z)$ is unknown), using CMA for example:

$$\mathbf{g}_{\text{MA}} = \arg\min_{g} \mathbb{E}\left[(|\mathbf{g}_{\text{MA}}^{\dagger} * \mathbf{y}[n]|^2 - 1)^2 \right]$$

Note that since $\mathbf{e}$ is FIR, the filter to be optimized blindly is FIR of the same order, independently of the noise level and of the diversity factor. With this structure, the number of filter taps to be optimized blindly is greatly reduced.

3.3. Implementation

Since the estimation of $a(z)$ is very sensitive to the errors of estimation of the spectral density $s_{\mathbf{y}^{\dagger}\mathbf{y}}(z)$ and to the noise variance σ^2, we propose to approximate this recursive filter

[1] When U and V are L-long vectors, $(I + UV^{\dagger})^{-1} = I - (1 + V^{\dagger}U)^{-1}UV^{\dagger}$.

by a transversal filter of order L_a. Ideally, the length of this filter should be infinite. We will discuss the choice of the value L_a in the sequel and suppose only in this section that $L_a \gg L_e$. This transversal filter $a(z)$ can also be viewed as the linear predictor of order L_a of the received signal $\mathbf{y}^\dagger$. Since $\frac{1}{\sigma^2+\sum_{l=0}^{L-1} e^{(l)*}(z)e^{(l)}(z^{-1})}$ has less variations than $\frac{1}{\sigma^2+e^*(z)e(z^{-1})}$, the predictor is better estimated with a FIR filter of fixed length L_a. This can be viewed as another benefit from spatio-temporal diversity.

Using a finite block of data, the proposed processing structure, prefiltered CMA, can be described by:

1. estimation of $a(z)$, solving the Yule-Walker equations:
$$(\hat{\mathcal{R}}_{\mathbf{y}^\dagger\mathbf{y}} - (L-1)\hat{\sigma}^2 I)\mathbf{a} = \hat{\mathbf{p}} \tag{6}$$
where $\hat{\mathcal{R}}_{\mathbf{y}^\dagger\mathbf{y}}$ is the estimated covariance matrix of $\mathbf{y}^\dagger$, $\hat{\mathbf{p}} = \left[\hat{r}_{\mathbf{y}^\dagger\mathbf{y}}[1]\hat{r}_{\mathbf{y}^\dagger\mathbf{y}}[2]\ldots\hat{r}_{\mathbf{y}^\dagger\mathbf{y}}[L_a-1]\right]^T$ and $\mathbf{a} = [a[0]a[1]\ldots a[L_a]]^T$.

2. filtering by $a(z)$. Then, considering the filtered data block in reverse order, filtering by $a^*(z)$,

3. blind optimization of the remaining transversal FIR filter using block CMA. We use a deterministic gradient descent CMA, described in the complex case by:
$$\mathbf{g}_{\text{MA}}^{(i+1)} = \mathbf{g}_{\text{MA}}^{(i)} - \mu\nabla_{\mathbf{g}_{\text{MA}}} J_{\text{CM}}(\mathbf{g}_{\text{MA}}^{(i)}),$$
where $\nabla_{\mathbf{g}_{\text{MA}}} J_{\text{CM}}(\mathbf{g}_{\text{MA}}^{(i)}) = 2(I\otimes\mathbf{g}_{\text{MA}}^\dagger)\hat{Q}_{\mathbf{y}}(\mathbf{g}_{\text{MA}}^*\otimes I)\mathbf{g}_{\text{MA}} - 4\hat{R}_{\mathbf{y}\mathbf{y}^\dagger}\mathbf{g}_{\text{MA}}^*$, $\otimes$ denotes the Kronecker product. The estimated covariance matrix is given by $\hat{\mathcal{R}}_{\mathbf{y}} = \frac{1}{N_e}\sum_{n=0}^{N_e-1}\mathbf{y}_{L_g}[n]\mathbf{y}_{L_g}^\dagger[n]$ and the estimated quadricovariance matrix by $\hat{Q}_{\mathbf{y}} = \frac{1}{N_e}\sum_{n=0}^{N_e-1}\left(\mathbf{y}_{L_g}[n]\mathbf{y}_{L_g}^\dagger[n]\otimes\mathbf{y}_{L_g}[n]\mathbf{y}_{L_g}^\dagger[n]\right)$ with N_e being the number of samples used for estimation and $\mathbf{y}_{L_g} = [y[n]y[n-1]\ldots y[n-L_g+1]]^T$.

4. SIMULATIONS

Simulation results confirm the good behaviour and performance of the proposed method. The diversity factor L is equal to 2 in these experiments. $s[n]$ is a BPSK i.i.d. sequence. The length of both components of the channel L_e is equal to 4. Each plot shows the results of 10 Monte Carlo runs, where the data length for BER estimation in each one was $N_s = 10000$. The solid lines represent recursive approaches, whereas the dashed lines represent transversal approaches. The length of the recursive scalar prefilter $g_{\text{AR}}(z)$ is equal to the channel length L_e, as for the second step filter (optimized by the CM criterion (curve with $*$) and by the MSE criterion (curve with o)). For sake of comparison, we use a transversal filter of length $2L_e$: both methods (recursive and transversal) have still approximatively an equivalent computation cost.

4.1. Test Case 1: influence of the number of samples used for estimation

In the first case, the batch algorithm of section 3.3 was tested for both subchannels $e^{(1)}$ and $e^{(2)}$ defined by their zeros ($[\,0.8\ -0.4\ 0.5\,]$ and $[\,0.85\ -0.7\ 0.2\,]$). The algorithm was run with different numbers of symbols (from 200 to 700) used for the estimation of the covariance and the quadricovariance matrices. The noise level was set to SNR=9dB. As depicted in fig.2, the performance obtained by the prefiltered CMA and the recursive MMSE are similar (MMSE ≈ 0.24 and MSE of prefiltered CMA ≈ 0.26). Even for a small numbers of symbols used for estimation, $N_e = 250$, the difference between their errors does not exceed 0.04. Performance obtained by the transversal CMA filter is much more affected by the number of samples for the estimation (MSE grows from 0.32 to 0.36 if N_e decreases from 500 to 250). As expected, estimation of the second order statistics is satisfactory even if N_e is less than 500.

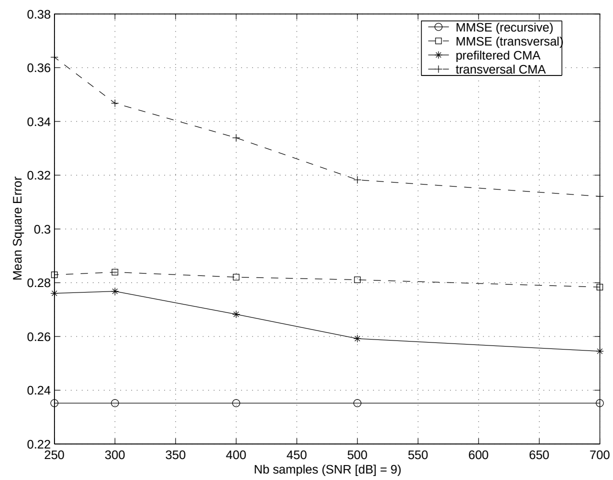

Fig. 2. Mean Square Error versus number of samples, at SNR = 9dB.

4.2. Test Case 2: comparison between the recursive structure and the transversal structure.

In this experiment, the zeros of the subchannels are respectively $[\,0.95\ -0.4\,]$ and $[\,0.96\ -0.5\,]$ to make the channel very difficult (close to common zeros near the unit circle). The number of samples used for estimation is equal to 250. Fig.3 shows the limitation of the structure using a transversal filter to equalize such a channel. Also, with nearly the same complexity (less for the proposed method), the prefiltered CMA outperforms the transversal CMA (improvement of $4dB$ for $N_e = 250$) and is very close to the Wiener performance.

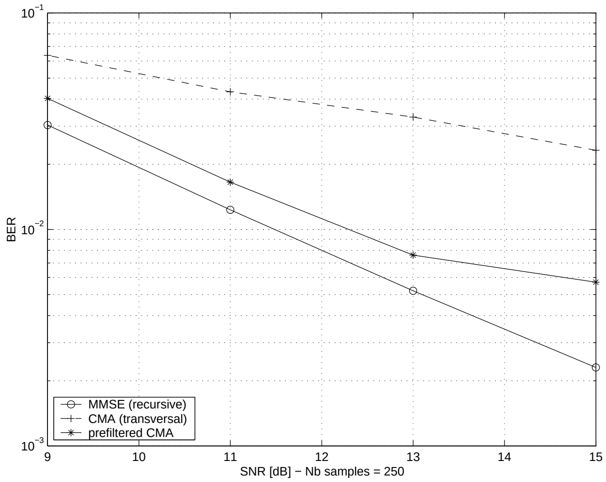

Fig. 3. Bit Error Rate versus Signal to Noise Ratio [dB]. (number of samples used for estimation is equal to 250).

5. COMPARISON WITH PREWHITENED CMA

In this section, we compare the proposed prefiltered CMA with the prewhitened CMA (as in [2] in its linear phase). We know that prewhitened CMA has improved convergence since the blind optimization must equalize a unitary transformation [3]. This is not true for the prefiltered CMA. However, because of its Wiener-like structure, the proposed method outperforms the prewhitened CMA in fig.4. The simulation setting is that of test case 2.

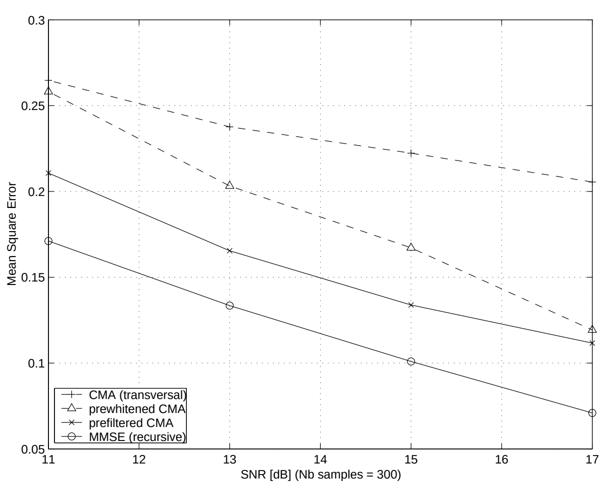

Fig. 4. Mean Square Error vs. Signal to Noise Ratio [dB]. (number of samples used for estimation is equal to 300).

This behavior can be explained by the fact that CMA with spatio-temporal diversity is known to converge to an equalizer response collinear to that of Wiener (see [4]). Our blind algorithm is constrained to mimic the Wiener solution with less equalizer taps to be optimized blindly than the prewhitened CMA. We can show that, assuming perfect prefiltering is achieved, the CM minima is achieved by

$$\mathbf{g}(z) \approx \alpha \mathbf{g}_{\text{MMSE}}(z)$$

with

$$\alpha^2 = \frac{\rho}{2\left(A - (1-\frac{\rho}{2})B\right)}$$

where A and B are defined from $q(e^{j\omega}) = \frac{\|\mathbf{e}(e^{j\omega})\|^2}{\sigma^2 + \|\mathbf{e}(e^{j\omega})\|^2}$ as,

$$A = \int q(e^{j\omega})d\omega, \; B = \frac{\sum_k |q[k]|^4}{A} \text{ with } \rho = \frac{\mathbb{E}[|s|^4]}{\mathbb{E}[|s|^2]^2}.$$

(The proof can be provided by e-mail). In the noiseless case, α is indeed equal to one.

6. CONCLUSION

We have proposed a new blind equalization algorithm taking full benefit from spatio-temporal diversity at the receiver. Our approach is derived from the recursive part of the Wiener solution which can be deduced from the data spectral density. As a result, the remaining transverse filter to be optimized blindly has the same length as the channel. Simulations show that our prefiltered CMA outperforms other algorithms when only a small amount of data is available. The achieved performance are very close to that of the Wiener solution.

7. REFERENCES

[1] J.R. Treichler and B.G. Agee, “A new approach to multipath correction of constant modulus signals,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 31, no. 4, pp. 459–472, April 1983.

[2] J. Labat, O. Macchi, and C. Laot, “Adaptive Decision Feedback Equalization: Can You Skip the Training Period?,” *IEEE Transactions on Communications*, vol. 46, no. 7, July 1998.

[3] J. Leblanc and I. Fijalkow, “Analysis of Pre-Whitened CMA,” *submitted to IEEE Signal Processing Letters*, 2000.

[4] C.R. Johnson Jr., P. Schnitter, I. Fijalkow, and al., *Blind Deconvolution II*, chapter The core of FSE-CMA behavior theory, Wiley, Haykin edition, 2000.

[5] I. Fijalkow, “Multichannel equalization lower bound: a function of channel noise and disparity,” *In Proceeding of SSAP*, vol. 31, no. 4, pp. 344–7, June 1996.

[6] A. Bouttier, *Egalisation autodidacte des modulations linéaires par filtrage récursif*, Ph.D. thesis, Université de Rennes, 2000.