A NEW PENCIL CRITERIUM FOR MULTICHANNEL BLIND DECONVOLUTION IN DATA COMMUNICATION SYSTEMS[1]

Santiago Zazo; José Manuel Páez Borrallo;*

Universidad Alfonso X El Sabio, Madrid. Spain
** ETS Ingenieros de Telecomunicación - Universidad Politécnica de Madrid. Spain*
Phone: 341-8109153; Fax: 341- 8109101; e-mail: szazo@uax.es

ABSTRACT

It is well known that blind channel deconvolution enables the receiver to equalize the channel simply by analyzing the received digital signal. Much of the work in 1990's faces the challenge presented by multiple-output systems, exploiting cyclostationarity properties and multivariate formulation of the incoming data. Our proposal is twofold: on one hand, we develop a theoretical analysis of a new blind channel deconvolution scheme by the exploitation of some shifting properties of the autocorrelation matrices of the source, in order to propose an appropriate cost function; on the other hand, an efficient programming is considered based on a Generalized Rayleigh Quotient formulation by using a Conjugate Gradient algorithm.

1. INTRODUCTION

Blind system deconvolution is a fundamental system technology that retrieves unknown information regarding a system or channel by analyzing the characteristics of its output only and probably some information (deterministic or statistical) about the system or the transmitted sequence, but not the sequence itself [2].

Where much of the work in the 1980's was limited to single-output systems, work in the 1990's faces the challenge presented by multiple output systems, as the result of temporal and/or spatial oversampling. This new representation (not really new because it had been used for long time as a robust synchronization-equalization scheme known as Fractionally Spaced Equalization FSE) has supposed a major breakthrough in digital signal processing applied to data communication systems: algorithms for single-output schemes usually suffer from two drawbacks: they are subjected to local minima and shows an slow convergence rate. In order to overcome these drawbacks, it has been presented several new methods allowing the blind identification or deconvolution using only second order statistics. For a system to be identifiable, all channels must be distinguishable from each other, the input sequence must be complex enough, and there must be a sufficient number of output samples available.

In our opinion, there are two main approaches exploiting this formulation: on one hand, methods dealing with the proposal of an identification criteria for the system function: i.e., maximum likelihood method or subspace methods. Likewise, other criteria focused on the estimation of the transmitted sequence, as the input subspace method, the mutually referenced equalizer and also linear prediction methods. The choice between direct system function estimation and direct input sequence estimation depends on the application [1].

Our proposal is based on the works by *Tong et al*, but introducing a different perspective. Their works have shown an identification method based on the knowledge of the 'forward-shift' structure of the source [7]. The change in the rank of these correlation matrices provide enough information to identify the multichannel matrix. By our own, we are not concerned with the identification but with the problem of multichannel blind equalization: that is, our goal is the proposal and theoretical analysis of a new cost function for blind multichannel equalization. As well, we have realized that this criterium can be formulated as a Generalized Rayleigh Quotient, therefore providing a fast implementation with low computational complexity through conjugate gradient techniques. Several simulations over standard communication channels demonstrate the performance of the method.

2. PROBLEM STATEMENT

First, let us recall the problem formulation, showing the classical block diagram in fig.1:

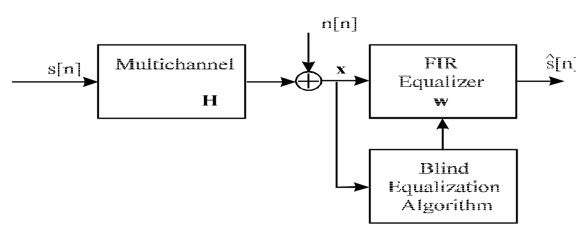

Fig.1. Block diagram of multichannel transmission and equalization scheme

[1] This work has been supported by National Project TIC95-0026

Consider a mathematical model where the input and the output are discrete , the system operator **H** is lineal and shift invariant, the system is driven by a single input sequence $s[n]$ and yields several output sequences. The multiple channel is required to be FIR for practical applications, but obviously an IIR channel can be well approximated by a FIR channel for an arbitrary length, large enough.

The behavior of the system can be summarized as follows: an independent sequence $s[n]$ is transmitted through a multichannel obtained either by an array receiver, either oversampling or both of them. In this situation, the received sequence is cyclostationary, and therefore a second order algorithm would be enough for identifying or equalizing a possible non minimum phase channel. It is well known that any cyclostationary signal could be considered as a stationary vector obtained by filtering by the full column rank matrix **H** (block Toeplitz structure). Sequence $n[n]$ is a Gaussian white noise independent of the transmitted sequence. The model usually assumed, establishes a linear model stated as follows:

$$\mathbf{x}(n) = \mathbf{H}\mathbf{s}(n) + \mathbf{n}(n) \tag{1}$$

Neglecting the noise effect, let us formulate the objective of any blind deconvolution scheme as the estimation of one vector **w** holding the following condition:

$$\mathbf{w} \mid \mathbf{w}^H \mathbf{x}[n] = \mathbf{w}^H \mathbf{H}\,\mathbf{s}[n] = e^{j\phi}\; s[n-k] \tag{2}$$

where ϕ is a phase ambiguity inherent to the problem.

Therefore, our goal is to design an appropriate cost function to find this vector **w**, uniquely, in order to deconvolve the transmitted sequence. In this approach, the information we want to use is only based on the estimation of several shifted correlation matrices $\mathbf{R}_\mathbf{x}(k)$ defined as follows:

$$\mathbf{R}_x(k) = E\left\{\mathbf{x}[n]\mathbf{x}^H[n-k]\right\} \tag{3}$$

(operator E means statistical expectation and superscript $(\ )^H$ means hermitian). Let us remark at this point that all the information involved in our criterium should be easily estimated from the incoming data as autocorrelation matrices for several shifting indexes.

3. OUR PROPOSAL

First of all, let us derive our criterium in this next section, providing a theoretical analysis and proof. This approach is based on the feasible condition for blind identification usually assumed when the virtual channels do not share any common zeros [1].

<u>*Step 1*</u>. Any matrix $\mathbf{H}_{mxd}$ ($m \geq d$) could be decomposed in the following form (Singular Value Decomposition)

$$\mathbf{H} = \mathbf{U}_{mxm}\mathbf{D}_{mxd}\mathbf{V}_{dxd} \tag{4}$$

where matrices dimension are specifically expressed for convenience. As it is shown in [7], equation (4) can be expressed as:

$$\mathbf{H} = \mathbf{U}_{\mathbf{H}_{mxd}}\mathbf{D}_{\mathbf{H}_{dxd}}\mathbf{V}_{dxd}\; e^{j\phi} \tag{5}$$

where $\mathbf{U_H}$ are the 'd' left singular vectors, $\mathbf{D_H}$ is the diagonal matrix collecting the non zeros singular values of **H**, and ϕ is a real phase.

<u>*Step 2.*</u> In lemma 1 [7] is shown that both matrices $\mathbf{U_H}$, $\mathbf{D_H}$ can be identified from the eigenvalue decomposition of the unshifted autocorrelation matrix $\mathbf{R}_x(0)$.

<u>*Step 3.*</u> Let us define the following linear transformation applied to the received data:

$$\mathbf{z}[n] = \mathbf{D}_\mathbf{H}^{-1}\,\mathbf{U}_\mathbf{H}^H\,\mathbf{x}[n] = \mathbf{D}_\mathbf{H}^{-1}\mathbf{U}_\mathbf{H}^H\mathbf{H}\,\mathbf{s}[n] = \mathbf{V}\mathbf{s}[n] \tag{6}$$

where **V** is a unitary matrix

<u>*Step 4.*</u> Exploiting the unitary character of **V**, it is obvious that any vector proportional to one of the right singular vectors (let us denote it by $\mathbf{v}_0$) will behave as a proper equalizer, that is:

$$\mathbf{v}_0^H\mathbf{z}[n] = \mathbf{v}_0^H\mathbf{V}\mathbf{s}[n] = e^{j\phi}\; s[n-k] \tag{7}$$

<u>*Step 5.*</u> Let us define an auxiliary correlation matrix **R** of vector **z**[n] as a linear combination of several shifted correlation matrices for arbitrary shifting indexes:

$$\begin{aligned}\mathbf{R} &= \sum_k E\left\{\mathbf{z}[n]\mathbf{z}^H[n-k]\right\} = \\ &= \sum_k E\left\{\mathbf{V}\mathbf{s}[n]\mathbf{s}^H[n-k]\mathbf{V}^H\right\} = \\ &= \sum_k \mathbf{V}\mathbf{R}_s(k)\mathbf{V}^H\end{aligned} \tag{8}$$

Assuming that **s**[n] is a white zero mean stationary process, their autocorrelation shifted matrices have the following form:

$$\mathbf{R}_s(k) = \begin{cases}\mathbf{J}^k, & k \geq 0 \\ \left(\mathbf{J}^H\right)^{|k|}, & k < 0\end{cases} \tag{9}$$

where **J** is dxd 'shifting' matrix as is defined in [7]. Equation (8) can be written then in a very compact formula just generalizing the effect of **J** by a new matrix **G** defined as:

$$\begin{aligned}\mathbf{R} &= \sum \mathbf{V}\mathbf{R}_s(k)\mathbf{V}^H = \\ &= \mathbf{V}\left(\sum_{k\geq 0}\mathbf{J}^k + \sum_{k<0}\left(\mathbf{J}^H\right)^{|k|}\right)\mathbf{V}^H = \mathbf{V}\mathbf{G}\mathbf{V}^H\end{aligned} \tag{10}$$

where, of course, a specific structure and condition for the square matrix **G** must be imposed in order to achieve the desired behavior. Let us remark at this point that equation (10) is a similarity transformation with some interesting properties [5].

<u>*Step 6.*</u> It is well known that if matrix **G** is normal, that is, real symmetric or hermitian, it has a diagonal Jordan form, and therefore there exists a similarity transformation involving a diagonal matrix $\mathbf{\Lambda}$:

$$\mathbf{G} = \mathbf{U}_\mathbf{G}\mathbf{\Lambda}\mathbf{U}_\mathbf{G}^H \tag{11}$$

where $\mathbf{U_G}$ is a unitary matrix (collecting the eigenvectors) and $\mathbf{\Lambda}$ is diagonal with the eigenvalues of **G**. Regarding that **R** and **G** are related by a similarity transformation (eq. 10), matrix **R** can also be decomposed as a similarity transformation with the same diagonal matrix, i.e.

$$\mathbf{R} = \mathbf{U_R} \Lambda \mathbf{U_R}^H \tag{12}$$

Several interesting properties are associated to both matrices (**R** and **G**), but let us remark just one of them: 'Vector $\mathbf{v_0}$ is an eigenvector of **G** associated with an eigenvalue λ, if and only if $\mathbf{V}^H\mathbf{x}$ is an eigenvector of **R** associated with the same eigenvalue'.

Step 7. It has been pointed out in steps 5-6 some mathematical background in order to present our deconvolution criterium: recall that, at this point, three main ideas must be remarked:

a) Matrix **R** is easily estimated from data $\{x\}$.

b) We only need the knowledge of one possibly scaled right singular vector of **V**.

c) Eigenvectors of **G** and **R** are related by a unitary similarity transformation.

Let us observe then that when imposing one of the eigenvectors of **G** to be $\mathbf{x}=[\mathbf{0}\ 1\ \mathbf{0}]$, the eigenvector of **R** associated to this eigenvalue will be $\mathbf{v_0}= \mathbf{V}^H\mathbf{x}$ as one of the columns of **V**, i.e. a proper equalizer. Therefore, in order to properly deconvolve the transmitted sequence, let us proceed as follows:

1. Matrix **G** must be symmetric, rank(d-1) with a null subspace whose associated eigenvector is $\mathbf{x}=[\mathbf{0}\ 1\ \mathbf{0}]$ (the choice is obvious as we can see in the following two examples for d=5)

$$\mathbf{G}_1 = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}, \ \mathbf{G}_2 = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{bmatrix} \tag{13}$$

2. Solve the related eigensystem:

$$\begin{aligned} \mathbf{v}_0^H \mathbf{R}_1 \mathbf{v}_0 &= 0 \\ \mathbf{v}_0^H \mathbf{v}_0 &= 1 \end{aligned} \tag{14}$$

4. PRACTICAL IMPLEMENTATION

In the previous section, we have discussed about the design of an equalization criterium based on a proper manipulation of some shifted correlation matrices of incoming data. However, so far, the practical implementation of our method seems to be expensive in computational requirements because subspace analysis are involved.

Let us go then into depth on equation (14) in order to develop a more feasible implementation. Applying the following linear transformation to **w**

$$\mathbf{v}_0 = \mathbf{D_H}^H \mathbf{U_H}^H \mathbf{w} \tag{15}$$

on equation (14) yields:

$$\begin{aligned} \mathbf{w}^H \mathbf{U_H} \mathbf{D_H} \mathbf{V} \mathbf{G} \mathbf{V}^H \mathbf{D_H}^H \mathbf{U_H}^H \mathbf{w} &= 0 \\ \mathbf{w}^H \mathbf{U_H} \mathbf{D_H} \mathbf{D_H}^H \mathbf{U_H}^H \mathbf{w} &= 1 \end{aligned} \tag{16}$$

which after some algebra shows an attractive form:

$$\begin{aligned} \mathbf{w}^H \mathbf{HGH}^H \mathbf{w} &= \lambda_0 \quad (\lambda_0 = 0) \\ \mathbf{w}^H \mathbf{HH}^H \mathbf{w} &= 1 \end{aligned} \tag{17}$$

these formulas are well known as the Generalized Eigenvalue problem (related with the null eigenvalue) and matrices $\mathbf{A}=\mathbf{HH}^H$, $\mathbf{B}=\mathbf{HGH}^H$ are known as pencil matrices. Let us remark again that theycan be estimated directly from data.

More indeed, equation (17) can be represented in a closed form through a Generalized Rayleigh Quotient (**G** is hermitian by definition):

$$\rho(\mathbf{w}) = \frac{\mathbf{w}^H \mathbf{Bw}}{\mathbf{w}^H \mathbf{Aw}} = 0 \tag{18}$$

i.e., the problem of the minimization of a quadratic function with a quadratic constraint.

In order to achieve a competitive computational complexity with a fast convergence of the associated adaptive process we should remark that eigenproblems are efficiently supported by using conjugate Gradient Techniques [6]. Also, it must be remembered the nice property of the Rayleigh quotient with singular points at which the gradient is zero are either unstable saddle points or the global minimum [8].

However, considering in our formulation that **B** is indefinite, the minimization process applied to equation (18) will lead us to the minimum (negative) eigenvalue instead of the desirable minimum absolute eigenvalue (see eq. (14)). This situation states that only by several iterations could be determined the lowest absolute eigenvalue by conjugate gradient techniques [5]. In order to increase the algorithm efficiency, this drawback can be avoided just considering and alternative implementation based on a modified criterium:

$$\min_{\mathbf{w}} \left(\frac{\mathbf{w}^H \mathbf{BB}^H \mathbf{w}}{\mathbf{w}^H \mathbf{Aw}} \right) \tag{19}$$

where $\mathbf{BB}^H$ is now positive definite. Although this equation provides different solutions than (18), observe that the eigenvector related with the minimun eigenvalue also holds condition (14).

Let us finally recall that this criterium was proposed by the authors in [9] but developed from another point of view. In that work, the authors proposed a criterium based on Bezout's theorem [3], leading to a set of quadratic equations identical to eq. (14). This optimization criterium was implemented also through efficient techniques based on Conjugate Gradient showing a promising feature. In the present contribution, a rigorous mathematical approach supports the previous ideas, also generalizing the structure of matrix **G**. Fuerthermore, a deep understanding on similarity transformation as provide an alternative formulation for our proposal.

5. COMPUTER SIMULATION RESULTS

In order to test the performance and convergence properties of our equalization algorithm, we have considered the scenario given in table 1, where several independent sequences are transmitted through four virtual complex channels [4]. White Gaussian noise is added to the output and a Signal to Noise Ratio (SNR) of 40 dB is considered.

In practical situations, the ensemble average of the signal correlation matrices are not known. Therefore, in the implementation of equation (19), the true matrices have been replaced by their recursive estimation. For example, choosing $\mathbf{G}_1$ in (13) we just need to estimate three data correlation matrices (using $\mathbf{G}_2$ in we should need the estimation of five shifted correlation matrices), i.e.:

$$\mathbf{A} = \mathbf{H}\mathbf{H}^H = \mathbf{R}_x(0)$$
$$\mathbf{B} = \mathbf{H}\mathbf{G}\mathbf{H}^H = \mathbf{R}_x(3) + \mathbf{R}_x(-3) \quad (20)$$

From a theoretical point of view, matrix **G** could be any symmetric rank-one null subspace, providing a null eigenvector **x**=[**0** 1 **0**]. That is, several combinations of shifted correlation matrices can be considered, all showing the desired performance. However, in order to reduce the computational requirements and also the estimation convergence, our proposal is limited to the estimation of the unshifted correlation matrix $\mathbf{R}_x(0)$ and two symmetric indexes for the shifted matrices $\mathbf{R}_x(k)$, $\mathbf{R}_x(-k)$ where k must be chosen as $k=(d+1)/2$ (assumed odd).

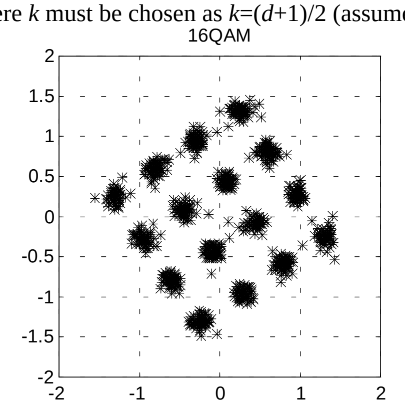

Fig. 2a. Deconvolved sequence for 16-QAM

In Fig.2 we show the deconvolved constellations for 16-QAM, and 8-PSK data transmission. The convergence speed could be very competitive to standard algorithms, also providing a simple implementation by Conjugate Gradient techniques: Fig.2a) for 16QAM at 1000 samples, and Fig. 2b), for 8PSK at 300 samples. It should be remarked that the different convergence speeds are mainly related with the constellations dimension in order to obtain a proper matrices estimation.

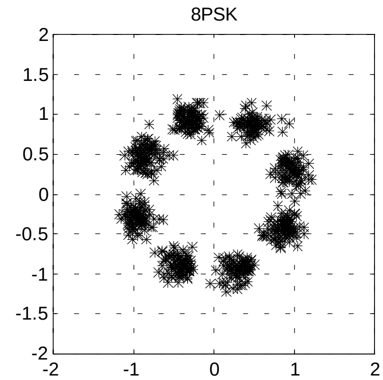

Fig. 2b. Deconvolved sequence for 8 - PSK

7. REFERENCES

[1]: Abed-Meraim, Karim, Qiu, W., Hua, Y. Blind system identification. *Proc. IEEE,* vol. 85, No8, August 1997, pp1310-1322.

[2]: Haykin, S. *Blind Deconvolution*, Englewood Cliffs, Prentice-Hall 1994.

[3]:Kaylath, T. *Linear systems*. Englewood Cliffs, Prentice Hall 1990.

[4]: Moulines, E., Duhamel, P., Cardoso, J.F., Mayrarge, S. Subspace methods for the blind identification of multiple FIR filters. *IEEE Transactions on Signal Processing*, Vol.43, No.2, February 1995, pp.516-525.

[5]:Noble, B., Daniel, J.W. *Applied linear algebra*. Prentice Hall 1988.

[6]:Sarkar, T.K., Yang, X. Application of the conjugate gradient and steepest descent for computing the eigenvalues of an operator. *Signal Processing 17*, 1989, Elsevier Sciences Publishers B.V., pp. 31-38.

[7]: Tong, L., Xu, G., Kailath, T. Blind identification and equalization based on second order statistics: a time domain approach. *IEEE Transactions on Information Theory*, Vol.40, No.2, March 1994, pp. 340-349.

[8]: Yang, X., Sarkar, T.K., Arvas, E. A survey of conjugate gradient algorithms for solution of extreme eigenproblems of a symmetric matrix. *IEEE Transactions on Acoustic, Speech and Signal Processing*, Vol.37, No.10, October 1989, pp. 1550-1555.

[9]: Zazo, S., Páez-Borrallo, J.M. A new multichannel blind equalization criterium based on a generalized Rayleigh quotient. *Proc. ICASSP'97*, Munich, pp. 3457-3460.

$\mathbf{h}_0$	-0.049+j 0.359	0.482-j 0.569	-0.556+j 0.587	1	-0.171+j 0.061
$\mathbf{h}_1$	0.443-j 0.0364	1	0.921-j 0.194	0.189-j 0.208	-0.087-j 0.054
$\mathbf{h}_2$	-0.211-j 0.322	-0.199+j 0.918	1	-0.284-j 0.524	0.136-j 0.19
$\mathbf{h}_3$	0.417+j 0.030	1	0.873+j 0.145	0.285+j 0.309	-0.049+j 0.161

Table 1. Scenario considered for simulations. Four virtual complex channels