Blind system identification using cross-relation methods: further results and developments

Abdeldjalil Aissa El Bey, Mimi Grebici, Karim Abed-Meraim, Adel Belouchrani

BLIND SYSTEM IDENTIFICATION USING CROSS-RELATION METHODS: FURTHER RESULTS AND DEVELOPMENTS

A. Aïssa-El-Bey, M. Grebici*, K. Abed-Meraim** and A. Belouchrani**

* EEE Dept., Ecole Nationale Polytechnique (El-Harrach), Algiers, Algeria.
** Signal & Image Proc. Dept., Telecom Paris, 46 rue Barrault, 75013 Paris, France.

ABSTRACT

We consider the problem of blind identification of FIR systems using the cross-relations (CR) method first introduced in [1]. Our contribution in this paper are as follows: (**i**) We introduce an extended formulation of the CR identification criterion which generalizes the standard CR criterion used in [1]. It can be shown that many existing multichannel blind identification methods belong to the class of generalized CR methods. (**ii**) We introduce a new identification method referred to as Minimum Cross-Relations (MCR) method which exploits with minimum redundancy the spatial diversity among the channel outputs. Simulation-based performance analysis of the MCR method and comparisons with CR method are also presented. (**iii**) Then, we present a modified version of the MCR referred to as the "unbiased MCR" (UMCR) method that leads to unbiased estimation of the channel parameters and better estimation performances without need of noise whitening as in the MCR. (**iv**) Finally, we discuss the multi-input case and show how additional difficulties arise due to the non-linear parameterization of the noise vectors in terms of the channel parameters.

1. INTRODUCTION

Blind system identification (BSI) is a fundamental signal processing technology aimed at retrieving a system's unknown information from its outputs only. This problem has received a lot of attention in the signal processing literature and a plethora of methods and techniques have been proposed to solve the BSI over the last 2 decades [6, 7]. Since 1991, it has been shown that using spatial and/or temporal diversity leads to efficient and simplified BSI methods using only the second order statistics of the outputs or even deterministic approaches. The CR method introduced in [1] is one of the simplest and efficient methods for blind identification of FIR SIMO systems. This paper focuses on the CR method and introduces several improvements and new developments related to this technique. At first, we reformulate the CR problem in such a way to provide a general framework where a large class of BSI methods can be seen as CR-like methods. Then we introduce several improvements /simplifications of the original CR method referred to as MCR (for Minimum Cross-Relations) and UMCR (for Unbiased Minimum Cross-Relations) method. Finally, we discuss the MIMO case and explain why the CR method cannot be extended "in a simple way" to solve the MIMO-BSI problem.

2. PROBLEM FORMULATION

Consider a SIMO system of q outputs given by:

$$\mathbf{y}(n) = \sum_{k=0}^{M} \mathbf{h}(k)s(n-k) + \mathbf{w}(n) \tag{1}$$

where $\mathrm{h}(z) = \sum_{k=0}^{M} \mathbf{h}(k)z^{-k}$ is an unknown causal FIR $q \times 1$ transfer function satisfying $\mathrm{h}(z) \neq 0,\ \forall z,\ s(n)$ a scalar (non-observable) stationary process and $\mathbf{w}(n)$ is an additive q-dimensional spatial white noise, i.e. $E[\mathbf{w}(n)\mathbf{w}^H(n)] = \sigma^2 \mathbf{I}_q$.

Given a finite set of observation vectors $\mathbf{y}(1), \ldots, \mathbf{y}(T)$ and based on the channel entries co-primness (i.e. $\mathrm{h}(z) \neq 0\ \forall z$), the objective here is to estimate the channel coefficients vector $\mathbf{h} = [\mathbf{h}(0)^T, \cdots, \mathbf{h}(M)^T]^T$ up to a scalar constant (this is an inherent indeterminacy of the BSI problem as shown in [5]) using CR-like techniques. Before proceeding, we review next the basic idea and principle of the original CR method in [5].

3. CR-LIKE METHODS

This section is devoted to the development of the MCR and UMCR methods as well as the generalized formulation of the CR criterion. For that, we start first by a brief review of the CR principle.

3.1. Review of the CR method

From (1), the noise-free outputs $y_i(k)$, $1 \leq i \leq q$ are given by:

$$y_i(k) = h_i(k) * s(k), \quad 1 \leq i \leq q \tag{2}$$

where "$*$" denotes convolution. Using commutativity of convolution, it follows:

$$h_j(k) * y_i(k) = h_i(k) * y_j(k), \quad 1 \leq i < j \leq q \tag{3}$$

This is a linear equation satisfied by every pairs of channels.

It was shown that based on $q(q-1)/2$ possible cross-relations, the channel parameters can be uniquely identified according to [5]:

Theorem 1 *Under the data model assumptions, the set of cross-relations (in the noise free case):*

$$y_i(k) * h'_j(k) - y_j(k) * h'_i(k) = 0, \quad 1 \leq i < j \leq q \tag{4}$$

where $\mathrm{h}'(z)$ *is a* $q \times 1$ *polynomial vector of degree* M*, is satisfied if and only if* $\mathrm{h}'(z) = \alpha \mathrm{h}(z)$ *for a given scalar constant* α*.*

By collecting all possible pairs of q channels, one can easily establish a set of linear equations. In matrix form, this set of equations can be expressed as:

$$\ _q\mathbf{h} = 0 \quad (5)$$

where $\ _q$ is defined by:

$$\mathbf{Y}_2 = [\mathbf{Y}_{(2)},\ -\mathbf{Y}_{(1)}]$$

$$\mathbf{Y}_l = \left[\begin{array}{ccc|c} & \mathbf{Y}_{l-1} & & \mathbf{0} \\ \hline \mathbf{Y}_{(l)} & & \mathbf{0} & -\mathbf{Y}_{(1)} \\ & \ddots & & \vdots \\ \mathbf{0} & & _{(l)} & -\mathbf{Y}_{(l-1)} \end{array}\right] \quad (6)$$

with $l = 3, \ldots, q$ and:

$$\mathbf{Y}_{(l)} = \begin{bmatrix} y_l(M) & \ldots & y_l(0) \\ \vdots & & \vdots \\ y_l(N-1) & \ldots & y_l(N-M-1) \end{bmatrix} \quad (7)$$

In the presence of noise, equation (5) can be naturally solved in the least-square (LS) sense according to:

$$\hat{\mathbf{h}}_{CR} = \arg\min_{\|\mathbf{h}\|=1} \mathbf{h}^H \mathbf{Y}_q^H \ \ _q\mathbf{h} \quad (8)$$

The CR method is referred to as the LS method in [5] because it represents the least-squares solution to the CR equation (5).

3.2. Minimum CR Method

In the same spirit as in the MNS (Minimum Noise Subspace) method [4], we show here that only $q-1$ (instead of $q(q-1)/2$) cross-relations can be used for channel identification. We have the following theorem:

Theorem 2 *Let* $\{p_1, \ldots, p_{q-1}\}$, $p_i = (i_1, i_2)$, $i = 1, \ldots, q-1$, *be a set of* $q-1$ *pairs of channels which form a tree structure, then the noise-free cross-relations* ($\mathrm{h}'(z)$ *being a polynomial vector of degree* M*):*

$$h'_{i_1}(k) * y_{i_2}(k) - h'_{i_2}(k) * y_{i_1}(k) = 0, \quad i = 1, \ldots, q-1 \quad (9)$$

yield a unique identification of the channel transfer function $\mathrm{h}(z)$*, i.e.* $\mathrm{h}'(z) = \alpha\mathrm{h}(z)$ *for a given scalar constant* α*.*

In figure 1, we consider the following example corresponding to $q = 5$ and $p_1 = (1,2)$, $p_2 = (1,3)$, $p_3 = (3,4)$ and $p_4 = (3,5)$ that is a set of four (i.e. : $q-1$) pairs forming a tree structure. To solve (9) in presence of noise, we estimate the channel parameters in the least squares sense according to:

$$\hat{\mathbf{h}}_{MCR} = \arg\min_{\|\mathbf{h}\|=1} \mathbf{h}^H \overline{}^H \overline{} \mathbf{h} \quad (10)$$

where $\overline{\mathbf{Y}}$ is a block matrix which (k,l) block entry is zero if $l \notin \{k_1, k_2\}$ with $p_k = (k_1, k_2)$ and is equal to $\mathbf{Y}_{(k_2)}$ if $l = k_1$ and $-\mathbf{Y}_{(k_1)}$ if $l = k_2$. Contrary to the CR method, the MCR requires noise whitening as the mean value of the noise term in the quadratic form of (10) is not proportional to identity, but to a positive diagonal matrix. In the illustrative example of figure 1 and under the white noise assumption the noise term satisfies:

$$E(\overline{\mathbf{W}}^H \overline{\mathbf{W}}) \propto diag(2,1,3,1,1)$$

where matrix $\overline{\mathbf{W}}$ is defined from the noise term similarly to $\overline{\mathbf{Y}}$. This would require a noise whitening to obtain unbiased estimation of the channel parameters.

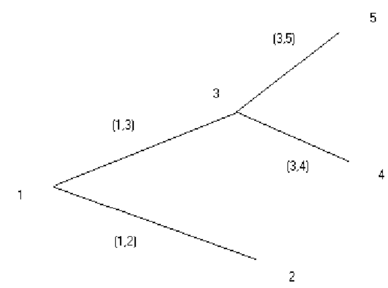

Figure 1: Example of a tree structure for $q = 5$.

3.3. Unbiased MCR Method

We introduce here a modified version of the MCR in such a way that the contribution of the noise term becomes proportional to identity. More precisely, instead of using $(q-1)$ cross-relations as in the MCR, we do use q cross-relations corresponding to the following pairs:

$$\begin{cases} p_1 = (1,2) \\ p_2 = (2,3) \\ \quad\vdots \\ p_{q-1} = (q-1, q \\ p_q = (q,1) \end{cases} \quad (11)$$

where the first $(q-1)$ pairs correspond to the MCR, and the last one represents a redundancy chosen such that all system outputs are used similarly[1].

This has also the advantage of rendering the performances of the method independent from a specific choice of the selected tree structure of the MCR. This leads to a slight performance gain as observed in the simulation results (see section 5). In addition, in the UMCR, noise whitening is not necessary as stated by the following theorem:

Theorem 3 *Under the data model assumptions, the channel parameter estimate given by the least squares solution of the cross-relations of (11) is asymptotically unbiased.*

4. GENERALIZED CR CRITERION

Equation (3) can be rewritten in a more compact form as:

$$[\mathbf{F}(h)] * \mathbf{y}(k) = 0 \quad (12)$$

or equivalently (dual form):

$$[\mathbf{G}(y)] * \mathbf{h}(k) = 0 \quad (13)$$

[1]This is not the case in the MCR as certain system outputs are used more than others. For example, in the first $(q-1)$ pairs of (11) system outputs 1 and q are used only once while the others are used twice.

where $[\mathbf{F}(h)]$ and $[\mathbf{G}(y)]$ are two matrix-valued operators depending linearly on the channel parameters and observation signals, respectively, according to:

$$\mathbf{F}(h) = \begin{bmatrix} h_2 & -h_1 & 0 & \dots & 0 \\ h_3 & 0 & -h_1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ h_q & 0 & \dots & 0 & -h_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & h_q & -h_{q-1} \end{bmatrix} \tag{14}$$

$$\mathbf{G}(y) = \begin{bmatrix} y_2 & -y_1 & 0 & \dots & 0 \\ y_3 & 0 & -y_1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ y_q & 0 & \dots & 0 & -y_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & y_q & -y_{q-1} \end{bmatrix} \tag{15}$$

In this case $[\mathbf{F}(h)]$ and $[\mathbf{G}(y)]$ have well defined specific forms given by (14) and (15). However, this specific forms are not necessary to achieve unique identification of the system parameters and thus a large class of functions $[\mathbf{F}(h)]$ and $[\mathbf{G}(y)]$ different from those used in (14) and (15) can provide admissible identification criteria (i.e. a criterion is said to be admissible if it yields a unique identification of the system parameters).

Also, we allow $[\mathbf{F}(h)]$ and $[\mathbf{G}(y)]$ to be linear or non-linear, explicit or implicit function of the channel parameters and observation signals respectively.

Note that the second form of criterion (15) is generally preferred for channel identification since it depends linearly on the unknown channel parameters. However, this latter is not always possible to derive from the first form (14) when $[\mathbf{F}(h)]$ is a non-linear or an implicit function of **h** (see the multi-input case below).

Using this general formulation it can be shown that methods as maximum likelihood [3], Subspace [2], MNS (Minimum Noise Subspace) [4], etc. are special members of the class of generalized CR methods.

For example the subspace method in [2] consists of estimating the channel parameters using the signal and noise subspace orthogonality. The estimation criterion can be written as:

$$\mathbf{\Pi} \;\; \mathcal{T}_N(h) = 0 \tag{16}$$

or equivalently:

$$\overline{\mathbf{\Pi}}(k) * \mathbf{h}(k) = 0 \tag{17}$$

where $\mathbf{\Pi}$ represents the noise subspace projection, $\mathcal{T}_N(h)$ is a block Sylvester matrix, and $\overline{\mathbf{\Pi}}$ is a filtering matrix computed from $\mathbf{\Pi}$ as:

$$\overline{\mathbf{\Pi}}_{ij}(k) = \mathbf{\Pi}_{i,j+kq} \tag{18}$$

Equation (17) has the form of (13) where $[\mathbf{G}(y)]$ is given here by $\overline{\mathbf{\Pi}}$, that is a non-linear, implicit function of the observation process $\mathbf{y}(k)$ (recall that the noise projection $\mathbf{\Pi}$ is computed from the data $\mathbf{y}(k)$ through the eigen decomposition of its covariance matrix).

5. THE MULTI-INPUT CASE

In the multi-input case the transfer function becomes a $q \times p$ matrix (z), $1 < p < q$ being the number of input signals. The noise-free observation can be written as:

$$\begin{aligned} \mathbf{y}_{(p)}(k) &= [\mathrm{H}_{(p)}(z)]\mathbf{s}(k) \\ y_{p+1}(k) &= [\mathrm{H}_{p+1,:}(z)]\mathbf{s}(k) \\ &\vdots \\ y_q(k) &= [\mathrm{H}_{q,:}(z)]\mathbf{s}(k) \end{aligned} \tag{19}$$

with $\mathbf{y}_{(p)}(k) = [y_1(k), \dots, y_p(k)]^T$, where $\mathrm{H}_{(p)}(z)$ is the top $p \times p$ sub-matrix of $\mathrm{H}(z)$, the $\mathrm{H}_{i,:}(z)$ is the $1 \times p$ ith row vector of $\mathrm{H}(z)$ and $[\mathrm{H}(z)]\mathbf{s}(k)$ represents the system outputs, corresponding to a transfer function $\mathrm{H}(z)$ excited by $\mathbf{s}(k) \stackrel{\mathrm{def}}{=} [s_1(k), \cdots, s_p(k)]^T$ the p-dimensional vector of independent source signals.

For unique channel identifiability, $\mathrm{H}(z)$ is assumed to be irreducible [5] i.e:

$$Rank[\mathrm{H}(z)] = p, \;\; \forall z$$

In addition, we assume here that the $p \times p$ sub-matrix $\mathrm{H}_{(p)}(z)$ is full rank, i.e:

$$\det(\mathrm{H}_{(p)}(z)) \not\equiv 0$$

In that case, the following $q - p$ cross-relations can be obtained by:

$$\mathbf{s}(k) = [H^{-1}_{(p)}(z)]\mathbf{y}_{(p)}(k) = [\frac{\mathrm{com}(\mathrm{H}_{(p)}(z))}{\det(\mathrm{H}_{(p)}(z))}]\mathbf{y}_{(p)}(k) \tag{20}$$

or equivalently:

$$\begin{aligned} [\det(\mathrm{H}_{(p)}(z))]y_{p+1}(k) &= [\mathrm{H}_{p+1,:}(z)\mathrm{com}(\mathrm{H}_{(p)}(z))]\mathbf{y}_{(p)}(k) \\ &\vdots \\ [\det(\mathrm{H}_{(p)}(z))]y_q(k) &= [\mathrm{H}_{q,:}(z)\mathrm{com}(\mathrm{H}_{(p)}(z))]\mathbf{y}_{(p)}(k) \end{aligned} \tag{21}$$

where $\mathrm{com}(\mathbf{A})$ and $\det(\mathbf{A})$ denote the co-factor matrix and determinant of **A**, respectively. Under above assumptions, this set of cross-relations yields a unique identification[2] of the channel parameters (up to a constant $p \times p$ non-singular matrix which represents in fact the inherent indeterminacy of the MIMO-BSI problem).

Unfortunately, as we can see in the case $p > 1$, the noise vectors are non-linear functions[3] of channel parameters. Therefore, a simple extension of the cross-relations algorithm to the multi-input case seems not to be possible.

6. SIMULATION RESULTS

We present here some numerical simulations to assess the performances of the proposed CR-like methods. We consider a SIMO system with $q = 6$ outputs represented by polynomial transfer function of degree $M = 4$. The channel coefficients are generated randomly (at each Monte-Carlo run) following the complex gaussian distribution, i.e. the amplitude of each channel coefficient is Rayleigh distributed with unit-variance while its phase is uniformly distributed in $[0, 2\pi]$. The input signal is a 4QAM iid sequence of length $T = 256$. The observation is corrupted by addition white gaussian noise with a variance σ^2 chosen such that the $SNR \quad \frac{\|\mathbf{h}\|^2}{\sigma^2}$ varies in the range $[0, 30]dB$.

[2] The proof is omitted here due to space limitation.

[3] In the mono-input case, noise vectors are expressed as linear functions of the channel parameters as shown by eq.(14).

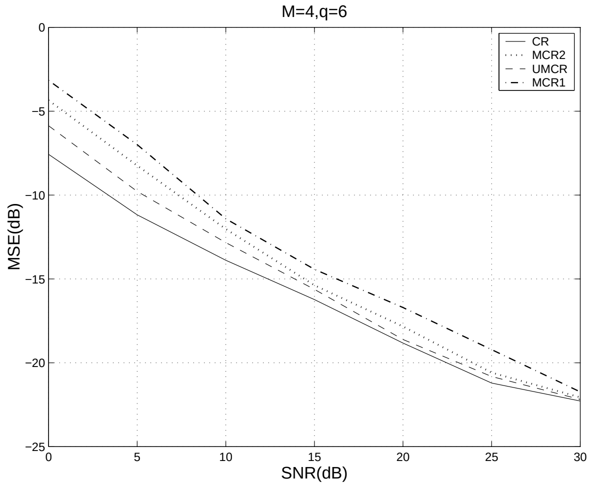

Figure 2: Performance comparison of the CR, MCR and UMCR for $q = 6$.

Statistics are evaluated over N_r 1000 Monte-Carlo runs and estimation performances are given by the normalized mean-square error criterion:

$$MSE \quad \frac{1}{N}\sum_{r=1}^{N_r}\frac{\|\hat{\mathbf{h}}_r - \mathbf{h}\ \|^2}{\|\mathbf{h}\ \|^2}$$

Where $\hat{\mathbf{h}}_r$ denotes the estimated channel coefficient vector at the r-th Monte-Carlo run.

In figure 2, we compare the performances of the CR, the MCR referred to as MCR1 with $p_i = (1, i)$, $i = 2, \ldots, q$, the UMCR and the MCR with $p_i = (i-1, i)$, $i = 2, \ldots, q$ referred to as MCR2.

We can observe that the choice of the structure underlying the MCR method can affect significantly the performances of the method. Also, we observe a slight loss of estimation performances of the UMCR compared to the CR method, but the former remains computationally much more efficient.

7. CONCLUSION

In this paper we have presented several extensions of the CR method originally introduced in [5]. These extensions consist of a general formulation of the CR criterion, a minimum CR (MCR) method, an unbiased MCR method and a discussion of the CR method for the MIMO case.

The MCR and UMCR methods presented in this paper are simplified version of CR that might reduce significantly the computational cost of the blind channel estimation especially for large systems.

8. REFERENCES

[1] H. Liu, G. Xu, and L. Tong, "A deterministic approach to blind identification of multichannel FIR systems," *in Proc. ICASSP*, vol. 4, pp. 581–584, 1994.

[2] E. Moulines, P. Duhamel, J. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Tr. on Sig. Proc.*, pp. 516–525, Feb. 1995.

[3] Y. Hua, "Fast maximum likelihood for blind identification of multiple FIR channels," *IEEE Tr. on Sig. Proc.*, Mar. 1996.

[4] Y. Hua, K. Abed-Meraim and M. Wax, "Blind system identification using minimum noise subspace," *8th IEEE Workshop on SSAP, Corfu, Greece*, June 1996.

[5] G. Xu, H. Liu, L. Tong, and T. Kailath, "A least-squares approach to blind channel identification." *IEEE Trans. Signal Processing*, vol. 43, pp. 2982–2993, Dec. 1995.

[6] K. Abed-Meraim, W. Qiu, and Y. Hua, "Blind system identification," *IEEE Tr. on Sig. Proc*, vol. 85, No 8, Aug 1997.

[7] L. Tong and S. Perreau, "Multichannel blind identification: from subspace to maximum likelihood methods", *Proceedings of the IEEE*, Volume: 86 Issue: 10 , pp. 1951–1968, Oct. 1998

9. APPENDIX

Proof of Theorem 2: Let consider the Z-transform of the $(q-1)$ cross-relations of (5) in the noiseless case. That leads to :

$$\widetilde{\mathrm{H}}(z)^T\mathrm{h}'(z) = 0$$

where $\mathrm{h}'(z) = [h_1'(z), \ldots, h_q'(z)]^T$ and $\widetilde{\mathrm{H}}(z)$ is a $q \times (q-1)$ polynomial matrix which k-th column is the zero valued vector except for the k_1-th and k_2-th entries that are equal to $h_{k_2}(z)$ and $-h_{k_1}(z)$ respectively (with $p_k = (k_1, k_2)$).

According to [4] and thanks to the tree structure, the columns of $\widetilde{\mathrm{H}}(z)$ form a basis of the rational subspace Range$\{\mathrm{h}(z)\}^\perp$, i.e., the orthogonal rational subspace to Range$\{\mathrm{h}(z)\}$. As a consequence $\mathrm{h}'(z)$ belongs to Range$\{\mathrm{h}(z)\}$ and since it is a polynomial vector with degree M equal to that of $\mathrm{h}(z)$, we have

$$\mathrm{h}'(z) = \alpha\mathrm{h}(z)$$

for a given scalar constant α.

Proof of Theorem 3: In presence of noise, matrix $\overline{\ \ }$ in (10) constructed from the set of pairs in (11) becomes :

$$\overline{\ \ } \quad \overline{\mathbf{X}} + \overline{\mathbf{W}}$$

where $\overline{\mathbf{W}}$ represents the additive noise term. Under the spatial and temporal white noise assumption and the independence of the noise and signal term, we have:

$$E(\overline{\mathbf{Y}}^H\overline{\mathbf{Y}}) = E(\overline{\mathbf{X}}^H\overline{\mathbf{X}}) + E(\overline{\mathbf{X}}^H\overline{\mathbf{W}}) + E(\overline{\mathbf{W}}^H\overline{\mathbf{X}}) + E(\overline{\mathbf{W}}^H\overline{\mathbf{W}}$$

the second and third term in the right side of above equation is equal to zero because of the independence of noise and signal terms. The last term is equal to :

$$E(\overline{\mathbf{W}}^H\overline{\mathbf{W}}) = 2\sigma^2(T-M)\mathbf{I}_{q(M+1)}$$

where σ^2 is the noise power, T the sample size and M the channel polynomial degree. Consequently, the channel estimate given by the least eigenvector of $E(\overline{\mathbf{Y}}^H\overline{\mathbf{Y}})$ coincide with that of the noiseless covariance matrix $E(\overline{\mathbf{X}}^H\overline{\mathbf{X}})$.