ANALYTICAL SOLUTION OF THE BLIND IDENTIFICATION PROBLEM FOR MULTICHANNEL FIR SYSTEMS BASED ON SECOND ORDER STATISTICS

Konstantinos I. Diamantaras[1] *and Theophilos Papadimitriou*[2]

[1] Department of Informatics
TEI of Thessaloniki
Sindos 57400, Greece
kdiamant@it.teithe.gr

[2] Dept. of Int. Econ. Relat. & Devel.
Democritus University of Thrace
Komotini 69100, Greece
papadimi@ierd.duth.gr

ABSTRACT

We present a novel method for the blind identification problem of multichannel FIR systems based on the analytical solution of a linear system of equations derived from the output signal second order statistics. The method is very efficient computationally since it does not involve iterative optimization procedures. In addition, it is very robust to noise as shown by our experiments. We also demonstrate that it compares favorably in terms of estimation accuracy compared to classical second order methods.

Index Terms— System identification, Blind identification, Multichannel Systems

1. INTRODUCTION

The System Identification problem refers to the identification of the impulse response of a linear system. Blind Identification refers to System Identification knowing only the system output signals (and the least possible information amount on the input signals). The described problem is of great importance as it can be applied in many engineering areas: from classic problems in wireless telecommunications and sound recording, to hard disk data recovery.

The previous work on the Blind Identification problem can be roughly separated in two general classes: methods using the Higher-Order Statistics (HOS) or Second-Order Statistics (SOS) of the output signals (often called mixtures). Methods of the former group exploit the higher order spectra of the signals in order to identify the channels [1], [2], [3]. A drawback of these methods is that they are limited in systems driven by non-Gaussian sources. However the biggest problem is that they usually rely on a minimization process with low convergence. That kind of process requires large dataset samples and important computational cost. Methods of the latter case exploit the special structure of mixture covariance matrices [4], [5], [6], [7]. The solution is often non-iterative (batch) making these methods computational attractive. Although they can treat system with Gaussian sources, their main disadvantage comes from the fact that they must be driven by white sources.

A recently treated system (born on wireless telecommunication scenarios) supposes that the channels are time-varying [7], [8], [9]. These methods usually suggest that the channels are ”slowly” changing in time and the identification process adapts to the new data. Another approach tries to transforms the time-varying systems into equivalent stable ones and solving them with well-known techniques [10], [11].

The first Blind Identification method treated a Single-Input Single-Output (SISO) system by transforming it into an equivalent Single-Input Multiple-Output (SIMO) system and then solving it using a matrix-pencil based method [12, 13]. Many methods exist nowadays treating the more complex problem of Mingle-Input Multiple-Output (MIMO) usually when the mixtures are more than the sources [3] [6] or even in cases where we have more sources than mixtures [14].

In this paper we will present a novel blind identification method of multichannel FIR systems based on the second order statistics of the output vector. The problem is transformed into a set of linear equations generated with the help of the SVD of the lagged covariance matrix. We shall demonstrate the conceptual simplicity of the approach and we shall compare against standard second order methods to establish the numerical robustness and efficiency of the proposed technique.

2. SYSTEM MODEL AND PROBLEM DESCRIPTION

Let us consider a complex $m \times 1$, FIR channel of length L, excited by the complex input $s(k)$, generating an m-dimensional vector sequence $\mathbf{x}(k)$:

$$\mathbf{x}(k) = \sum_{l=0}^{L-1} \mathbf{h}(l)s(k-l) + \mathbf{e}(k) \quad k = 1, \cdots, N \tag{1}$$

This work has been supported by the “EPEAEK Archimides-II” Programme, project “BSP-GRID”, funded in part by the European Union (75%) and in part by the Greek Ministry of National Education and Religious Affairs (25%).

The element h_i of $\mathbf{h}$ is the unknown complex filter connecting the input s with the i-th output x_i, while $\mathbf{e}(k)$ is additive white noise vector.

It is well known that if we form the vector sequence $\bar{\mathbf{x}}(k) = [\mathbf{x}(k)^T\, \mathbf{x}(k-1)^T \cdots \mathbf{x}(k-W+1)^T]^T$ by stacking observations in consecutive time windows of size W, then we can write

$$\bar{\mathbf{x}}(k) = \bar{\mathbf{H}}\bar{\mathbf{s}}(k) + \bar{\mathbf{e}}(k) \tag{2}$$

where

$$\bar{\mathbf{H}} = \begin{bmatrix} \mathbf{h}(0) & \cdots & \mathbf{h}(L-1) & & 0 & \\ & \ddots & & & \ddots & \\ 0 & & \mathbf{h}(0) & & \cdots & \mathbf{h}(L-1) \end{bmatrix} \tag{3}$$

is the generalized Sylvester matrix associated with the filter $\mathbf{h}$ and

$$\bar{\mathbf{s}}(k) = [\mathbf{s}(k)\, \mathbf{s}(k-1) \cdots \mathbf{s}(k-L-W+2)]^T,$$
$$\bar{\mathbf{e}}(k) = [\mathbf{e}(k)\, \mathbf{e}(k-1) \cdots \mathbf{e}(k-W+1)]^T.$$

Our assumptions regarding the input and noise signals are described below:
[A.1] The input process is wide sense stationary and the input samples are i.i.d., so

$$E\{s(k)s(k-l)^*\} = \delta(l), \quad \text{all } k. \tag{4}$$

[A.2] The noise samples are i.i.d.

$$E\{e_i(k)e_i(k-l)^*\} = \sigma^2\delta(l), \quad \text{all } i,\, k. \tag{5}$$

[A.3] The noise components are independent to each other and to the input signal

$$E\{e_i(k)e_j(k-l)^*\} = 0, \quad \text{all } i,\, j,\, k,\, l, \tag{6}$$
$$E\{e_i(k)s(k-l)^*\} = 0, \quad \text{all } i,\, k,\, l. \tag{7}$$

3. ANALYTICAL APPROACH

Let us inspect the delayed covariance matrix $\mathbf{R}_{\bar{s}}(l) \triangleq E\{\bar{\mathbf{s}}(k)\, \bar{\mathbf{s}}(k-l)^H\}$ of the source vector $\bar{\mathbf{s}}$. Using the assumptions of section 2 we can easily compute the covariance for delays $l = 0$ and $l = 1$ to be $\mathbf{R}_{\bar{s}}(0) = \mathbf{I}$ and $\mathbf{R}_{\bar{s}}(1) = \mathbf{J}_1$, respectively, where

$$\mathbf{J}_1 = \begin{bmatrix} 0 & 0 & & 0 \\ 1 & \ddots & \ddots & \\ & \ddots & \ddots & 0 \\ 0 & & 1 & 0 \end{bmatrix}.$$

Similarly, for the noise we have $\mathbf{R}_{\bar{e}}(0) = \sigma^2\mathbf{I}$ and $\mathbf{R}_{\bar{e}}(1) = \sigma^2\mathbf{J}_1$. Thus the covariances for the output vector $\bar{\mathbf{x}}$ is

$$\begin{aligned} \mathbf{R}_{\bar{x}}(0) &= \bar{\mathbf{H}}\bar{\mathbf{H}}^H + \sigma^2\mathbf{I}, \\ \mathbf{R}_{\bar{x}}(1) &= \bar{\mathbf{H}}\mathbf{J}_1\bar{\mathbf{H}}^H + \sigma^2\mathbf{J}_1. \end{aligned}$$

Our proposed method is based on the observation that

$$\bar{\mathbf{H}}\mathbf{J}_1\bar{\mathbf{H}}^H = \bar{\mathbf{H}}_F\bar{\mathbf{H}}_L^H,$$

where $\bar{\mathbf{H}}_F$ ($\bar{\mathbf{H}}_L$) is equal to the matrix $\bar{\mathbf{H}}$ except for the missing first (last) column. The first step is to estimate the noise variance σ^2 from the last eigenvalues of $\mathbf{R}_{\bar{x}}(0)$ and then remove the noise component from $\mathbf{R}_{\bar{x}}(1)$ by subtracting $\hat{\sigma}^2\mathbf{J}_1$, to obtain

$$\bar{\mathbf{R}}(1) = \mathbf{R}_{\bar{x}}(1) - \hat{\sigma}^2\mathbf{J}_1 = \bar{\mathbf{H}}_F\bar{\mathbf{H}}_L^H. \tag{8}$$

A key assumption is
[A.5] The matrices $\bar{\mathbf{H}}_F$, $\bar{\mathbf{H}}_L$ are tall, so

$$mW > L + W - 2. \tag{9}$$

Assumption A.5 holds for positive values of W iff $m > 1$, whence

$$W > (L-2)/(m-1). \tag{10}$$

Furthermore, we assume that
[A.6] The matrix $\bar{\mathbf{H}}$ has full column span.

The matrix $\bar{\mathbf{R}}(1)$ has size $(mW) \times (mW)$ but its rank is $L + W - 2 < mW$ since it is formed by the outer product (8) of two tall matrices with $L + W - 2$ columns. Therefore, the column span of $\bar{\mathbf{H}}_F$ and $\bar{\mathbf{H}}_L$ can be obtained from the SVD of $\bar{\mathbf{R}}(1)$:

$$\bar{\mathbf{R}}(1) = \mathbf{U}\Sigma\mathbf{V}^H. \tag{11}$$

Let $\bar{\mathbf{U}}$ and $\bar{\mathbf{V}}$ be the parts of the matrices $\mathbf{U}$ and $\mathbf{V}$ corresponding to the non-zero singular values of $\bar{\mathbf{R}}(1)$, then

$$\text{colspan}(\bar{\mathbf{H}}_\mathrm{F}) = \text{colspan}(\bar{\mathbf{U}}), \tag{12}$$
$$\text{colspan}(\bar{\mathbf{H}}_\mathrm{L}) = \text{colspan}(\bar{\mathbf{V}}). \tag{13}$$

$\bar{\mathbf{U}}$ and $\bar{\mathbf{V}}$ have size $(mW) \times (L + W - 2)$. It follows that there exist two square, invertible $(L+W-2) \times (L+W-2)$ matrices $\mathbf{T}_F$, $\mathbf{T}_L$, such that

$$\bar{\mathbf{H}}_F = \bar{\mathbf{U}}\mathbf{T}_F, \tag{14}$$
$$\bar{\mathbf{H}}_L = \bar{\mathbf{V}}\mathbf{T}_L. \tag{15}$$

Let us focus for a moment on the equation (14), since an identical discussion holds for (15) as well. We know that the matrix $\bar{\mathbf{H}}_F = [h_{ij}^F]$ has a special structure: it comes from a block-Toeplitz matrix (see (3)) by erasing the first column. Therefore

(a) there exists a set $\mathcal{Z}$, of index pairs (i, j) such that $h_{ij}^F = 0$, iff $(i, j) \in \mathcal{Z}$

(b) for all index pairs $(i, j) \notin \mathcal{Z}$ and $i > m$, we have $h_{ij}^F = h_{i-m,j-1}^F$.

We thus form the following system of linear constraints

$$\begin{array}{ll}\bar{\mathbf{u}}_i^H \mathbf{t}_{F,j} = 0, & \text{for all } (i,j) \in \mathcal{Z} \\ \bar{\mathbf{u}}_i^H \mathbf{t}_{F,j} - \bar{\mathbf{u}}_{i-m}^H \mathbf{t}_{F,j-1} = 0, & \text{for all } (i,j) \notin \mathcal{Z}, i > m\end{array} \quad (16)$$

where $\bar{\mathbf{u}}_i^H$ is the i-th row of $\bar{\mathbf{U}}$. We may rewrite (16) as

$$\Theta \bar{\mathbf{t}} = 0$$

where $\bar{\mathbf{t}}$ is the long vector $[\mathbf{t}_{F,1}^T \cdots \mathbf{t}_{F,L+W-2}^T]^T$ and the matrix Θ has size $C \times (L+W-2)^2$ with C being the number of equations in (16). Each row of Θ corresponds to one equation in the system and it contains the appropriate values $\bar{\mathbf{u}}_i^H$ and/or $-\mathbf{u}_{i-m}^H$ accordingly.

A careful counting of the indexes shows that system (16) comprises of $m(W-1)(L+W-1)$ constraints whereas the total number of unknowns $t_{F,i,j}$ is $(L+W-2)(L+W-2)$. By assumption A.5 we have more constraints than unknowns, therefore, system (16) admits at most one solution. However, according to (14), we know that one such solution exists thus the system has exactly one solution. In practice, instead of solving the system, we minimize the squared error

$$\min_{\bar{\mathbf{t}}} \{\bar{\mathbf{t}}^H \Theta \Theta^H \bar{\mathbf{t}}\}$$

by computing the eigenvalue decomposition of the matrix $\Theta\Theta^H$ and selecting $\bar{\mathbf{t}}$ to be the least eigenvector (i.e. the one associated with the smallest eigenvalue).

We are thus able to obtain the matrix $\mathbf{T}_F$ and consequently we can form the matrix $\bar{\mathbf{H}}_F$ using (14). Subsequently we can form an estimate of the filter $\hat{\mathbf{h}}^F(l)$, $l = 0, \cdots, L-1$, by averaging the block diagonals of $\bar{\mathbf{H}}_F$. An entirely similar procedure can be performed using $\mathbf{T}_L$ and $\bar{\mathbf{H}}_L$ obtaining a second estimate $\hat{\mathbf{h}}^L(l)$. The average of the two estimates forms the final estimate $\hat{\mathbf{h}}$.

4. SIMULATIONS

We compared our proposed method against a well established second order approach of Tong-Xu and Kailath [13] subsequently referred to as the TXK method. The randomly generated model has $m = 2$ outputs:

$$\begin{aligned} h_1(z) &= 1.6424 - 0.5825i + (1.0135 - 0.2525i)z^{-1} \\ &\quad +(-0.5314 + 1.3851i)z^{-2} \\ &\quad +(-1.4366 + 1.6141i)z^{-3} \\ h_2(z) &= 2.6369 + 0.6517i + (-0.1642 + 0.9942i)z^{-1} \\ &\quad +(1.7606 - 0.7157i)z^{-2} \\ &\quad +(0.0796 - 0.1402i)z^{-3} \end{aligned}$$

The performance index is the estimation accuracy measured by the normalized MSE defined as follows:

$$NMSE = \frac{\|\mathbf{h}' - \hat{\mathbf{h}}'\|^2}{\|\hat{\mathbf{h}}'\|^2}$$

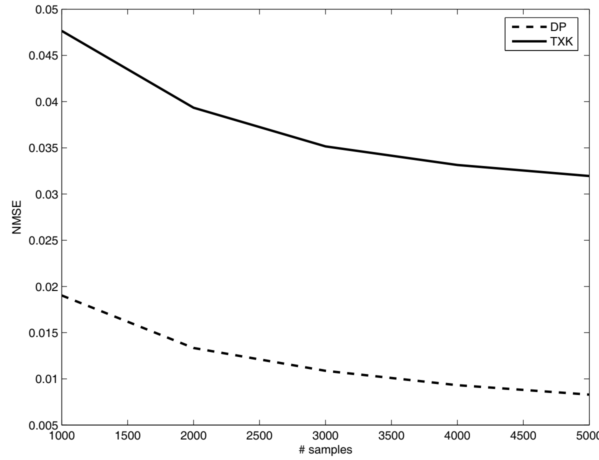

Fig. 1. Comparison between the estimation accuracy between the Diamantaras-Papadimitriou method (DP) and the Tong-Xu-Kailath method (TXK) as a function of the sample size N. The SNR level is constant at 20dB. The plots are mean values generated from 500 Monte Carlo experiments.

where $\mathbf{h}' = [h_1(0) \cdots h_m(0) \cdots h_1(L-1) \cdots h_m(L-1)]$ and $\hat{\mathbf{h}}' = [\hat{h}_1(0) \cdots \hat{h}_m(0) \cdots \hat{h}_1(L-1) \cdots \hat{h}_m(L-1)]$. The estimate $\hat{\mathbf{h}}'$ is scaled to best fit the original system vector $\mathbf{h}'$.

Figure 1 compares the proposed method against the TXK method for various sample sizes N keeping the noise level constant at $20dB$. The results are averaged over 500 Monte Carlo runs. The figure shows a clear advantage of the proposed method. The same conclusion is drawn from Figure 2 where the two methods are compared for different levels of SNR. The performance improvement is apparent especially in low SNR cases. In this experiment the number of samples is fixed at $N = 2000$, while again, the plots are averages after 500 Monte Carlo experiments.

5. CONCLUSION

In this paper we proposed a new analytical method for the blind identification of multichannel FIR systems. The method is conceptually simple as it boils down to the solution of a linear system of equations derived from second order statistics of the output signal. The method is quite fast since it does not involve iterative optimization procedures. Our simulations show that it outperforms classical second order methods both in terms of robustness to noise and in terms of performance with a fixed sample size.

6. REFERENCES

[1] Y. Inouye, and K. Hirano, "Cumulant-Based Blind Identification of Linear Multi-Input Multi-Output Systems

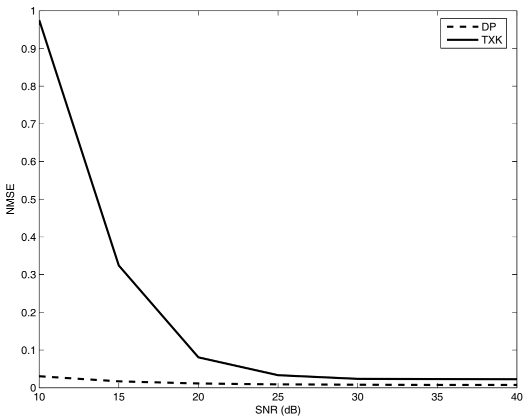

Fig. 2. Comparison between the estimation accuracy between the Diamantaras-Papadimitriou method (DP) and the Tong-Xu-Kailath method (TXK) as a function of the SNR. The number of samples in each experiment are 2000. The plots are mean values generated from 500 Monte Carlo experiments.

Driven by Colored Inputs", IEEE Trans. Signal Processing, vol. 45, no. 6, pp. 1543-1552, June 1997.

[2] J. K. Tugnait, "Adaptive blind separation of convolutive mixtures of independent linear signals," Signal Process., vol. 73, pp. 139-152, 1999.

[3] B. Chen and A. Petropulu, "Frequency domain blind MIMO system identification based on second and higher order statistics," IEEE Trans. Signal Processing, vol. 49, pp. 1677-1688, Aug. 2001.

[4] K. Abed-Meraim, P. Loubaton, and E. Moulines, "A subspace algorithm for certain blind identification problems," IEEE Trans. Inf. Theory, vol. 43, no. 2, pp. 499-511, Mar. 1997.

[5] Z. Ding, "Matrix outer-product decomposition method for blind multiple channel identification," IEEE Trans. Signal Process., vol. 45, no. 12, pp. 3053-3061, Dec. 1997.

[6] A. Gorokhov and P. Loubaton, "Blind identification of MIMO-FIR systems: A generalized linear prediction approach," Signal Process., vol. 73, pp. 105-124, Feb. 1999.

[7] E. Moulines, P. Duhamel, J. Cardoso, and S. Mayrargue, "Subspace methods for blind identification of multichannel FIR filters," IEEE Trans. Signal Proc., vol. 43, pp. 516-525, Feb. 1995.

[8] K. Abed-Meraim, E. Moulines, and P. Loubaton, "Prediction error method for second-order blind identification," IEEE Trans. Signal Processing, vol. 45, pp. 694-705, Mar. 1997.

[9] J. K. Tugnait, L. Tong, and Z. Ding, "Single-user channel estimation and equalization," IEEE Signal Processing Mag., vol. 17, pp. 16-28, May 2000.

[10] M. K. Tsatsanis and G. B. Giannakis, "Subspace methods for blind estimation of time-varying FIR channels," IEEE Trans. Signal Proc., vol. 45, pp. 3084-3093, Dec. 1997.

[11] H. Liu and G. B. Giannakis, "Deterministic approaches for blind equalization of time-varying channels with antenna arrays," IEEE Trans. Signal Processing, vol. 46, pp. 3003-3013, Nov. 1998.

[12] L. Tong, G. Xu, and T. Kailath, "A new approach to blind identification and equalization of multipath channels," in Proc. 25th Asilomar Conf., Pacific Grove, CA, 1991, pp. 856-860.

[13] L. Tong, G. Xu, and T. Kailath, "Blind Identification and Equalization Based on Second-Order Statistics: A Time Domain Approach", IEEE Trans. Information Theory, vol. 40, pp. 340-349, March 1994

[14] A. Ferreol, L. Albera, and P. Chevalier, "Fourth-Order Blind Identification of Underdetermined Mixtures of Sources (FOBIUM)", IEEE Trans. Signal Processing, vol. 53, pp. 1640-1653, May. 2005.