AN ALGEBRAIC APPROACH TO BLIND MIMO IDENTIFICATION

Lieven De Lathauwer, Bart De Moor and Joos Vandewalle

K.U.Leuven - E.E. Dept. (ESAT) - SISTA/COSIC
Kard. Mercierlaan 94, B-3001 Leuven (Heverlee), Belgium
tel: 32/16/321805 fax: 32/16/321970
email: Lieven.DeLathauwer@esat.kuleuven.ac.be

ABSTRACT

This paper deals with the blind identification of multiple-input multiple-output (MIMO) finite impulse response (FIR) filters. We limit ourselves to the case of 2 outputs and 2 inputs with values in $\mathbb{R}$. After a classical prewhitening, the remaining problem is the blind identification of a paraunitary filter. For this task, we derive a multilinear algebraic algorithm. This procedure is a generalization of the algorithm for independent component analysis (ICA) described in [1]. The performance is illustrated by means of some numerical experiments.

1. INTRODUCTION

We consider the following basic model for blind identification:

$$Y(n) = \sum_{k=0}^{K} \mathbf{H}(k) X(n-k),$$

in which $Y(n) \in \mathbb{R}^2$ is the vector of observations, $X(n) \in \mathbb{R}^2$ is the vector of inobservable sources, and $\{\mathbf{H}(k) \in \mathbb{R}^{2\times 2}\}$ are the Markov parameters of an invertible FIR mixing system. We assume that the sources (1) have statistically independent components, (2) are independent identically distributed, and (3) have at most one zero cumulant. In this paper we consider real-valued data; the extension to the complex case is discussed in [3].

This research was partially supported by the Flemish Government: (1) Research Council K.U.Leuven: Concerted Research Actions GOA-MIPS and GOA-MEFISTO-666, (2) the Fund for Scientific Research-Flanders (F.W.O.) projects G.0240.99 and G.0256.97, (3) the F.W.O. Research Communities ICCoS and ANMMM, and by the Belgian State, Prime Minister's Office - Federal Office for Scientific, Technical and Cultural Affairs: the Interuniversity Poles of Attraction Programmes IUAP P4-02 and IUAP P4-24. L. De Lathauwer is a post-doctoral researcher with the F.W.O. B. De Moor is a senior Research Associate with the F.W.O. and an Associate Professor with the K.U.Leuven. J. Vandewalle is a Full Professor with the K.U.Leuven. The scientific responsibility is assumed by the authors.

The goal is to find a demixing system

$$\mathbf{G}(z) = \sum_{k=0}^{K} \mathbf{G}(k) z^{-k},$$

in which z^{-1} is the backward-shift operator, such that

$$\mathbf{G}(z)\mathbf{H}(z) = \mathbf{\Lambda P},$$

in which $\mathbf{\Lambda}$ is an invertible constant diagonal matrix and $\mathbf{P}$ a permutation matrix.

Resorting only to second-order statistics, the best one can achieve is to linearly transform $Y(n)$ into a sequence $Z(n)$ such that

$$\mathrm{E}\{Z(n)Z^T(n-\tau)\} = \mathbf{I}\,\delta(\tau),$$

in which $\mathbf{I}$ is the identity matrix and $\delta(\tau)$ the Kronecker delta. Full identification is not possible since, as is well-known, second-order statistics do not carry information about the phase. In the rest of the paper we will assume that this classical prewhitening has been carried out.

After the prewhitening step, the remaining mixing system is a paraunitary filter $F(z)$, i.e.,

$$\mathbf{F}(z)\mathbf{F}^H(1/z^*) = \mathbf{I},$$

with parametrization

$$\mathbf{F}(z) = \mathbf{Q}(\theta_K)\cdot\Omega\cdot\ldots\cdot\Omega\cdot\mathbf{Q}(\theta_2)\cdot\Omega\cdot\mathbf{Q}(\theta_1), \qquad (1)$$

where

$$\mathbf{Q}(\theta_k) = \begin{pmatrix} \cos\theta_k & -\sin\theta_k \\ \sin\theta_k & \cos\theta_k \end{pmatrix},$$

$$\Omega = \begin{pmatrix} 1 & 0 \\ 0 & z^{-1} \end{pmatrix}.$$

This paraunitary filter can be blindly identified by means of the information carried by the higher-order cumulants of $Z(n)$. In this paper we consider fourth-order cumulants; the third-order case can be dealt with in a similar way.

In Section 2 we discuss the principles of an algebraic approach for the blind identification of the paraunitary filter. The approach is a generalization for convolutive mixtures of the well-known algorithm [1]. In Section 3 we discuss a cheaper implementation of the technique proposed in Section 2. In Section 4 the performance is illustrated by means of a numerical experiment. Section 5 is the conclusion.

2. DYNAMIC INDEPENDENT COMPONENT ANALYSIS

The effect of a multiplication with Ω in Eq. (1) is that the region over which the fourth-order cumulant $\mathcal{C}_Z(\tau_1, \tau_2, \tau_3)$ is defined, is expanded with an extra delay. For example, if we consider the filter $\mathbf{Q}(\theta_k) \cdot \ldots \cdot \Omega \cdot \mathbf{Q}(\theta_2) \cdot \Omega \cdot \mathbf{Q}(\theta_1)$ and the subsequent multiplication with Ω, then the boundary $-k+1 \leqslant \tau_1 \leqslant k-1$ is replaced by the boundary $-k \leqslant \tau_1 \leqslant k$. The effect of a multiplication with $\mathbf{Q}(\theta_r)$ is an orthogonal transformation of the column space, row space, ... of all the fourth-order cumulant tensors, corresponding to different time lags. For example, let, for a specific choice of τ_1, τ_2, τ_3, $\mathcal{C}_Z(\tau_1, \tau_2, \tau_3)$ be abbreviated to $\mathcal{C}$, and let us briefly denote $\mathbf{Q}(\theta_r)$ by $\mathbf{Q}$, then the tensor $\mathcal{C}$ is replaced by the tensor $\mathcal{C}'$, defined by the element-wise equation

$$c'_{i_1 i_2 i_3 i_4} = \sum_{j_1, j_2, j_3, j_4} q_{i_1 j_1} q_{i_2 j_2} q_{i_3 j_3} q_{i_4 j_4} c_{j_1 j_2 j_3 j_4}.$$

A crucial observation is that, after multiplication with Ω, the new cumulant tensors that are formed in the borders of the (τ_1, τ_2, τ_3)-domain, are not full tensors; instead, most of their entries are zero. Let us denote the cumulant of the sequence $\tilde{Z}_k(n)$, defined by

$$\tilde{Z}_k(z) = \mathbf{Q}(\theta_k) \cdot \ldots \cdot \Omega \cdot \mathbf{Q}(\theta_2) \cdot \Omega \cdot \mathbf{Q}(\theta_1) \cdot X(z),$$

by $\mathcal{C}_{\tilde{Z}}^{(k)}(\tau_1, \tau_2, \tau_3)$. Let us further denote the cumulant of the sequence $Z_k(n)$, defined by

$$Z_k(z) = \Omega \cdot \tilde{Z}_k(z),$$

by $\mathcal{C}_Z^{(k)}(\tau_1, \tau_2, \tau_3)$. For the principal domain, defined by $0 \leqslant \tau_1 \leqslant \tau_2 \leqslant \tau_3$, the multiplication with Ω results in the following substitutions:

$$(\mathcal{C}_Z^{(k)})_{1222}(k,k,k) \leftarrow (\mathcal{C}_{\tilde{Z}}^{(k)})_{1222}(k-1,k-1,k-1) \quad (2)$$

$$(\mathcal{C}_Z^{(k)})_{1222}(\tau_1,k,k) \leftarrow (\mathcal{C}_{\tilde{Z}}^{(k)})_{1222}(\tau_1-1,k-1,k-1) \quad (0<\tau_1<k), \quad (3)$$

$$(\mathcal{C}_Z^{(k)})_{1122}(\tau_1,k,k) \leftarrow (\mathcal{C}_{\tilde{Z}}^{(k)})_{1122}(\tau_1,k-1,k-1) \quad (0<\tau_1<k), \quad (4)$$

$$(\mathcal{C}_Z^{(k)})_{1122}(0,k,k) \leftarrow (\mathcal{C}_{\tilde{Z}}^{(k)})_{1122}(0,k-1,k-1), \quad (5)$$

$$(\mathcal{C}_Z^{(k)})_{1222}(\tau_1,\tau_2,k) \leftarrow (\mathcal{C}_{\tilde{Z}}^{(k)})_{1222}(\tau_1-1,\tau_2-1,k-1) \quad (0<\tau_1 \leqslant \tau_2<k), \quad (6)$$

$$(\mathcal{C}_Z^{(k)})_{1122}(\tau_1,\tau_2,k) \leftarrow (\mathcal{C}_{\tilde{Z}}^{(k)})_{1122}(\tau_1,\tau_2-1,k-1) \quad (0<\tau_1 \leqslant \tau_2<k), \quad (7)$$

$$(\mathcal{C}_Z^{(k)})_{1212}(\tau_1,\tau_2,k) \leftarrow (\mathcal{C}_{\tilde{Z}}^{(k)})_{1212}(\tau_1-1,\tau_2,k-1) \quad (0<\tau_1 \leqslant \tau_2<k), \quad (8)$$

$$(\mathcal{C}_Z^{(k)})_{1112}(\tau_1,\tau_2,k) \leftarrow (\mathcal{C}_{\tilde{Z}}^{(k)})_{1112}(\tau_1,\tau_2,k-1) \quad (0<\tau_1 \leqslant \tau_2<k), \quad (9)$$

$$(\mathcal{C}_Z^{(k)})_{1122}(0,\tau_2,k) \leftarrow (\mathcal{C}_{\tilde{Z}}^{(k)})_{1122}(0,\tau_2-1,k-1) \quad (0<\tau_2<k), \quad (10)$$

$$(\mathcal{C}_Z^{(k)})_{1112}(0,\tau_2,k) \leftarrow (\mathcal{C}_{\tilde{Z}}^{(k)})_{1112}(0,\tau_2,k-1) \quad (0<\tau_2<k), \quad (11)$$

$$(\mathcal{C}_Z^{(k)})_{1112}(0,0,k) \leftarrow (\mathcal{C}_{\tilde{Z}}^{(k)})_{1112}(0,0,k-1). \quad (12)$$

At the same time, in the slices adjacent to the border slices, zeros are introduced; the corresponding entries are not explicitly enumerated here, for brevity. One could now estimate $\mathbf{Q}(\theta_k)$ as the rotation that minimizes the numerical values for the latter entries and maximizes the entries enumerated in (2-12), both in a quadratic sense. This approach follows the spirit of [1], where for the determination of a static mixture the off-diagonal cumulant tensor entries, which theoretically vanish, are minimized, while the diagonal entries are maximized.

The overall procedure starts then with the estimation of $\mathbf{Q}(\theta_K)$, followed by an inverse time shift, after which $\mathbf{Q}(\theta_{K-1})$ is estimated, again followed by an inverse time shift, etc. In other words, in each step a layer of sequence (1) is peeled off, until finally only the estimation of $\mathbf{Q}(\theta_1)$ remains, which is a classical blind source separation problem.

A disadavantage of this approach is that the estimation of a rotation matrix $\mathbf{Q}(\theta_k)$ is based on the cumulants of the outer slices of the (τ_1, τ_2, τ_3)-domain, which form a decreasing part of the domain of support as k increases. Note that the sum of the squared Frobenius-norms of all the cumulant tensors in the domain of support is constant, as we consider only orthogonal transformations and time shifts, so that the outer slices usually contain less information as k increases. In other words, the longer the filter length, the harder the job.

After multiplication with $\mathbf{Q}(\theta_k)$ the entries under consideration are homogeneous polynomials of degree 4 in $\sin\theta_k$ and $\cos\theta_k$. For the determination of the optimal rotation angle the derivative of a linear combination of squares of these entries has to be rooted, which leads to the determination of the roots of a polynomial of degree 8 in $\tan\theta_k$ and selection of the best root. If $k = 1$, then the expressions can be simplified; in this case it is sufficient to compute the roots of a polynomial of degree 4 [1]. (These problems can also be considered in terms of the best rank-1 approximation of a higher-order tensor [2, 4].) This approach would rightfully be called "Dynamic ICA".

3. ALGORITHM

In this section we will present a simplified, cheaper version of the scheme described in the previous section. Denote $\mathbf{Q}(\theta_k) = (Q_1\ Q_2)$, and let α, β, γ be arbitary values in $\{1, 2\}$. Then we have, in the absence of noise, the following proportionalities from (2-12):

$$(\mathcal{C}_Z^{(k)})_{:,\alpha,\beta,\gamma}(k,k,k) \sim Q_1, \quad (13)$$
$$(\mathcal{C}_Z^{(k)})_{\alpha,:,\beta,\gamma}(k,k,k) \sim Q_2, \quad (14)$$
$$(\mathcal{C}_Z^{(k)})_{:,\alpha,\beta,\gamma}(\tau_1,k,k) \sim Q_1 \quad (0 < \tau_1 < k), \quad (15)$$
$$(\mathcal{C}_Z^{(k)})_{\alpha,\beta,:,\gamma}(\tau_1,k,k) \sim Q_2 \quad (0 < \tau_1 < k), \quad (16)$$
$$(\mathcal{C}_Z^{(k)})_{:,\alpha,\beta,\gamma}(0,k,k) \sim Q_1, \quad (17)$$
$$(\mathcal{C}_Z^{(k)})_{\alpha,\beta,:,\gamma}(0,k,k) \sim Q_2 \quad (18)$$
$$(\mathcal{C}_Z^{(k)})_{:,\alpha,\beta,\gamma}(\tau_1,\tau_2,k) \sim Q_1 \quad (0 < \tau_1 \leqslant \tau_2 < k), \quad (19)$$
$$(\mathcal{C}_Z^{(k)})_{\alpha,\beta,\gamma,:}(\tau_1,\tau_2,k) \sim Q_2 \quad (0 < \tau_1 \leqslant \tau_2 < k), \quad (20)$$
$$(\mathcal{C}_Z^{(k)})_{:,\alpha,\beta,\gamma}(0,\tau_2,k) \sim Q_1 \quad (0 < \tau_2 < k), \quad (21)$$
$$(\mathcal{C}_Z^{(k)})_{\alpha,\beta,\gamma,:}(0,\tau_2,k) \sim Q_2 \quad (0 < \tau_2 < k), \quad (22)$$
$$(\mathcal{C}_Z^{(k)})_{:,\alpha,\beta,\gamma}(0,0,k) \sim Q_1, \quad (23)$$
$$(\mathcal{C}_Z^{(k)})_{\alpha,\beta,\gamma,:}(0,0,k) \sim Q_2, \quad (24)$$

in which we used MATLAB notation. These proportionalities can be rewritten as an homogeneous set of linear equations, of which Q_1 is the theoretical solution. For example, if we have that

$$\begin{pmatrix} a \\ b \end{pmatrix} \sim Q_1,$$

then we obtain the equation

$$\begin{pmatrix} -b & a \end{pmatrix} \cdot Q_1 = 0.$$

Proportionalities of the form

$$\begin{pmatrix} a \\ b \end{pmatrix} \sim Q_2,$$

can be rewritten as

$$\begin{pmatrix} a & b \end{pmatrix} \cdot Q_1 = 0.$$

The influence of noise is mitigated by the fact that this set is largely overdetermined. Hence, Q_1 can be found as the second right singular vector of a $(P \times 2)$-matrix, with $P \gg 2$, which is a quadratic problem; $\mathbf{Q}(\theta_k)$ follows immediately (up to the sign, which is irrelevant).

Finally, it is worth mentioning that, theoretically, $\mathbf{Q}(\theta_k)$ can be determined by resorting to only one of the relations (13-24). This means that in the absence of noise the blind identification problem can be solved as an ordinary system of linear equations.

4. SIMULATIONS

We consider two sources drawn from a binary distribution with an equal probability of signal values +1 and −1. The filter length is 2. The data length is 50. We assume a perfect prewhitening. $\sin\theta_1$, $\sin\theta_2$ and $\sin\theta_3$ are drawn from a uniform distribution over $(-1, +1]$. The noise is additive and Gaussian. The experiment consists of 100 Monte Carlo runs.

In Fig. 1 we plot the Root Mean Square Error (RMSE) between the true sources and their estimates as a function of the Signal-to-Noise Ratio (SNR):

$$\mathrm{RMSE} = \sqrt{\mathrm{E}\{\|X - \hat{X}\|^2\}/2},$$

in which $\hat{X}$ is the estimate of the true source vector X, and in which the factor 2 is a normalization due to the fact that there are two components; we considered the optimal ordering and sign of the source estimates. Next, we make use of the prior knowledge of the source distribution: source estimate values $\hat{x}$ are mapped to $\mathrm{sign}(\hat{x})$. In Fig. 2 the corresponding Bit Error Rate (BER) is shown, i.e., the number of errors divided by the total number of source values. We see that for an SNR $\geqslant$ 10 dB nearly all the source values are reconstructed correctly, despite the fact that only 50 samples are available. An SNR of 5 dB is sufficient for a correct reconstruction of 95 % of the source values.

5. CONCLUSION

We have proposed a new line of approach to the problem of MIMO blind identification / deconvolution. The core of our contribution is the derivation of a (multi)linear algebraic technique for the blind identification of a paraunitary filter. Such a filter is the dynamic equivalent of the orthogonal (unitary) matrix that remains unidentified after a prewhitening in the case of static mixtures. We developed a similar reasoning as in the well-known ICA paper [1], but for convolutive mixtures. In this paper we only considered the two-input two-output case. The results can be generalized to deal with N-input N-output systems; this is a topic of current research.

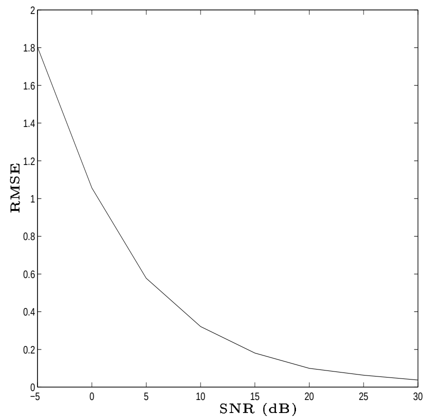

Figure 1: RMSE as a function of SNR

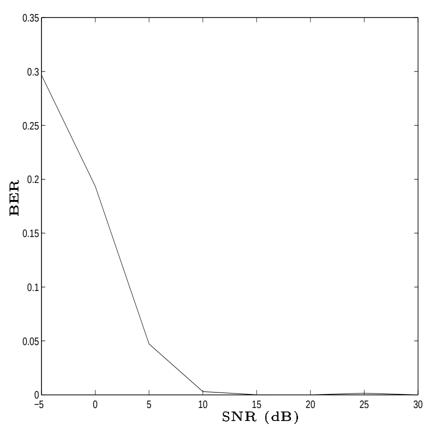

Figure 2: BER as a function of SNR

6. REFERENCES

[1] P. Comon, "Independent Component Analysis, A New Concept?" *Signal Processing*, Special Issue *Higher Order Statistics*, Vol. 36, No. 3, April 1994, pp. 287-314.

[2] L. De Lathauwer, P. Comon, B. De Moor, J. Vandewalle, "Higher-Order Power Method - Application in Independent Component Analysis", *Proc. Int. Symp. on Nonlinear Theory and Its Applications (NOLTA'95)*, Las Vegas, U.S.A., Dec. 10-14, 1995, Vol. 1, pp. 91-96.

[3] L. De Lathauwer, B. De Moor, J. Vandewalle, "An Algebraic Approach to the Blind Identification of Para-Unitary Filters", *IEEE Wireless Communications and Networking Conference (WCNC 2000)*, Chicago, USA, Sept. 23-28, 2000.

[4] L. De Lathauwer, B. De Moor, J. Vandewalle, "On the Best Rank-1 and Rank-$(R_1, R_2, \ldots, R_N)$ Approximation of Higher-Order Tensors", Tech. Report No. 97-75, ESAT/SISTA, K.U.Leuven. To appear in: *SIAM J. Matrix Anal. Appl.*