3rd Int. Conf. on Independent Component Analysis, San Diego, 9-13 December, 2001

BLIND SEPARATION OF CONVOLUTIVE MIXTURES A CONTRAST-BASED JOINT DIAGONALIZATION APPROACH

Pierre Comon[1], *Éric Moreau*[2], *and Ludwig Rota*[1]

(1) I3S, Algorithmes-Euclide-B, BP 121, F-06903 Sophia-Antipolis Cedex, France
(2) SIS, ISITV, Univ. Toulon, BP 56, F-83162 La Valette du Var Cedex, France
`comon@i3s.unice.fr` `moreau@isitv.univ-tln.fr`

ABSTRACT

Blind Separation of convolutive mixtures and Blind Equalization of Multiple-Input Multiple-Output (MIMO) channels are two different ways of naming the same problem, which we address here. The numerical algorithm, subsequently presented in detail, is based on theoretical results on contrasts recently published by the authors [1]. This algorithm consists of Partial Approximate Joint Diagonalization (PAJOD) of several matrices, containing some values of output cumulant multi-correlations.

1. INTRODUCTION

Blind equalization (that is, without observing the inputs) of linear time-invariant systems has been studied extensively during the last decade. Single Input Single Output (SISO) equalization often requires High-Order Statistics (HOS) [2] [3] [4]; this can be implicit through constant modulus [5] [6] or constant power [7] criteria. For multiple channels (SIMO or MIMO), HOS can be used [8] [9], but second-order statistics can suffice, provided mild identifiability conditions are satisfied, but a number of limitations have been identified [10] [11] [12] [13] [14]; see [15] and references therein. Because of their robustness to hypotheses, HOS-based methods remain very attractive.

The case of static mixtures (as opposed to convolutive) has also retained a lot of attention, because its simpler form allows a deeper treatment, but often requires resorting to HOS. Put in simple words, one can say that the problem can be viewed as diagonalizing a tensor [16] [17], but can be addressed by diagonalizing approximatively a subset of matrix slices [18]. The latter algorithms are efficient when applied to short data records (or fast varying channels) because they are of block type (*i.e.*. off-line), but on-line algorithms have also been devised [19].

Our main contribution consists of a block algorithm dedicated to blind MIMO equalization. This algorithm has been shown to maximize a well-defined *contrast* [1] [20], as pointed out in section 3. On-line versions of this algorithm would be easy to implement, and are not studied in this paper. Instead, we concentrate on the algorithm description, and on its performances on very short data records (*e.g.*, 200 to 400 symbols).

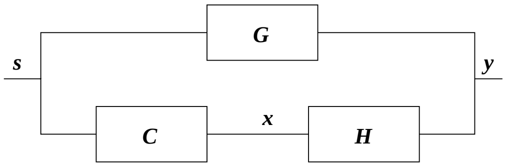

Fig. 1. Observation $\boldsymbol{x}$ is equalized by $\boldsymbol{H}$; the global system is denoted $\boldsymbol{G}$.

2. PROBLEM AND NOTATION

Consider the following linear time-invariant invertible system:

$$\boldsymbol{x}(n) = \sum_{k=-\infty}^{\infty} \boldsymbol{C}(k)\boldsymbol{s}(n-k) \tag{1}$$

where $\boldsymbol{s}(n)$ denotes the N-dimensional source vector, whose components $s_i(n)$ are independent, $\boldsymbol{x}(n)$ is the N-dimensional observation, and $\{\boldsymbol{C}\} \stackrel{\text{def}}{=} \{\boldsymbol{C}(n), n \in \mathbb{Z}\}$ denotes the $N \times N$ impulse response matrix sequence. For convenience, vectors and matrices are denoted with bold lowercase and bold uppercase letters, respectively. The problem consists of finding a filter $\{\boldsymbol{H}\} \stackrel{\text{def}}{=} \{\boldsymbol{H}(n), n \in \mathbb{Z}\}$ from the sole observation of the channel outputs, $\boldsymbol{x}(n)$, aiming at delivering an estimate $\boldsymbol{y}(n)$ of the inputs $\boldsymbol{s}(n)$.

The following hypotheses are assumed:

H1. Sources $s_i(n)$, $i \in \{1,\ldots,N\}$ are mutually independent i.i.d. zero-mean processes, with unit variance.

H2. $\boldsymbol{s}(n)$ is stationary up to the considered order, $r \in \mathbb{N}^*$, *i.e.* $\forall i \in \{1,\ldots,N\}$, the order-$r$ cumulant, $\mathsf{Cum}\,[s_i(n),\ldots,s_i(n)]$, does not depend on n, and will be denoted $\mathsf{C}_r[s_i]$.

H3. At most one source has a zero marginal cumulant of order r.

H4. The global transfer matrix, $\boldsymbol{G}(z) = \boldsymbol{H}(z)\boldsymbol{C}(z)$, satisfies the property

$$\boldsymbol{G}(z)\boldsymbol{G}^{\mathrm{H}}(1/z^*) = \boldsymbol{I}$$

where $\boldsymbol{I}$ denotes the $N \times N$ identity matrix; in other words, $\boldsymbol{G}(z)$ is *para-unitary*.

Remark 1. More generally, if sources are non i.i.d. linear processes, our approach holds valid. It suffices to still assume (**H1**) in a first stage in order to equalize the channel, and to extract the original sources in a second stage by linear regression between each equalizer output and the observations. In fact, the equalizer outputs are the driving processes of the sources.

Remark 2. The Hypothesis **H4** is not restrictive. Indeed, one can always whiten the observations (in a non unique manner), by using a filter that factorizes the second-order power spectrum.

3. CONTRASTS

We briefly report in this section the results stated by the authors in [1], showing how contrast-based blind MIMO equalization can be posed in terms of a Partial Joint Approximate Diagonalization (PAJOD) of a set of cumulant matrices. To start with, define:

$$\begin{aligned}&\mathsf{C}^{\boldsymbol{y}}_{p,r}[i,\boldsymbol{j},\boldsymbol{\ell}] = \\ &\mathsf{Cum}[\underbrace{y_i(n),\ldots,y_i(n)}_{p \text{ terms}},\underbrace{y_{j_1}(n-\ell_1),\ldots,y_{j_q}(n-\ell_q)}_{q=r-p \text{ terms}}]\end{aligned} \tag{2}$$

where $r = p+q$ is the cumulant order, $\boldsymbol{j} = (j_1,\ldots,j_q)$, and $\boldsymbol{\ell} = (\ell_1,\ldots,\ell_q)$. Then we can prove the two propositions below [1]:

Proposition 1 *Let r, p and q be three integers such that $r \geq 3$, $2 \leq p \leq r$ and $q = r-p$, then the functional*

$$\mathcal{J}_{p,r}(\boldsymbol{y}) = \sum_{i=1}^{N}\sum_{\boldsymbol{j}\in\mathsf{J}}\sum_{\boldsymbol{\ell}\in\mathsf{L}}\left|\mathsf{C}^{\boldsymbol{y}}_{p,r}[i,\boldsymbol{j},\boldsymbol{\ell}]\right|^2 \tag{3}$$

is a contrast when observations $\boldsymbol{x}(n)$, and hence the outputs $\boldsymbol{y}(n)$ of the para-unitary equalizer, are standardized (i.e., second-order white with unit covariance).

Also define the cumulant tensor of observations:

$$\begin{aligned}\boldsymbol{T}_{\boldsymbol{a},\boldsymbol{b}}(\boldsymbol{\alpha},\boldsymbol{\beta}) = \mathsf{Cum}[&x_{a_1}(n-\alpha_1),\\ &x_{a_2}(n-\alpha_2),\, x_{b_1}(n-\beta_1),\ldots,x_{b_q}(n-\beta_q)]\end{aligned} \tag{4}$$

where $\boldsymbol{a}$ and $\boldsymbol{\alpha}$ are vectors of size 2, and $\boldsymbol{b}$ and $\boldsymbol{\beta}$ are of size $q = r-2$, the entries of $\boldsymbol{a}$ and $\boldsymbol{b}$ belonging to $\{1,...,N\}$. If L delays are considered, the entries of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ belong to $\{0,...,L-1\}$.

This tensor can be stored in a set of $NL \times NL$ matrices, denoted $\boldsymbol{M}(\boldsymbol{b},\boldsymbol{\beta})$. for any fixed $(\boldsymbol{b},\boldsymbol{\beta})$, the entries of these matrices are given by:

$$\boldsymbol{M}_{\eta\mu}(\boldsymbol{b},\boldsymbol{\beta}) = \boldsymbol{T}_{\boldsymbol{a},\boldsymbol{b}}(\boldsymbol{\alpha},\boldsymbol{\beta}), \tag{5}$$

with

$$\eta = \alpha_1\,N + a_1,\ \mu = \alpha_2\,N + a_2$$

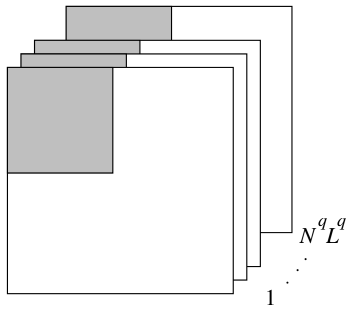

Fig. 2. The semi-unitary matrix $\mathcal{H}$ aims at diagonalizing jointly the $N \times N$ leading submatrices.

Proposition 2 *The contrast $\mathcal{J}_{2,r}(\boldsymbol{y})$ can be rewritten as a PAJOD criterion of a set of N^qL^q matrices:*

$$\mathcal{J}_{2,r}(\boldsymbol{y}) = \sum_{\boldsymbol{b}}\sum_{\boldsymbol{\beta}}||\mathbf{Diag}\{\mathcal{H}^{\mathrm{H}}\,\boldsymbol{M}(\boldsymbol{b},\boldsymbol{\beta})\,\mathcal{H}\}||^2 \tag{6}$$

where $\mathcal{H}$ is semi-unitary, i.e., satisfies $\mathcal{H}^{\mathrm{H}}\mathcal{H} = \boldsymbol{I}$.

See [1] for the proofs and further details.

Remark 3. The above criterion differs from that proposed in [18] in several respects: (i) the matrices $\boldsymbol{M}(\boldsymbol{b},\boldsymbol{\beta})$ are built differently, (ii) the matrix sought for is not square unitary but rectangular, which involves quite different calculations, as will be subsequently seen.

From now on, we shall assume that the channel length L is known, and that the equalizer has the same length. The robustness with respect to this assumption will be investigated in a companion paper.

4. NUMERICAL ALGORITHM

Take the particular case where cumulants of order $r = 4$ are used, and choose $p = q = 2$. Vectors $\boldsymbol{a}$, $\boldsymbol{\alpha}$, $\boldsymbol{b}$, and $\boldsymbol{\beta}$ are thus of size 2. The propositions of the previous section teach us that a semi-unitary matrix, $\mathcal{H}$, of size $NL \times N$, must be found, which should diagonalize approximately and jointly the set of N^2L^2 matrices, $\boldsymbol{M}(b_1, b_2, \beta_1, \beta_2)$. Each of these matrices is of size $NL \times NL$. The goal is to maximize the sum of the squared moduli of the N first diagonal entries of the N^2L^2 matrices, as depicted in figure 2.

4.1. Jacobi sweeping

In order to reach this goal, one looks for a $NL \times NL$ unitary matrix, $\boldsymbol{U}$, whose $\mathcal{H}$ will be the the leading $NL \times N$ submatrix. This unitary matrix can be built by accumulating Givens rotations, as proposed in the Jacobi algorithm [16] [21]:

$$\boldsymbol{U} = \prod_{1 \leq i < j \leq NL} \Theta[i, j],$$

where $\Theta[i, j]$ coincides with the identity matrix except for 4 entries, namely:

$$\Theta[i, j]_{ii} = \Theta[i, j]_{jj} = \cos(\theta[i, j])$$

and

$$\Theta[i, j]_{ji} = -\Theta[i, j]^*_{ij} = \sin(\theta[i, j])\, e^{\jmath\psi[i,j]}.$$

This rotation can indeed always be imposed to have a real cosine [21] [16]. The cosine, c, and sine, s, must be determined so as to maximize, successively for every pair $[i, j]$:

$$\mathcal{J}_{2,4} = \sum_{\boldsymbol{b},\boldsymbol{\beta}} \sum_{k=1}^{N} \left| \sum_{\eta,\mu=1}^{NL} \Theta^*_{\eta k}[i, j] \Theta_{\mu k}[i, j] M_{\eta\mu}(\boldsymbol{b}, \boldsymbol{\beta}) \right|^2 \quad (7)$$

4.2. Processing every pair

As emphasized earlier, indices $[i, j]$ do not describe all possible pairs from the set $\{1, ..., NL\}^2$. In fact, since $k \leq N$, it suffices that $i \leq N$; in addition, we also have that $i < j$. As a consequence, two cases must be distinguished, depending on the fact that $j \leq N$ or not.

In the two cases, we have to find the roots of a polynomial of degree 4. But in the first case, with the help of a change of variables, this rooting can be converted into the solving of two trinomials of degree 2, as in [22]. This transformation is not possible in the second case, and the (still analytical) rooting of the fourth degree polynomial is mandatory:

- **Case** $j \leq N$: this is the classical one [22]. One maximizes the sum of the 2 diagonal terms on which one has some action:

  $$\mathcal{J}_{2,4} = \textstyle\sum_{\boldsymbol{b},\boldsymbol{\beta}} |c^2 M_{ii} + cs^* M_{ji} + csM_{ij} + ss^* M_{jj}|^2 \\ +|ss^* M_{ii} - cs^* M_{ji} - csM_{ij} + c^2 M_{jj}|^2$$

  In the real case, stationary values in $t = \tan\theta$ are the roots of the degree-4 polynomial:

  $$t^4 + \rho t^3 - 6t^2 - \rho t + 1 = 0$$

  where

  $$\rho = \frac{\sum_{\boldsymbol{b}\,\boldsymbol{\beta}} (M_{ii} - M_{jj})^2 - 4M^2_{ij}}{\sum_{\boldsymbol{b}\,\boldsymbol{\beta}} (M_{ii} - M_{jj}) M_{ij}}$$

  Because of symmetries in the coefficients, one can instead root the trinomial in variable $Z = \tan 2\theta$:

  $$Z^2 + \frac{\rho}{2} Z + 1 = 0$$

  and go back to the angle tangent by rooting in a second stage $Zt^2 + 2t - Z = 0$.

- **Case** $j > N$: here only one diagonal term should be maximized:

  $$\mathcal{J}_{2,4} = \sum_{\boldsymbol{b},\boldsymbol{\beta}} |c^2 M_{ii} + cs^* M_{ji} + csM_{ij} + ss^* M_{jj}|^2$$

  In the real case, this amounts to solving the degree-4 polynomial in $t = \tan\theta$:

  $$\rho_3 t^4 + (\rho_2 - \rho_4) t^3 + 3(\rho_1 - \rho_3) t^2 + (\rho_0 - \rho_2) t - \rho_1 = 0$$

  with

  $$\begin{aligned} \rho_0 &= \textstyle\sum M^2_{ii} \\ \rho_1 &= \textstyle\sum M_{ii} M_{ij} \\ \rho_2 &= \textstyle\sum M_{ii} M_{jj} + 2M^2_{ij} \\ \rho_3 &= \textstyle\sum M_{ij} M_{jj} \\ \rho_4 &= \textstyle\sum M^2_{jj} \end{aligned}$$

  The selection of the best angle among the four candidates can be done by simply calculating the contrast value at these four points, for instance using

  $$\mathcal{J}_{2,4} = (1 + t^2)^{-2} (\rho_0 + 4\rho_1 t + 2\rho_2 t^2 + 4\rho_3 t^3 + \rho_4 t^4)$$

5. PERFORMANCES

5.1. Working example

One considers a Finite Impulse Response (FIR) real mixture of length $L = 3$ of $N = 2$ real white processes. Cumulants of order $r = p + q = 4$ are chosen with $p = 2$. Thus, there are $N^q L^q = 36$ square matrices, each of size $NL = 6$, and the goal is to jointly and approximately diagonalize their 2×2 leading matrix by congruent transform. With this goal, a real orthogonal 6×6 matrix, $\boldsymbol{U}$ is estimated. Matrix $\mathcal{H}$ corresponds to the two first columns of $\boldsymbol{U}$.

The channel is para-unitary, to preserve second-order whiteness as explained in section 2. It has been generated as follows:

$$\boldsymbol{C}(z) = \boldsymbol{R}(\phi_1) \cdot \boldsymbol{Z} \cdot \boldsymbol{R}(\phi_2) \cdot \boldsymbol{Z} \cdot \boldsymbol{R}(\phi_3),$$

where

$$\boldsymbol{Z} = \begin{pmatrix} 1 & 0 \\ 0 & z^{-1} \end{pmatrix},$$

and

$$\boldsymbol{R}(\phi) = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}$$

Note the absence of phase terms, because real channels are desired. Because of the 3 free parameters above, we have some control on the location of zeros of the 4 length-3 SISO channels. Here, we have chosen $\phi_1 = 52^o$, $\phi_2 = 53^o$, and $\phi_3 = 70^o$. As a result, the zeros of $C_{12}(z)$ are inside the unit disk, and those of $C_{21}(z)$ are outside; $C_{11}(z)$ and $C_{22}(z)$ each have one zero inside and one outside, as shown in figure 3. So, $\boldsymbol{C}(z)$ has a stable inverse, whereas the components $C_{12}(z)$ and $C_{21}(z)$ have not.

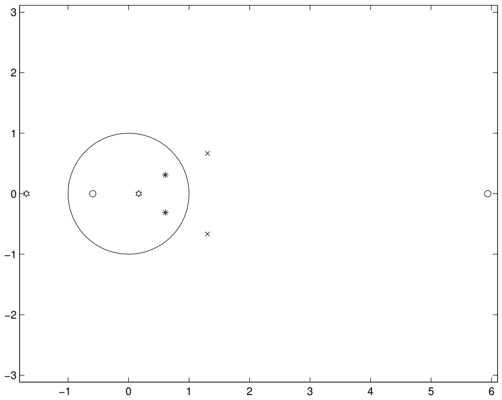

Fig. 3. Zeros of the 4 channels in $\boldsymbol{C}(z)$.

5.2. Performance criteria

When evaluating performances of MIMO equalizers, a difficulty to overcome stems from inherent indeterminacies. In fact, equalizer $\boldsymbol{H}(z)$, and hence global filter $\boldsymbol{G}(z)$, can be estimated only up to a multiplicative matrix of the form $\boldsymbol{D}(z) = \boldsymbol{\Lambda}(z)\boldsymbol{P}$, as defined in section 2.

5.2.1. Distance criterion

Let the global transfer function

$$\boldsymbol{G}(z) = \sum_{n=0}^{2L-1} \boldsymbol{G}(n) z^{-n}.$$

One can decide to store matrices $\boldsymbol{G}(n)$ in a $N \times N(2L-1)$ array, $\mathcal{G}$, by merely stacking the matrices one after the other. Then finding the best matrix $\boldsymbol{D}(z)$ amounts to searching every row of $\mathcal{G}$ for the entry of largest modulus, under the constraint that their column index are different modulo N.

Let us now explain how this is done in our example where $N = 2$ and $L = 3$:

- Case 1: search for column j_1 (resp. j_2) containing the entry of largest modulus, $\mathcal{G}_{1,j_1}$ (resp. $\mathcal{G}_{2,j_2}$), among the entries of row 1 (resp. 2) of odd column index (resp. even). Normalize row 1 (resp. 2) of $\mathcal{G}$ by $\mathcal{G}_{1,j_1}$ (resp. $\mathcal{G}_{2,j_2}$). Compute the Froebenius distance between matrix $\bar{\mathcal{G}}$ obtained this way and matrix $\mathcal{D}$, of same size, containing two 1's at locations $(1, j_1)$ and $(2, j_2)$ and zero elsewhere.

- Case 2: search for column k_1 (resp. k_2) containing the entry of largest modulus, $\mathcal{G}_{2,k_1}$ (resp. $\mathcal{G}_{1,k_2}$), among the entries of row 1 (resp. 2) of even column index (resp. odd). Repeat the same normalizing operations and distance calculations as in case 1.

- Choose the case leading to the minimal distance, $\epsilon(\mathcal{G}) = ||\bar{\mathcal{G}} - \mathcal{D}||$, which actually corresponds to:

$$\epsilon(\mathcal{G}) \stackrel{\text{def}}{=} \operatorname*{Min}_{\boldsymbol{P},\nu_1,\alpha_1,\nu_2,\alpha_2} ||\mathcal{G}\,\boldsymbol{P}\,\mathbf{Diag}(\nu_1 z^{\alpha_1},\ \nu_2 z^{\alpha_2})|| \tag{8}$$

This procedure can be easily extended to $N \geq 2$; then, $N!$ cases must be tested instead of 2.

5.2.2. Symbol Error Rate

Another criterion is focussed on source estimates instead of the channel estimate itself. As before, we have to get rid of the possible delay α, the possible factor $\nu = e^{\jmath\psi}$, and the possible permutation σ, that may be present in the estimate, and minimize $\nu\, a_{\sigma(i)}(n-\alpha) - \hat{a}_i(n)$.

This can be done in a similar manner as in the previous section, even if it turns out to be more computationally complex. In fact, under each of the two hypotheses, one must explore 10 different cases (5 possible delays for each row). In general, one must calculate the error rate of $N!\,N(2L-1)$ potential estimators (instead of 20 for $N =$

2), which can become quite costly. In such cases, one may want to assume the matrix $\mathcal{D}$ obtained via the distance criterion.

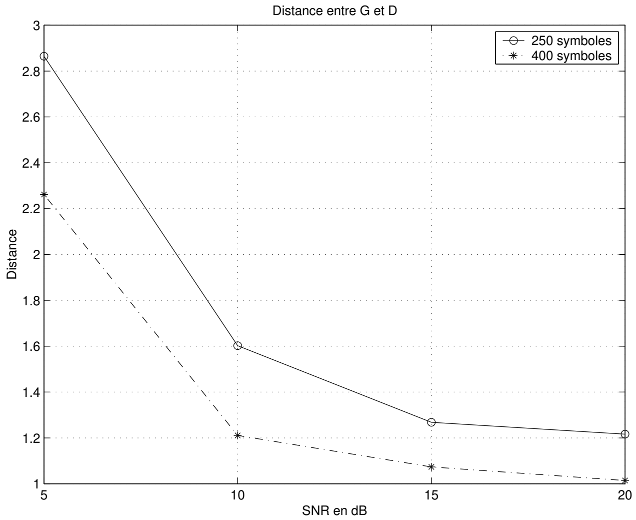

Fig. 4. Error $\epsilon(\mathcal{G})$ on the impulse response.

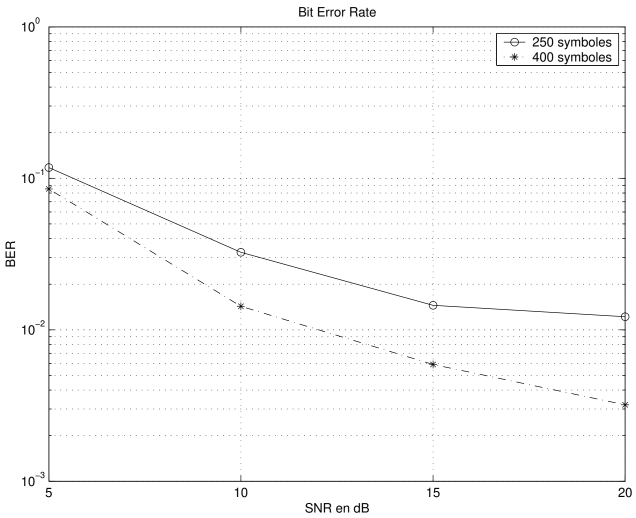

Fig. 5. Mean source Bit Error Rate (BER).

5.3. Simulation results

The two sources generated are i.i.d. binary sequences, and take their values in $\{-1, 1\}$. The equalizer has been estimated from data lengths of 250 to 400 symbols (thus quite short), and perturbed by an additive isotropic white Gaussian noise:

$$\boldsymbol{x}(n) = \sum_{k=0}^{L} \boldsymbol{C}(k)\boldsymbol{a}(n-k) + A\boldsymbol{v}(n)$$

Performances have been averaged over 500 independent realizations. The distance and error rate criteria are plotted in figures 4 and 5 as a function of the Signal to Noise Ratio (SNR), $-20\log_{10} A$.

In the presence of a low noise, matrices $\mathcal{D}$ that have been obtained are the same for both criteria. On the other hand, for larger SNR's, it might happen that the two criteria do not estimate the same matrix $\mathcal{D}$, as reported in figure 6.

6. CONCLUDING REMARKS

Based on the theoretical results obtained in [1], a first numerical algorithm has been developed. Even if it is not yet optimized, this algorithm demonstrates that it is possible to equalize blindly FIR MIMO channels from short data records (typically 300 symbols).

Subjects that need to be further addressed include (i) the robustness to channel length misadjustment, (ii) improvements on convergence and accuracy of the algorithm, (iii) more complete computer experiments, for instance for a wider variety of channels and sources, and (iv) the development of the code in the case of complex channel and sources (the theoretical results indeed hold true in that case).

7. REFERENCES

[1] P. COMON and E. MOREAU, "Blind MIMO equalization and joint-diagonalization criteria," in *ICASSP'01*, Salt Lake City, May 7-11, 2001.

[2] O. SHALVI and E. WEINSTEIN, "New criteria for blind deconvolution of nonminimum phase systems," *IEEE Trans. Inf. Theory*, vol. 36, no. 2, pp. 312–321, Mar. 1990.

[3] D. DONOHO, "On minimum entropy deconvolution," in *Applied time-series analysis II*, pp. 565–609. Academic Press, 1981.

[4] D. YELLIN and E. WEINSTEIN, "Criteria for multichannel signal separation," *IEEE Trans. Sig. Proc.*, vol. 42, no. 8, pp. 2158–2168, Aug. 1994.

[5] A. J. van der VEEN and A. PAULRAJ, "An analytical constant modulus algorithm," *IEEE Trans. Sig. Proc.*, vol. 44, no. 5, pp. 1136–1155, May 1996.

[6] I. FIJALKOW, A. TOUZNI, and J. R. TREICHLER, "Fractionally spaced equalization using CMA: Robustness to channel noise and lack of disparity," *IEEE Trans. Sig. Proc.*, vol. 45, no. 1, pp. 56–66, Jan. 1997.

[7] P. COMON and O. GRELLIER, "Analytical blind identification of a SISO communication channel," in *IEEE Workshop on Statistical Signal and Array Processing*, Pocono Manor, Pennsylvania, Usa, Aug. 14-16, 2000, pp. 206–210.

[8] J. K. TUGNAIT, "Identification and deconvolution of multichannel non-gaussian processes using higher order statistics and inverse filter criteria," *IEEE Trans. Sig. Proc.*, vol. 45, pp. 658–672, Mar. 1997.

[9] C. B. PAPADIAS, "A multi-user kurtosis algorithm for blind source separation," in *ICASSP*, Istanbul, June 5-9 2000.

[10] C. SIMON, P. LOUBATON, C. VIGNAT, C. JUTTEN, and G. d'URSO, "Blind source separation of convolutive mixtures by maximization of fourth-order cumulants : the non i.i.d. case," in *ICASSP*, Seattle, May 12-15 1998.

[11] K. ABED-MERAIM, E. MOULINES, and P. LOUBATON, "Prediction error method for second-order blind identification," *IEEE Trans. Sig. Proc.*, vol. 45, no. 3, pp. 694–705, Mar. 1997.

[12] N. THIRION and E. MOREAU, "New criteria for blind signal separation," in *IEEE Workshop on Statistical Signal and Array Processing*, Pocono Manor, Pennsylvania, Usa, Aug. 14-16, 2000, pp. 344–348.

[13] P. COMON, "Contrasts for multichannel blind deconvolution," *Signal Processing Letters*, vol. 3, no. 7, pp. 209–211, July 1996.

[14] A. SWAMI, G. GIANNAKIS, and S. SHAMSUNDER, "Multichannel ARMA processes," *IEEE Trans. Sig. Proc.*, vol. 42, no. 4, pp. 898–913, Apr. 1994.

[15] P. LOUBATON, E. MOULINES, and P. A. REGALIA, "Subspace methods for blind identification and deconvolution," in *Signal Processing Advances in Wireless and Mobile Communications*, Giannakis, Hua, Stoica, and Tong, Eds., chapter 3. Prentice-Hall, 2001.

[16] P. COMON, "Independent Component Analysis," in *Proc. Int. Sig. Proc. Workshop on Higher-Order Statistics*, Chamrousse, France, July 10-12 1991, pp. 111–120, Republished in *Higher-Order Statistics*, J.L.Lacoume ed., Elsevier, 1992, pp 29–38.

[17] P. COMON, "Tensor decompositions, state of the art and applications," in *IMA Conf. Mathematics in Signal Processing*, Warwick, UK, Dec. 18-20, 2000, keynote address.

[18] J. F. CARDOSO and A. SOULOUMIAC, "Blind beamforming for non-Gaussian signals," *IEE Proceedings - Part F*, vol. 140, no. 6, pp. 362–370, Dec. 1993, Special issue on Applications of High-Order Statistics.

[19] N. DELFOSSE and P. LOUBATON, "Adaptive blind separation of independent sources: a deflation approach," *Signal Processing*, vol. 45, pp. 59–83, 1995.

[20] E. MOREAU, "A generalization of joint-diagonalization criteria for source separation," *IEEE Trans. Signal Processing*, vol. 49, no. 3, pp. 530–541, March 2001.

[21] G. H. GOLUB and C. F. VAN LOAN, *Matrix computations*, The John Hopkins University Press, 1989.

[22] J. F. CARDOSO and A. SOULOUMIAC, "Jacobi angles for simultaneous diagonalization," *SIAM Jour. matrix Analysis*, vol. 17, no. 1, pp. 161–164, Jan. 1996.

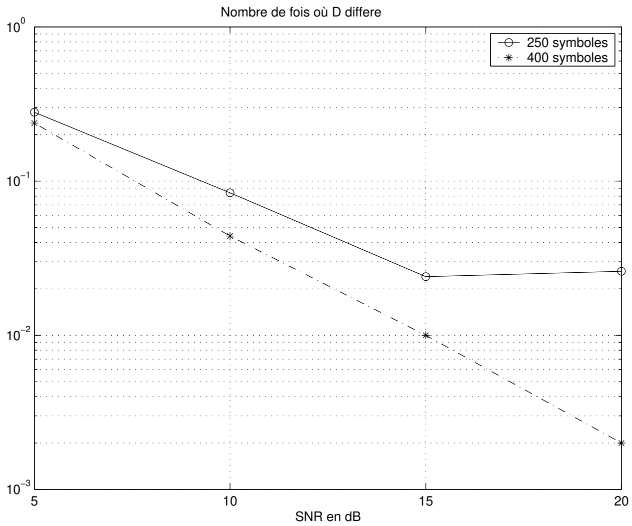

Fig. 6. Error rate on the estimation of matrix $\boldsymbol{D}(z) = \boldsymbol{\Lambda}(z)\boldsymbol{P}$.