ICASSP'03, Hong-Kong, 6-10 April 2003

BLIND MIMO PARAUNITARY EQUALIZER

Ludwig ROTA, Pierre COMON, and Sylvie ICART

I3S, Les Algorithmes/Euclide B - 2000 route des Lucioles,
BP 121, 06903 Sophia-Antipolis Cedex, France.

ABSTRACT

This paper introduces a new Blind Source Separation algorithm for convolutive mixtures. In addition to separate sources, this algorithm respects the paraunitary property of the model considered, obtained after whitening observations. In order to do this, the equalizer is factorized in a novel manner. After a presentation of theoretical results, a numerical algorithm is then derived. This algorithm is based on the solution of a polynomial system, which some values of output cumulant multi-correlations enter. Simulations and performances of the numerical algorithm are presented in the last section.

1. INTRODUCTION

The method presented in this paper is intended to Multiple Input Multiple Output (MIMO) paraunitary channels. The fact that the channel is considered as paraunitary is not restrictive since prewhitening can always be performed in a first stage (in a non unique manner).

Most blind MIMO equalization techniques use High Order Statistics (HOS) for separating signals [1] [2] [3] [4] ; this can be implicit through constant modulus [5] [6] or constant power [7] criteria. Indeed, this paper presents an algorithm based on HOS. Moreover, our algorithm is very attractive since it can be implemented "off-line". Contrary to "on-line" algorithms which need long data block to converge (typically from 10,000 to 100,000 symbols), "off-line" algorithms exhibit much shorter convergence times.

Algorithms like PAJOD [8] have already been proposed for MIMO channels. Unfortunately, the paraunitary constraint was not accurately verified for equalizers when it was considered for channels, especially for low SNR.

Our main contribution consists of a block algorithm dedicated to blind MIMO equalization. The goal of this algorithm is to build a paraunitary equalizer in order to correct channel mixing effects. It has been shown to maximize a well-defined *contrast*, as pointed out in section 3. Simulations and performances obtained are reported in the last section of the paper.

2. MODEL AND NOTATIONS

Throughout the paper, $(^{\mathsf{T}})$ stands for transposition, $(^{\mathsf{H}})$ for conjugate transposition, and $(^*)$ for complex conjugation, and $\jmath = \sqrt{-1}$. Vectors and matrices are denoted with bold lowercase and bold uppercase letters respectively. Next, let $\{\boldsymbol{G}(k), k \in \mathbb{Z}\}$ denote the matrix impulse response of the global system. Then, we denote its transfer function as $\boldsymbol{G}[z] \stackrel{\text{def}}{=} \sum_k \boldsymbol{G}(k) z^{-k}$. Furthermore, the entries of the matrix $\boldsymbol{G}$ are denoted G_{ij}, where subscript ij denotes the i-th row and the j-th column of $\boldsymbol{G}$.

Now, consider the linear time-invariant (LTI) invertible system of length L, mixing N white random processes. This system (depicted in figure 1) is described by:

$$\boldsymbol{w}(n) = \sum_{k=0}^{L} \boldsymbol{C}(n-k)\boldsymbol{s}(k) \tag{1}$$

where $\{\boldsymbol{C}(n), n \in \mathbb{Z}\}$ is a sequence of $N \times N$ matrices denoting the complex Finite Impulse Response (FIR) of channel $\boldsymbol{C}[z]$, $\boldsymbol{s} = (s_1, \ldots, s_N)^{\mathsf{T}}$ denotes the $N-$

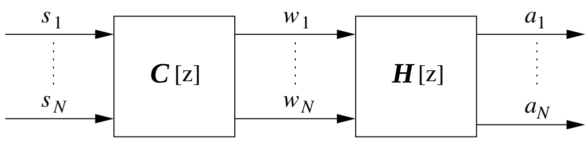

Figure 1: Source $\boldsymbol{s}$ is filtered by channel $\boldsymbol{C}[z]$ and observation $\boldsymbol{w}$ is equalized by $\boldsymbol{H}[z]$.

dimensional source vector of baseband complex signals, $\boldsymbol{w} = (w_1, \ldots, w_N)^{\mathsf{T}}$ the $N-$ dimensional observation vector, and $\boldsymbol{a} = (a_1, \ldots, a_N)^{\mathsf{T}}$ the $N-$ dimensional estimated source vector. All these vectors are *spatially* and *temporally* white at second order. Note that $L = 0$ corresponds to an instantaneous mixture.

The multichannel blind deconvolution problem consists of finding a LTI filter $\boldsymbol{H}[z]$, the *equalizer*, in order to retrieve the N input signals $s_i(n), i \in \{1, \ldots, N\}$, solely from the observations $\boldsymbol{w}(n)$ of the output of the unknown LTI channel $\boldsymbol{C}[z]$. The signals recovered may be reordered by a permutation matrix $\boldsymbol{P}$, and delayed by a diagonal filter $\boldsymbol{\Lambda}[z]$, so that $\boldsymbol{C}[z]\boldsymbol{H}[z] = \boldsymbol{\Lambda}[z]\boldsymbol{P}$. The estimated source vector is $\boldsymbol{a}(n) = \sum_k \boldsymbol{H}(n-k)\boldsymbol{w}(k)$ and the global LTI system $\boldsymbol{G}[z]$ is defined according to $\boldsymbol{a}(n) = \sum_k \boldsymbol{G}(n-k)\boldsymbol{s}(k)$.

Definition 1: ***Paraunitarity***. *A $N \times N$ polynomial matrix $\boldsymbol{H}[z]$ is said to be paraunitary [9] if :*

$$\boldsymbol{H}^{\mathsf{H}}[1/z^*]\boldsymbol{H}[z] = \boldsymbol{I}_N \tag{2}$$

where $\boldsymbol{I}_N$ is the $N \times N$ identity matrix.

The following hypotheses are assumed:

H1. Inputs $s_i(n), i \in \{1, \dots, N\}$, are mutually independent and identically distributed (i.i.d.) zero-mean random processes, with unit variance.

H2. The vector $\boldsymbol{s}(n)$ is stationary up to the considered order r, $r \geq 3$, *i.e.* $\forall i \in \{1, \dots, N\}$, the order-$r$ marginal cumulants,

$$\mathsf{C}_p^q[s_i] = \mathsf{Cum}[\underbrace{s_i(n), \dots, s_i(n)}_{p \text{ terms}}, \underbrace{s_i^*(n), \dots, s_i^*(n)}_{q=r-p \text{ terms}}]$$

do not depend on n. For definitions of cumulants, refer to [10] and references therein.

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

H4. $\boldsymbol{C}[z]$, $\boldsymbol{H}[z]$, and hence $\boldsymbol{G}[z] = \boldsymbol{H}[z]\boldsymbol{C}[z]$ are all paraunitary, as defined in definition 1.

Remark 1. The constraint of hypothesis **H4** is not restrictive. Indeed, one can always whiten the observations by using a filter that factorizes the second-order power spectrum, *i.e.* a classical prewhitening of the observations [11]. Thus, paraunitary filters can be easily obtained by standardization of observations (second order white with unit covariance).

Considering the previous hypotheses and models, and assuming that $N = 2$, we can make a first proposition:

Proposition 1: *A $N \times N$ FIR paraunitary filter of length $L \geq 0$, $\boldsymbol{H}[z]$, can be factorized in 3 filters:*

$$\boldsymbol{H}[z] = \boldsymbol{A}[z]\boldsymbol{W}_{\ell_b}\boldsymbol{B}[z] \tag{3}$$

where $\boldsymbol{A}[z]$ and $\boldsymbol{B}[z]$ are FIR paraunitary filters of length ℓ_a and ℓ_b respectively, with:

$$0 \leq \ell_a \leq L \text{ and } 0 \leq \ell_b \leq L$$
$$\ell_a + \ell_b = L.$$

and $\boldsymbol{W}_{\ell_b}$ is a $N \times N$ unitary matrix.

Proof. For convenience, we prove the proposition for $N = 2$. Extending the factorization of [9] to the non real case, one gets the following factorization:

$$\boldsymbol{H}[z] = \boldsymbol{W}_L\boldsymbol{Z}[z]\boldsymbol{W}_{L-1}\dots\boldsymbol{Z}[z]\boldsymbol{W}_0 \tag{4}$$

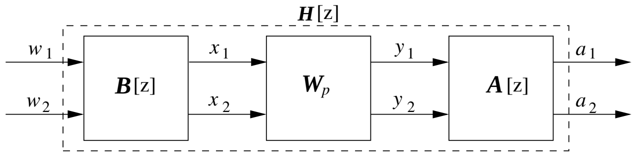

Figure 2: Factorization of the paraunitary equalizer in 3 filters.

where $\boldsymbol{W}_p$, $p \in \{0, \dots, L\}$, are 2×2 unitary, and $\boldsymbol{Z}[z]$ is 2×2 diagonal:

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

When $\ell_a = 0$ (respectively $\ell_b = 0$), we can replace $\boldsymbol{A}[z]$ (respectively $\boldsymbol{B}[z]$) by $\boldsymbol{I}_2$. When $\ell_a > 0$, filter $\boldsymbol{A}[z]$ is defined as the product:

$$\boldsymbol{A}[z] = \boldsymbol{W}_L\boldsymbol{Z}[z]\dots\boldsymbol{W}_{\ell_b+1}\boldsymbol{Z}[z] \tag{5}$$

and when $\ell_b > 0$, we have:

$$\boldsymbol{B}[z] = \boldsymbol{Z}[z]\boldsymbol{W}_{\ell_b-1}\dots\boldsymbol{Z}[z]\boldsymbol{W}_0. \tag{6}$$

Thus, between $\boldsymbol{A}[z]$ and $\boldsymbol{B}[z]$, it remains $\boldsymbol{W}_{\ell_b}$. Hence, we can factorize $\boldsymbol{H}[z]$ like in (3). $\diamondsuit$

For the sake of clarity, it will be now assumed that $N = 2$. In the remaining, we assume the following notation for cumulants, *e.g.* cumulants of vector $\boldsymbol{w}$:

$$\Gamma^{\boldsymbol{w}}_{eg,fh}(\boldsymbol{\nu}) = \mathsf{Cum}[w_e(n-\nu_1), w_f^*(n-\nu_2), w_g(n-\nu_3), w_h^*(n-\nu_4)]. \tag{7}$$

where $\{e, f, g, h\}$ take their values in $\{1, 2\}$, and $\nu_i \in \mathbb{N}, \forall i \in \{1, \dots, 4\}$.

Now, consider the following input-output relations for the convolutive model:

$$a_i(n) = \sum_{q,r,m} A_{iq}(m)W_{qr}x_r(n-m), \tag{8}$$

$$\text{and} \quad x_r(n-m) = \sum_{s,l} B_{rs}(l)w_s(n-m-l). \tag{9}$$

From (8), thanks to the multilinearity property of cumulants, we can express the input-output relations between cumulants of input $\boldsymbol{x}$ and output $\boldsymbol{a}$:

$$\Gamma^{\boldsymbol{a}}_{ik,jl} = \sum_{abcd}\sum_{\boldsymbol{\tau}}\sum_{qrst} A_{iq}(\tau_1)A^*_{jr}(\tau_2)A_{ks}(\tau_3) A^*_{lt}(\tau_4)W_{qa}W^*_{rb}W_{sc}W^*_{td}\Gamma^{\boldsymbol{x}}_{ac,bd}(\boldsymbol{\tau}) \tag{10}$$

with $\boldsymbol{\tau} = (\tau_1, \tau_2, \tau_3, \tau_4)$. The range of each τ_i is $[0, \dots, \ell_a]$ and indices $\{a, b, c, d, i, j, k, l, q, r, s, t\}$ take their values in

$\{1,2\}$. It is the same for input-output relations between observed cumulants of $\boldsymbol{w}$ and computed cumulants at the output of $\boldsymbol{B}[z]$ (*i.e.* cumulants of $\boldsymbol{x}$):

$$\Gamma^{\boldsymbol{x}}_{ac,bd}(\boldsymbol{\tau}) = \sum_{\boldsymbol{\rho}} \sum_{efgh} B_{ae}(\rho_1) B^*_{bf}(\rho_2) B_{cg}(\rho_3) B^*_{dh}(\rho_4) \Gamma^{\boldsymbol{w}}_{eg,fh}(\boldsymbol{\tau}+\boldsymbol{\rho}) \qquad (11)$$

with $\boldsymbol{\rho} = (\rho_1, \rho_2, \rho_3, \rho_4)$. The range of each ρ_i is $[0, \ldots, \ell_b]$. The global input-output relation of equalizer $\boldsymbol{H}[z]$ is not given in this paper since it is not necessary for the algorithm. Nevertheless it can be easily deduced by combining (10) and (11) as shown in [12].

Remark 2. For $N=2$, unitary matrices $\boldsymbol{W}_p$ can be generated $\forall p$ as:

$$\boldsymbol{W}_p = \begin{pmatrix} \cos\theta_p & \sin\theta_p e^{\jmath\phi_p} \\ -\sin\theta_p e^{-\jmath\phi_p} & \cos\theta_p \end{pmatrix}. \qquad (12)$$

Thus, only one pair of angles (θ_p, ϕ_p) is needed for each $\boldsymbol{W}_p$.

3. CONTRAST PROPOSED

The reader is invited to consult [13] for definitions and properties about *contrasts*.

Proposition 2: *The separation of sources, solely from outputs of the channel, can be performed by maximizing the following contrast:*

$$\Upsilon_{1,4} = \sum_{i=1}^{N} \left| \Gamma^{\boldsymbol{a}}_{ii,ii} \right|. \qquad (13)$$

Here, $\Gamma^{\boldsymbol{a}}_{ii,ii}$ includes entries of $\boldsymbol{W}_p$ matrices. So, in order to estimate $\boldsymbol{H}[z]$, the criterion can be written:

$$\boldsymbol{H} = Arg \max_{\mathbb{W}} \Upsilon_{1,4} \qquad (14)$$

where $\mathbb{W}$ stands for the set of $\boldsymbol{W}_p$, $p \in \{0, \ldots, L\}$.

It has been proved in [13] that $\Upsilon_{1,4}$ is a contrast. We maximize $\Upsilon_{1,4}$ with respect to each pair (θ_p, ϕ_p) in turn. The sequence of values of $\Upsilon_{1,4}$ obtained this way is monotonically increasing. Since it is also bounded above, it converges.

For sake of clarity, we drop index p. Thus we have to find all pairs (θ, ϕ) which maximize (13) independently from other pairs. To reach this goal, we have to simplify (10) firstly by expanding it, and secondly by collecting terms involving θ or ϕ. In this manner, we obtain the following equation for the output cumulants of $\boldsymbol{a}$:

$$\Gamma^{\boldsymbol{a}}_{ii,ii} = \sum_{\alpha=0}^{4} \left(\sum_{\beta=0}^{4-\alpha} \mathcal{K}_{\alpha}^{2\beta+\alpha-4} (\cos\theta)^{\alpha} (\sin\theta)^{4-\alpha} e^{\jmath(2\beta+\alpha-4)\phi} \right) \qquad (15)$$

where each $\mathcal{K}_{\alpha}^{2\beta+\alpha-4}$ denotes the product of 4 entries of $\boldsymbol{A}[z]$, depending on indices a, b, c, d and q, r, s, t, and in accordance with α and β.

Next, make the change of variables: $\cos\phi = \frac{1-t^2}{1+t^2}$, $\sin\phi = \frac{2t}{1+t^2}$ with $t = \tan\frac{\phi}{2}$, and $\cos\theta = \frac{1}{\sqrt{1+u^2}}$, $\sin\theta = \frac{u}{\sqrt{1+u^2}}$ with $u = \tan\theta$.

Now, in order to maximize contrast (13) , we find all the roots of polynomial system (16), *i.e.* stationary points of $\Upsilon_{1,4}$, thanks to derivatives:

$$\left. \begin{aligned} \Phi_1(u,t) &= \frac{\partial \Upsilon_{1,4}}{\partial u} = \sum_{k=0}^{4} \lambda_{4-k}(t)\, u^k \\ \Phi_2(u,t) &= \frac{\partial \Upsilon_{1,4}}{\partial t} = \sum_{k=0}^{3} \xi_{3-k}(t)\, u^k \end{aligned} \right\} \qquad (16)$$

Polynomial system (16) can be solved by using the resultant of a Sylvester matrix. Thus, considering only variable u for $\Phi_1(u,t)$ and $\Phi_2(u,t)$, and collecting terms of same degree in u, we obtain a Sylvester matrix of size 7×7. See [12] for more details about resolution of (16). When all roots are found, we plug them back in (15) in order to select the best solution for (14).

4. ALGORITHM

In this section we present the algorithm derived from previous statements. It has been implemented for $N=2$ and results are shown in section 5.

The algorithm is the following:

1. Compute the tensor of cumulants $\Gamma^{\boldsymbol{w}}_{eg,fh}(\boldsymbol{\tau}+\boldsymbol{\rho})$ defined in (7) and of length $L = \max\{\boldsymbol{\tau}\} + \max\{\boldsymbol{\rho}\}$.

2. Initialize equalizer $\boldsymbol{H}[z]$ with $(\theta_p = 0, \phi_p = 0)$, $\forall p \in \{0, \ldots, L\}$

3. Loop on $k = 0, \ldots, L$.

   (a) Compute the cumulant tensor $\Gamma^{\boldsymbol{x}}_{ac,bd}(\boldsymbol{\tau})$,

   (b) Search for pair (θ_k,ϕ_k) maximizing $\Upsilon_{1,4}$,

   (c) Plug back angles (θ_k, ϕ_k) in $\boldsymbol{W}_k$ (update filters $\boldsymbol{A}[z]$ and $\boldsymbol{B}[z]$ as in (5) and (6)).

4. Goto *3* until number of sweeps $\leq T$.

Of course, this algorithm considers that angles are all independent. The resulting tensor of (*3*a) is composed of $N^2(L-\ell_b+1)^2$ matrices each of size $N(L-\ell_b+1) \times N(L-\ell_b+1)$. In order to increase the precision of the angles, we suggest to execute $T = \lceil \sqrt{L}\, \rceil + 1$ sweeps. Actually, the first angles computed are not well defined since all other angles are null (set at stage *2*). Hence, when loop *3* is repeated T times, angles are better estimated.

5. COMPUTER RESULTS

One considers a FIR complex mixture of length $L = 3$ of $N = 2$ unit variance QPSK white processes. The channels are paraunitary in order to preserve second-order whiteness and are constructed as explained in section 2. The 8 angles θ_p, ϕ_p, for $0 \leq i \leq L$, are drawn according to a uniform distribution in $[0, 2\pi)$ in order to generate paraunitary channels. For each randomly generated channel, blocks of noisy observations are generated according to $\boldsymbol{w}(n) = \sum_{k=0}^{L} \boldsymbol{C}(n-k)\boldsymbol{s}(k) + \rho\boldsymbol{v}(n)$ where $\boldsymbol{v}(n)$ is a white circular complex Gaussian noise with identity covariance matrix, and $s_i(n)$ the source sequences. Parameter ρ is introduced in order to control the Signal to Noise Ratio per bit (SNR), and is defined as: $SNR_{dB} = -20\log_{10}\rho$.

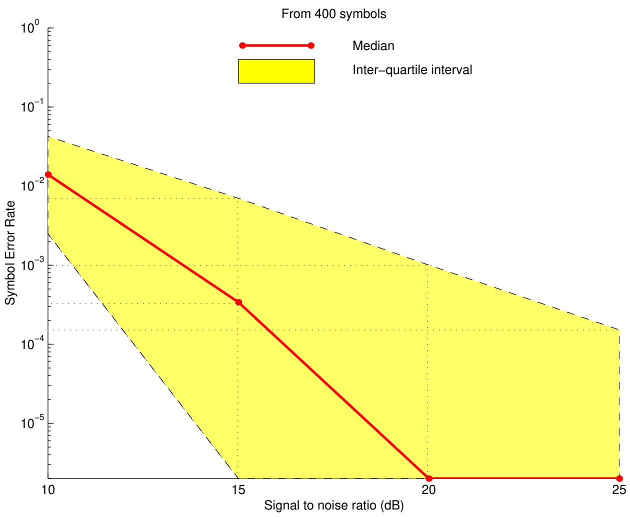

Figure 3: Symbol Error Rate obtained when a length-3 equalizer is built from blocks of 400 symbols. Zero values are replaced by the minimal resolution (2.10^{-6}).

Equalizers returned by the algorithm are then tested with two different white processes of 10000 symbols in order to compute the Symbol Error Rate (SER). Figure 3 shows median SER for blocks of 400 symbols, over 25 trials. Thus, the minimal resolution is $(2*10000*25)^{-1} = 2.10^{-6}$. The inter-quartile interval is also represented, *i.e.* 25% of values on both sides of the median, since it is a good representation of global performances for an algorithm. Simulations prove that our algorithm works well on short data length since from 400 symbols and with a noise of 20dB, the median SER is below the minimal resolution.

6. CONCLUDING REMARKS

Through this paper, we have presented a parametrization of paraunitary equalizers, in order to respect the paraunitarity of the channel after prewhitening. Then, from theoretical results of section 2, a numerical algorithm has been implemented and performances evaluated. Results obtained are very attractive since the algorithm works very well on data blocks as short as 400 symbols.

7. REFERENCES

[1] D. DONOHO, "On minimum entropy deconvolution," in *Applied time series analysis II* (A. Press, ed.), pp. 565–609, 1981.

[2] D. YELLIN and E. WEINSTEIN, "Criteria for multichannel signal separation," *IEEE Trans. Sig. Proc.*, vol. 42, pp. 2158–2168, Aug. 1994.

[3] O. SHALVI and E. WEINSTEIN, "New criteria for blind deconvolution of nonminimum phase systems," *IEEE Trans. Inf. Theory*, vol. 36, pp. 312–321, Mar. 1990.

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

[5] 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, pp. 56–66, Jan. 1997.

[6] A. J. van der VEEN and A. PAULRAJ, "An analytical constant modulud algorithm," *IEEE Trans. Sig. Proc.*, vol. 44, pp. 1136–1155, May 1996.

[7] O. GRELLIER, P. COMON, B. MOURRAIN, and P. TREBUCHET, "Analytical blind channel identification," *IEEE Trans. Signal Processing*, vol. 50, Sept. 2002.

[8] P. COMON, E. MOREAU, and L. ROTA, "Blind separation of convolutive mixtures: a contrast-based joint diagonalization approach," in *3rd Int. Conf. Independent Component Analysis*, (San Diego), pp. 686–691, Dec. 9-13 2001.

[9] P. P. VAIDYANATHAN, *Multirate Systems and Filter Banks*. London: Prentice-Hall, 1993.

[10] P. McCULLAGH, *Tensor Methods in Statistics*. Monographs on Statistics and Applied Probability, Chapman and Hall, 1987.

[11] A. H. SAYED and T. KAILATH, "A Survey of Spectral Factorization Methods," *Numer. Linear Algebra Appl.*, vol. 8, pp. 467–496, 2001.

[12] L. ROTA, P. COMON, and S. ICART, "Blind mimo equalization under paraunitary constraint," I3S Report RR-2002-35-FR, Laboratoire I3S, `http://www.i3s.unice.fr/`, Sept. 2002.

[13] P. COMON, "Contrasts for multichannel blind deconvolution," *IEEE Signal Processing Letters*, vol. 3, pp. 209–211, July 1996.