LABORATOIRE

INFORMATIQUE, SIGNAUX ET SYSTÈMES
DE SOPHIA ANTIPOLIS
UMR 6070

BLIND MIMO EQUALIZATION UNDER PARAUNITARY CONSTRAINT

Ludwig ROTA, Pierre COMON, Sylvie ICART

Projet ASTRE

Rapport de recherche
I3S/RR–2002-35–FR

Septembre 2002

LABORATOIRE I3S: Les Algorithmes / Euclide B – 2000 route des Lucioles – B.P. 121 –
06903 Sophia-Antipolis Cedex, France – Tél. (33) 492 942 701 – Télécopie : (33) 492 942 898
http://www.i3s.unice.fr/I3S/FR/

RÉSUMÉ :

MOTS CLÉS :

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 model considered, obtained after whitening observations. In order to respect this property, authors introduce a new model for equalizer, wisely factorized in 3 filters. After a presentation of theoretical results, a numerical algorithm is then derived. This algorithm is based on the solution of a polynomial system, containing some values of output cumulant multi-correlations. Simulations and performances of the numerical algorithm are presented in the last section.

KEY WORDS :

MIMO, Blind Equalization, Source Separation, Paraunitary matrix, Tensor, Cumulants, Sylvester matrix

Blind MIMO equalization under paraunitary constraint

Ludwig ROTA, Pierre COMON, and Sylvie ICART

Laboratoire I3S, Les Algorithmes/Euclide B - 2000 route des Lucioles,

BP 121, F-06903 Sophia-Antipolis Cedex, France.

I3S Report RR–2002–35–FR

September 10, 2002

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 respect this property, authors introduce a new model for equalizer, wisely factorized in 3 filters. After a presentation of theoretical results, a numerical algorithm is then derived. This algorithm is based on the solution of a polynomial system, containing some values of output cumulant multi-correlations. Simulations and performances of the numerical algorithm are presented in the last section.

Keywords: MIMO, Blind Equalization, Source Separation, Paraunitary matrix, Tensor, Cumulants, Sylvester matrix

1 Introduction

Blind channel identification has been studied extensively during the last decade. The method presented in this paper is intended to Multiple Input Multiple Output (MIMO) paraunitary channel. The fact that the channel is considered as paraunitary is not restrictive since prewhitening can always be performed in a first stage.

Most blind MIMO equalization techniques use High Order Statistics (HOS) for separating signals [5] [13] [10], even if second order statistics are sufficient; this can be implicit through constant modulus [6] [12] 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 only need approximatively 1000 to 2000 symbols.

Algorithms like PAJOD [4] have already been proposed for MIMO channels. Unfortunately, the paraunitary constraint was not accurately verified for equalizers when it was considered for channels.

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. Moreover some implementation tips are proposed in order to reduce time calculus for long-length channels. 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, with $\jmath^2 = -1$. Vectors and matrices are denoted with bold lowercase and bold uppercase letters respectively. Next, let $\{\boldsymbol{H}(k), k \in \mathbb{Z}\}$ denote a matrix impulse response. Then, we denote its transfer function as:

$$\boldsymbol{H}[z] \stackrel{\text{def}}{=} \sum_k \boldsymbol{H}(k) z^{-k}.$$

Furthermore, the entries of a matrix $\boldsymbol{H}$ are denoted H_{ij}, where subscript ij denotes the i-th row and the j-th column of $\boldsymbol{H}$.

Now, consider the linear time-invariant system (LTI) depicted in figure 1 like a Finite Impulse Response (FIR) complex mixture of length L of N

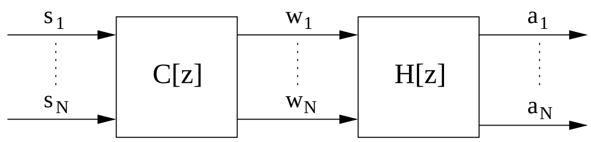

Fig. 1: Source $\boldsymbol{s}$ is filtered by channel $\boldsymbol{C}[z]$ and observation $\boldsymbol{w}$ is equalized by $\boldsymbol{H}[z]$.

white random processes. More precisely, denote $\boldsymbol{s} = (\mathrm{s}_1, \ldots, s_N)^\top$ the $N-$dimensional source vector of complex signals, $\boldsymbol{w} = (\mathrm{w}_1, \ldots, w_N)^\top$ the $N-$dimensional observation vector and $\boldsymbol{a} = (\mathrm{a}_1, \ldots, a_N)^\top$ the $N-$ dimensional estimated source vector. Hence, the multichannel LTI invertible system is described by the following equation:

$$\boldsymbol{w}(n) = \sum_k \boldsymbol{C}(n-k)\boldsymbol{s}(k) \tag{1}$$

where $\{\boldsymbol{C}(n), n \in \mathbb{Z}\}$ is a sequence of $N \times N$ matrices which denote the impulse response of the LTI mixing filter. Hence, the transfer function of the channel is the following:

$$\boldsymbol{C}[z] = \sum_k \boldsymbol{C}(k) z^{-k}. \tag{2}$$

The multichannel blind deconvolution problem consists of finding a LTI filter $\boldsymbol{H}[z]$, also named *equalizer*, in order to retrieve the N input signals $\{s_i(n), i = 1, \ldots, N\}$, solely from the observations $\boldsymbol{w}(n)$ of the output of the unknown LTI channel $\boldsymbol{C}[z]$. Hence, the estimated source vector is the following:

$$\boldsymbol{a}(n) = \sum_k \boldsymbol{H}(n-k)\boldsymbol{w}(k) \tag{3}$$

Define the global LTI system $\boldsymbol{G}[z]$ according to

$$\boldsymbol{a}(n) = \sum_k \boldsymbol{G}(n-k)\boldsymbol{s}(k) \tag{4}$$

with the transfer function

$$\boldsymbol{G}[z] = \sum_k \boldsymbol{G}(k) z^{-k}. \tag{5}$$

Now, we take the following definition of *paraunitary*:

Definition 1: ***Paraunitary***. *A $N \times N$ rational matrix $\boldsymbol{H}[z]$ is said to be paraunitary [11] if :*

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

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

The following hypotheses are assumed:

H1. Inputs $\{s_i(n), i = 1, \ldots, N\}$ 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.* $\forall i \in \{1, \ldots, N\}$, the order-$r$ marginal cumulants,

$$\mathsf{C}_p^q[s_i] = \mathsf{Cum}[\underbrace{s_i(n), \ldots, s_i(n)}_{p \text{ terms}}, \underbrace{s_i^*(n), \ldots, s_i^*(n)}_{q=r-p \text{ terms}}] \tag{6}$$

do not depend on n. For definitions of cumulants, refer to [9] and references therein.

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

$$\boldsymbol{G}[z]\boldsymbol{G}^{\mathsf{H}}[1/z^*] = \boldsymbol{I}_N \tag{7}$$

and hence

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

in other words, $\boldsymbol{G}[z]$ and $\boldsymbol{H}[z]$ are *paraunitary*.

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. Thus paraunitary filters can be easily obtained by standardization of observations (second order white with unit covariance).

Considering the previous hypotheses and models, we can make a first proposition:

Proposition 1: *A $N \times N$ FIR paraunitary filter of length L, $\boldsymbol{H}[z]$, can be factorized in 3 filters as shown in figure 2:*

$$\boldsymbol{H}[z] = \boldsymbol{A}[z].\boldsymbol{Q}.\boldsymbol{B}[z]$$

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

$$1 \leq L_a \leq L \text{ and } 1 \leq L_b \leq L$$
$$L_a + L_b = L + 1 \text{ with } L \in \mathbb{Z}_+^*.$$

and $\boldsymbol{Q}$ is a $N \times N$ unitary matrix.

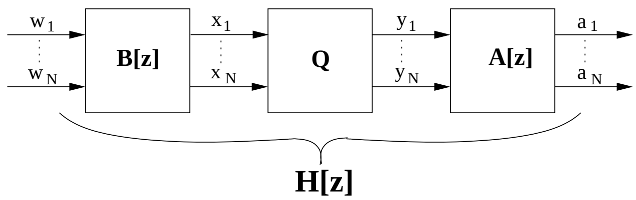

Fig. 2: Factorization of the paraunitary equalizer in 3 filters.

Proof. According to the general factorization of paraunitary matrices [11], the length L equalizer can be generated as follows:

$$\boldsymbol{H}[z] = \boldsymbol{W}_1 \boldsymbol{Z}_1[z] \boldsymbol{W}_2 \ldots \boldsymbol{Z}_{L-1}[z] \boldsymbol{W}_L$$

where, for the same number N of inputs and outputs, $\boldsymbol{W}_p$, $p \in \{1, \ldots, L\}$ denotes a $N \times N$ unitary matrix and $\boldsymbol{Z}_k[z]$, $k \in \{1, \ldots, L-1\}$ denotes a $N \times N$ diagonal matrix:

$$\boldsymbol{Z}_k[z] = \boldsymbol{P}_{i,N} \begin{pmatrix} \boldsymbol{I}_{N-1} & \boldsymbol{0} \\ \boldsymbol{0} & z^{-1} \end{pmatrix} \boldsymbol{P}_{i,N}$$

where $\boldsymbol{P}_{i,N}$ denotes a permutation matrix ($i.e$ swapping components Z_{ii} and Z_{NN}).Matrices $\boldsymbol{W}_k$ are generated thanks to Givens matrices, themselves function of 2 angles (θ and ϕ):

$$\boldsymbol{R}(\theta_i, \phi_i) = \begin{pmatrix} \boldsymbol{I} & \vdots & \boldsymbol{0} & \vdots & \boldsymbol{0} \\ \cdots & \cos\theta_i & \cdots & \sin\theta_i e^{\jmath\phi_i} & \cdots \\ \boldsymbol{0} & \vdots & \boldsymbol{I} & \vdots & \boldsymbol{0} \\ \cdots & -\sin\theta_i e^{-\jmath\phi_i} & \cdots & \cos\theta_i & \cdots \\ \boldsymbol{0} & \vdots & \boldsymbol{0} & \vdots & \boldsymbol{I} \end{pmatrix}$$

Hence, for $N = 2$, we take the following equality: $\boldsymbol{W}_p = \boldsymbol{R}(\theta_p, \phi_p)$. Case $L_a = 1$ (respectively $L_b = 1$) is equivalent to replace $\boldsymbol{A}[z]$ (respectively $\boldsymbol{B}[z]$) by $\boldsymbol{I}_N$. For $L_a > 1$, filter $\boldsymbol{A}[z]$ is defined as the product:

$$\boldsymbol{A}[z] = \boldsymbol{W}_1 \boldsymbol{Z}_1[z] \dots \boldsymbol{W}_{L_a-1} \boldsymbol{Z}_{L_a-1}[z]$$

and for filter $\boldsymbol{B}[z]$ when $L_b > 1$:

$$\boldsymbol{B}[z] = \boldsymbol{Z}_{L_a+1}[z] \boldsymbol{W}_{L_a+2} \dots \boldsymbol{Z}_{L-1}[z] \boldsymbol{W}_L.$$

Hence, the unknown unitary matrix $\boldsymbol{Q}$ of the equalizer corresponds to $\boldsymbol{W}_{L_a}$ in the general definition of a paraunitary filter. ♢

Example. In the case of a length 3 equalizer and for $N = 2$ inputs and outputs, we have 3 couples of angles to search. Tabular 3 describes filters $\boldsymbol{A}[z]$ and $\boldsymbol{B}[z]$ for each rotation matrix $\boldsymbol{Q}$.

Filter $\boldsymbol{A}[z]$	$\boldsymbol{Q}$ searched	Filter $\boldsymbol{B}[z]$
$\boldsymbol{I}_N$	$\boldsymbol{W}_1$	$\boldsymbol{Z}_1[z]\boldsymbol{W}_2\boldsymbol{Z}_2[z]\boldsymbol{W}_3$
$\boldsymbol{W}_1\boldsymbol{Z}_1[z]$	$\boldsymbol{W}_2$	$\boldsymbol{Z}_2[z]\boldsymbol{R}_3$
$\boldsymbol{W}_1\boldsymbol{Z}_1[z]\boldsymbol{W}_2\boldsymbol{Z}_2[z]$	$\boldsymbol{W}_3$	$\boldsymbol{I}_N$

Fig. 3: 3 cases for a length 3 equalizer.

Now, consider the followings input-output relations of our filters for $N = 2$:

$$\boldsymbol{x} = \boldsymbol{B}[z]\boldsymbol{w}, \quad \boldsymbol{y} = \boldsymbol{Q}\boldsymbol{x}, \quad \boldsymbol{a} = \boldsymbol{A}[z]\boldsymbol{y}.$$

Those relations can be expressed as a function of two elements: (i) time instant n with $n \in \mathbb{Z}$ and (ii) observed signal $\{\mathrm{w}_i(n), i \in [1, \dots, N]\}$. The previous relations became:

$$x_i(n) = \sum_{q,m} B_{iq}(m) w_q(n-m),$$

$$y_i(n) = \sum_{q} Q_{iq} x_q(n),$$

and

$$a_i(n) = \sum_{q,m} A_{iq}(m) y_q(n-m).$$

In the remaining, we take the following general notation for cumulants, *e.g.* cumulants of vector $\boldsymbol{w}$:

$$\Gamma^{\boldsymbol{w}}_{eg,fh}(\boldsymbol{\tau}+\boldsymbol{\mu}+\boldsymbol{\rho}) = \mathrm{C}um[w_e(n-\tau_1-\mu_1-\rho_1), w_f^*(n-\tau_2-\mu_2-\rho_2),$$
$$w_g(n-\tau_3-\mu_3-\rho_3), w_h^*(n-\tau_4-\mu_4-\rho_4)]. \quad (8)$$

Now, using the multilinearity property of cumulants, we can express the input-output relations between all cumulants of the system. Thus, between estimated source vector $\boldsymbol{a}$ and output vector $\boldsymbol{y}$ of $\boldsymbol{Q}$, we have a convolutive relation (through $\boldsymbol{A}[z]$) which leads to:

$$\Gamma^{\boldsymbol{a}}_{ik,jl}(\boldsymbol{\mu}) = \sum_{\boldsymbol{\tau}}\sum_{qrst} A_{iq}(\tau_1)A^*_{jr}(\tau_2)A_{ks}(\tau_3)A^*_{lt}(\tau_4)\Gamma^{\boldsymbol{y}}_{qs,rt}(\boldsymbol{\tau}+\boldsymbol{\mu}) \quad (9)$$

with $\boldsymbol{\mu} = (\mu_1,\mu_2,\mu_3,\mu_4)$ and $\boldsymbol{\tau} = (\tau_1,\tau_2,\tau_3,\tau_4)$. The range of each τ_i is $[0,\ldots,L_a-1]$. The input-output relation through $\boldsymbol{Q}$ is:

$$\Gamma^{\boldsymbol{y}}_{qs,rt}(\boldsymbol{\tau}+\boldsymbol{\mu}) = \sum_{abcd} Q_{qa}Q^*_{rb}Q_{sc}Q^*_{td}\Gamma^{\boldsymbol{x}}_{ac,bd}(\boldsymbol{\tau}+\boldsymbol{\mu}) \quad (10)$$

By combining (9) and (10), we deduce the following relation between computed cumulants of $\boldsymbol{x}$ and output cumulants of $\boldsymbol{a}$:

$$\Gamma^{\boldsymbol{a}}_{ik,jl}(\boldsymbol{\mu}) = \sum_{abcd}\sum_{\boldsymbol{\tau}}\sum_{qrst} A_{iq}(\tau_1)A^*_{jr}(\tau_2)A_{ks}(\tau_3)A^*_{lt}(\tau_4)$$
$$Q_{qa}Q^*_{rb}Q_{sc}Q^*_{td}\Gamma^{\boldsymbol{x}}_{ac,bd}(\boldsymbol{\tau}+\boldsymbol{\mu}) \quad (11)$$

Next, the input-output relation between observed cumulants of $\boldsymbol{w}$ and computed cumulants at the output of $\boldsymbol{B}[z]$ (*i.e.* cumulants of $\boldsymbol{x}$) is similar as equation (9):

$$\Gamma^{\boldsymbol{x}}_{ac,bd}(\boldsymbol{\tau}+\boldsymbol{\mu}) = \sum_{\boldsymbol{\rho}}\sum_{efgh} B_{ae}(\rho_1)B^*_{bf}(\rho_2)B_{cg}(\rho_3)B^*_{dh}(\rho_4)\Gamma^{\boldsymbol{w}}_{eg,fh}(\boldsymbol{\tau}+\boldsymbol{\mu}+\boldsymbol{\rho}) \quad (12)$$

with $\boldsymbol{\rho} = (\rho_1,\rho_2,\rho_3,\rho_4)$. The range of each ρ_i is $[0,\ldots,L_b-1]$. 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 (11) and (12).

3 Contrast proposed

Now, let us focus on the contrast used to carry out the equalization of the system. In the previous section, $\boldsymbol{Q}$ is an instantaneous filter since it is a

$N \times N$ unitary matrix (typically generated from Givens matrices). Thus, for $N = 2$, we define it as a function of 2 angles named ϕ and θ:

$$\boldsymbol{Q} = \begin{pmatrix} \cos\theta & \sin\theta e^{\jmath\phi} \\ -\sin\theta e^{-\jmath\phi} & \cos\theta \end{pmatrix}$$

Considering the use of HOS for separating signals, and more precisely of fourth-order complex cumulants, we have to determine the contrast to use for this new structure of equalizer.

Definition 2: ***Trivial filters.*** *The set $\mathcal{S}$ of source processes is characterized by assumptions, such as $\boldsymbol{A1}$. One defines the set $\mathcal{T}$ of trivial filters, as containing all filters that do not affect these assumptions. In other words, $\mathcal{S}$ is stable by the operation of $\mathcal{T}$. For instance, filters of the form $\boldsymbol{\Lambda}[z] \cdot \boldsymbol{P}$, where $\boldsymbol{P}$ is a permutation matrix, and $\boldsymbol{\Lambda}[z]$ a diagonal filter, do not affect mutual independence between components of $\boldsymbol{s}(n)$. If in addition $\boldsymbol{s}(n)$ is an i.i.d. non Gaussian process, $\boldsymbol{\Lambda}[z]$ should contain only pure delays, integer multiples of the sampling period, and fixed complex factors; in other words, the entries of $\boldsymbol{\Lambda}[z]$ are of the form λz^k, with $k \in \mathbb{Z}$.*

Definition 3: ***Contrast.*** *Let $\mathcal{H}$ be a set of filters, and denote $\mathcal{H} \cdot \mathcal{S}$ the set of processes obtained by operation of filters of $\mathcal{H}$ on processes of $\mathcal{S}$. An approximation criterion, $\Upsilon(\boldsymbol{H}; \boldsymbol{x})$, will be referred to as a contrast defined on $\boldsymbol{H} \in \mathcal{H}, \boldsymbol{x} \in \mathcal{H} \cdot \mathcal{S}$, if it satisfies the three properties below [1]:*

. **Invariance:** *The contrast should not change within the set of acceptable solutions, which means that $\forall \boldsymbol{x} \in \mathcal{H}\cdot\mathcal{S}$, $\forall \boldsymbol{H} \in \mathcal{T}$ then $\Upsilon(\boldsymbol{H}; \boldsymbol{x}) = \Upsilon(\boldsymbol{I}; \boldsymbol{x})$.*

. **Domination:** *If sources are already separated, any filter should decrease the contrast. In other words, $\forall \boldsymbol{x} \in \mathcal{S}$, $\forall \boldsymbol{H} \in \mathcal{H}$, then $\Upsilon(\boldsymbol{H}; \boldsymbol{x}) \leq \Upsilon(\boldsymbol{I}; \boldsymbol{x})$.*

. **Discrimination:** *The maximum contrast should be reached only for filters linked to each other via trivial filters: $\forall \boldsymbol{x} \in \mathcal{S}, \Upsilon(\boldsymbol{H}; \boldsymbol{x}) = \Upsilon(\boldsymbol{I}; \boldsymbol{x}) \Rightarrow \boldsymbol{H} \in \mathcal{T}$.*

For convenience, we would take $\mu_i = 0$, $\forall i \in \{1, \ldots, 4\}$ in equation (8) and followings. Hence, $\Gamma^{\boldsymbol{a}}_{ik,jl}(\boldsymbol{\mu})$ is rewritten $\Gamma^{\boldsymbol{a}}_{ik,jl}$ for simplicity.

Now, let us define the contrast to use for equalizing system $\boldsymbol{G}[z]$.

Proposition 2: *The contrast defined in [3]:*

$$\boldsymbol{H} = Arg \max_{\boldsymbol{H}} \Upsilon_{1,4} \quad with \quad \Upsilon_{\alpha,4} = \sum_{i=1}^{N} \left|\Gamma_{ii,ii}^{\boldsymbol{a}}\right|^{\alpha} \tag{13}$$

with $\alpha = 1$, *is suitable for estimating equalizer* $\boldsymbol{H}[z]$.

Remark 2. For $\alpha = 1$, the previous contrast is the simplest since polynomial degrees and then time calculus (hence implementation difficulty) increase proportionally to factor α. Thus, $\alpha = 1$ is simplest than $\alpha = 2$, however precision is reduced.

Proposition 3: *Since couple of angles* (θ_i, ϕ_i) *are mutually independent for all* i *in* $\{0, \ldots, L-1\}$, *criteria defined in (13) can be rewritten as the following:*

$$\boldsymbol{H} = Arg \max_{\boldsymbol{Q}} \Upsilon_{1,4} \tag{14}$$

Proof. Indeed, for $N = 2$, criteria (13) consist of finding couples of angles, independent from other couples, thanks to $\boldsymbol{Q}$ matrices (called $\boldsymbol{R}(\theta, \phi)$ in the proof of proposition 2) which compose the equalizer $\boldsymbol{H}[z]$. $\diamondsuit$

Now expanding (11) and collecting terms involving θ or ϕ, we obtain the following equation for the output cumulants of $\boldsymbol{a}$:

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

Consider $\eta = (2\beta + \alpha - 4)$, then $\mathcal{K}_{\alpha}^{\eta}$ is a coefficient depending of indices a, b, c, d and q, r, s, t, and in accordance with α and η. Expanding equation (15):

$$\begin{aligned} \Gamma_{ii,ii}^{\boldsymbol{a}} \quad = \quad & \mathcal{K}_4^0 \cos^4\theta + \mathcal{K}_3^1 \cos^3\theta \sin\theta e^{\jmath\phi} \\ & + \mathcal{K}_3^{-1} \cos^3\theta \sin\theta e^{-\jmath\phi} + \mathcal{K}_2^0 \cos^2\theta \sin^2\theta \\ & + \mathcal{K}_2^2 \cos^2\theta \sin^2\theta e^{2\jmath\phi} + \mathcal{K}_2^{-2} \cos^2\theta \sin^2\theta e^{-2\jmath\phi} \\ & + \mathcal{K}_1^1 \cos\theta \sin^3\theta e^{\jmath\phi} + \mathcal{K}_1^{-1} \cos\theta \sin^3\theta e^{-\jmath\phi} \\ & + \mathcal{K}_1^3 \cos\theta \sin^3\theta e^{3\jmath\phi} + \mathcal{K}_1^{-3} \cos\theta \sin^3\theta e^{-3\jmath\phi} \\ & + \mathcal{K}_0^0 \sin^4\theta + \mathcal{K}_0^4 \sin^4\theta e^{4\jmath\phi} + \mathcal{K}_0^{-4} \sin^4\theta e^{-4\jmath\phi} \\ & + \mathcal{K}_0^2 \sin^4\theta e^{2\jmath\phi} + \mathcal{K}_0^{-2} \sin^4\theta e^{-2\jmath\phi} \end{aligned} \tag{16}$$

when indices q, r, s, t and a, b, c, d take their values in $\{1, 2\}$ in case of 2 i.i.d. sources. Hence, in (16) we have a total of 15 coefficients $\mathcal{K}_\alpha^\eta$. For instance, $\mathcal{K}_4^0$ corresponds to the sum described in (11) when $q = a$, $r = b$, $s = c$, and $t = d$, for any delays τ_i, $i \in [0, \ldots, L-1]$:

$$\mathcal{K}_4^0 = \sum_{qrst} \sum_{\boldsymbol{\tau}} A_{iq}(\tau_1) A_{ir}^*(\tau_2) A_{is}(\tau_3) A_{it}^*(\tau_4) \Gamma_{qs,rt}^{\boldsymbol{x}}(\boldsymbol{\tau})$$

We can simplify (16) thanks to the circularity of cumulants $\Gamma_{ii,ii}^{\boldsymbol{a}}$. Actually, $\forall i \in \mathbb{N}^*$, $\Gamma_{ii,ii}^{\boldsymbol{a}}$ is in the real field. Consequently, we have the following equalities:

$$\begin{array}{rclcrcl} \mathcal{K}_3^1 &=& (\mathcal{K}_3^{-1})^* &;& \mathcal{K}_2^2 &=& (\mathcal{K}_2^{-2})^* \\ \mathcal{K}_1^1 &=& (\mathcal{K}_1^{-1})^* &;& \mathcal{K}_1^3 &=& (\mathcal{K}_1^{-3})^* \\ \mathcal{K}_0^4 &=& (\mathcal{K}_0^{-4})^* &;& \mathcal{K}_0^2 &=& (\mathcal{K}_0^{-2})^* \end{array}$$

Indeed, we work with fourth order circular cumulants; properties of circular cumulants are defined in [8]. Moreover, one can prove that $\Gamma_{ii,ii}^{\boldsymbol{a}}$ is real (and negative for N-PSK distributions) by applying Cauchy-Schwarz inequality. Thus, denoting $\Re$ the real part of a complex number and $\Im$ its imaginary part, we can rewrite equation (16) as:

$$\begin{aligned} \Gamma_{ii,ii}^{\boldsymbol{a}} \;=\;& \mathcal{K}_4^0 \cos^4\theta + 2\cos^3\theta \sin\theta \Re(\mathcal{K}_3^1 e^{\jmath\phi}) \\ & + \cos^2\theta \sin^2\theta \left(\mathcal{K}_2^0 + 2\Re(\mathcal{K}_2^2 e^{2\jmath\phi})\right) \\ & + 2\cos\theta \sin^3\theta \left(\Re(\mathcal{K}_1^1 e^{\jmath\phi}) + \Re(\mathcal{K}_1^3 e^{3\jmath\phi})\right) \\ & + \sin^4\theta \left(\mathcal{K}_0^0 + 2\Re(\mathcal{K}_0^4 e^{4\jmath\phi}) + 2\Re(\mathcal{K}_0^2 e^{2\jmath\phi})\right) \end{aligned} \tag{17}$$

Both cumulants $\Gamma_{1111}^{\boldsymbol{a}}$ and $\Gamma_{2222}^{\boldsymbol{a}}$ are real, and so we have the expression of contrast defined in (14):

$$\Upsilon_{1,4} = \sum_i \left|\Gamma_{ii,ii}^{\boldsymbol{a}}\right|$$

In this way, we can replace $e^{\jmath\phi}$ by Euler's formula so that $e^{\jmath\phi} = \cos\phi + \jmath \sin\phi$. The relations that ensue are the followings :

$$\begin{array}{rcl} \cos 4\phi &=& (\cos 2\phi)^2 - (\sin 2\phi)^2 \\ \sin 4\phi &=& 2\sin 2\phi \cos 2\phi \\ \cos 3\phi &=& \cos 2\phi \sin\phi - \sin 2\phi \cos\phi \\ \sin 3\phi &=& \sin 2\phi \cos\phi + \cos 2\phi \sin\phi \end{array}$$

$$\begin{aligned}
\cos 2\phi &= (\cos\phi)^2 - (\sin\phi)^2 \\
\sin 2\phi &= 2\sin\phi\cos\phi \\
\cos\phi = \frac{1-t^2}{1+t^2}, &\qquad \sin\phi = \frac{2t}{1+t^2}
\end{aligned}$$

with $t = \tan\frac{\phi}{2}$ and thus $\tan\phi = \frac{2t}{1-t^2}$, and

$$\cos\theta = \frac{1}{\sqrt{1+u^2}}, \qquad \sin\theta = \frac{u}{\sqrt{1+u^2}}$$

with $u = \tan\theta$.

In order to maximize contrast expressed in (14), we have to find roots of the derivatives of (17). The reasoning is the following:

1. rewrite (17) with variables t and u,
2. partially differentiate (17) according to variables t and u, individually,
3. find all pairs of roots (u,t) solving the system of polynomials,
4. test each pair to select the maximum of $\Upsilon_{1,4}$.

Thus, in a first stage, we obtain a system Σ of 2 polynomials denoted $\Phi_1(u,t)$ and $\Phi_2(u,t)$ in 2 unknowns (u and t) :

$$\Sigma = \begin{cases} \Phi_1(u,t) = \dfrac{\partial\Upsilon_{1,4}}{\partial u} \\[2ex] \Phi_2(u,t) = \dfrac{\partial\Upsilon_{1,4}}{\partial t} \end{cases}$$

Polynomial $\Phi_1(u,t)$ is of global degree 12 (8 in variable t and 4 in u) whereas $\Phi_2(u,t)$ is of global degree 11 (8 in t and 3 in u). Next, components of those polynomials are denoted likewise :

$$\Phi_1(u,t) = \sum_{k=0}^{4} \lambda_{4-k}(t)\ u^k \tag{18}$$

$$\Phi_2(u,t) = \sum_{k=0}^{3} \xi_{3-k}(t)\ u^k \tag{19}$$

System Σ can be solved by using the resultant of a Sylvester matrix. Hence, considering only variable z for $\Phi_1(u,t)$ and $\Phi_2(u,t)$, and collecting

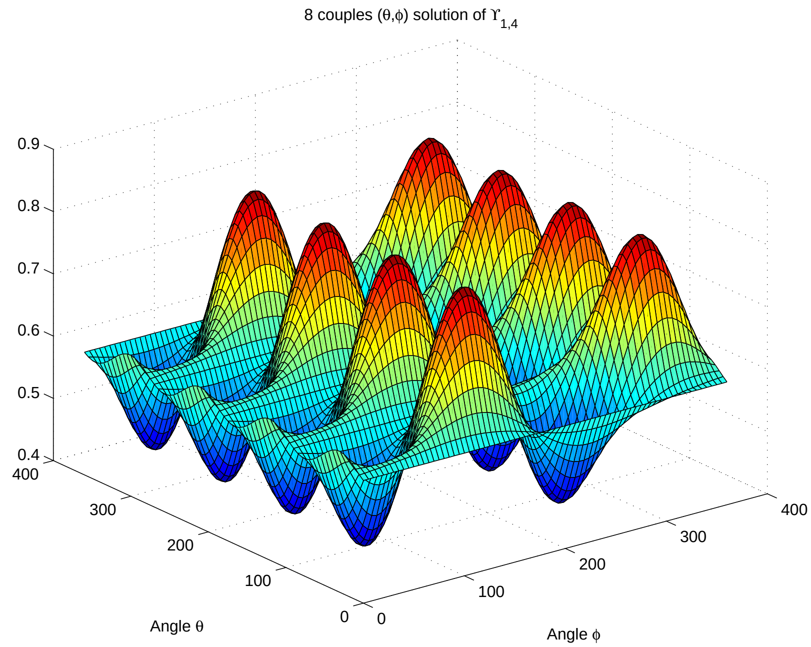

Fig. 4: Maximization of the criteria over θ and ϕ.

terms of same degree in u, we obtain a matrix of Sylvester of size 7×7. This size corresponds to the sum of degrees in u of the two polynomials, *i.e.* 4 for $\Phi_1(u,t)$ and 3 for $\Phi_2(u,t)$.

$$\begin{vmatrix} \xi_0(t) & 0 & 0 & 0 & \lambda_0(t) & 0 & 0 \\ \xi_1(t) & \xi_0(t) & 0 & 0 & \lambda_1(t) & \lambda_0(t) & 0 \\ \xi_2(t) & \xi_1(t) & \xi_0(t) & 0 & \lambda_2(t) & \lambda_1(t) & \lambda_0(t) \\ \xi_3(t) & \xi_2(t) & \xi_1(t) & \xi_0(t) & \lambda_3(t) & \lambda_2(t) & \lambda_1(t) \\ 0 & \xi_3(t) & \xi_2(t) & \xi_1(t) & \lambda_4(t) & \lambda_3(t) & \lambda_2(t) \\ 0 & 0 & \xi_3(t) & \xi_2(t) & 0 & \lambda_4(t) & \lambda_3(t) \\ 0 & 0 & 0 & \xi_3(t) & 0 & 0 & \lambda_4(t) \end{vmatrix}$$

Remark 3. Global degree of $\Phi_2(u,t)$ has been reduced from 12 to 11 (4 to 3 in variable u) thanks to an obvious root $u = 0$. Therefore equation (19) is already divided by u. Since a valid pair of roots (u,t) must satisfy both of the polynomials of Σ, we have 2 roots in variable t, reported in polynomial $\Phi_1(u,t)$:

$$t_{u=0} = \frac{y \pm \sqrt{y^2 + x^2}}{x}$$

where $x = \Re\left[\mathcal{K}_3^1\right] + \Re\left[(\mathcal{K}_3^1)^*\right] = 2\Re\left[\mathcal{K}_3^1\right]$ and $y = \Im\left[(\mathcal{K}_3^1)^*\right] - \Im\left[\mathcal{K}_3^1\right] = -2\Im\left[\mathcal{K}_3^1\right]$. Moreover, other relations like those are used in the algorithm in order to avoid imaginary parts residual. Indeed, only real parts are used in computing sequences.

Furthermore, in order to reduce time calculus, we compute only the tensor of cumulants observed from the output of filter $\boldsymbol{B}[z]$. Hence we need only (11) for solving system Σ. As a result, we do not take care of $\boldsymbol{B}[z]$ in the system and thus we reduce time computation cost.

In addition, we can reduce loops in the algorithm by using a $N^4 \times N^4$ matrix for $\Gamma^{\boldsymbol{a}}_{ik,jl}$. Actually, (11) uses 8 indices denoted a, b, c, d, q, r, s, t each varying in $\{1, \ldots, N\}$. Thus, we have N^8 values for each τ_i fixed (*e.g.* 256 values for $N = 2$).

4 Algorithm

In this section we present the new algorithm derived from previous statements. It has been implemented and results are shown in section 5. Tensors are grouped like in [4]. The algorithm, for $N = 2$ sources, is summarized in figure 5. Describe each step of this algorithm:

(1) <u>Compute</u> tensor of observations (8)
(2) <u>Initialize</u> $\boldsymbol{A}$ and $\boldsymbol{B}$
(3) <u>for</u> $k = 1$ to L,
    (a) <u>Compute</u> tensor (12)
    (b) <u>Maximize</u> contrast (14)
    (c) <u>Actualize</u> $\boldsymbol{A}$ and $\boldsymbol{B}$ with angles
  end;

Fig. 5: Summary of the algorithm

1. Compute the tensor of cumulants $\Gamma^{\boldsymbol{w}}_{eg,fh}(\boldsymbol{\tau} + \boldsymbol{p})$ defined in (8) and of length $L = \max\{\boldsymbol{\tau}\} + \max\{\boldsymbol{p}\}$. Cumulants are standardized since channel is paraunitary.

2. Initialize the equalizer with zero angles for filters $\boldsymbol{A}[z]$ and $\boldsymbol{B}[z]$,

3. Loop on the sweeps: $k = 1, \ldots, L$.

   (a) Compute the cumulant tensor $\Gamma^{\boldsymbol{x}}_{ac,bd}(\boldsymbol{\tau})$ of $\boldsymbol{x}$ (defined in (12)). The resulting tensor is composed of $N^2(L - L_b + 1)^2$ matrices each of size $N(L - L_b + 1) \times N(L - L_b + 1)$.

(b) Search couple (θ_k,ϕ_k) maximizing $\Upsilon_{1,4}$ defined in (14). Then angles are reported in $\boldsymbol{R}(\theta_k, \phi_k)$ for the global equalizer.

(c) Actualize filters $\boldsymbol{A}[z]$ and $\boldsymbol{B}[z]$ with all known angles (θ_i, ϕ_i) for $i \in [1, \ldots, k)$.

4. Build equalizer $\boldsymbol{H}[z]$ and separate observations $\boldsymbol{w}$.

This algorithm consider that angles are all independent.

Remark 4. This algorithm is not yet optimized. Indeed, we suggest two optimization tips:

- *Increasing precision* : in order to accurate precision of angles, a loop on $j = 1, \ldots, \varepsilon$ can be inserted bellow the loop on k. Effectively, in the first loop, *i.e.* $k = 1$, angles (θ_i, ϕ_i) for $i \in [2, \ldots, L-1)$ are null and hence values of (θ_1, ϕ_1) are approximative. If a second loop on j is inserted, it will re-calculate and increase precision of (θ_1, ϕ_1) and following couple since $(\theta_{L-1}, \phi_{L-1})$ and previous couple have been determined approximatively in the first loop. Hence, the largest range for k we put, the highest precision we get.

- *Reducing time calculus* : during initialization $\boldsymbol{A}[z]$ and $\boldsymbol{B}[z]$ are build from $\boldsymbol{Z}[z]$ and $\boldsymbol{R}(\theta_i, \phi_i)$ with $\forall i, \theta_i = 0, \phi_i = 0$. Hence, tensors components of $\boldsymbol{x}$, $\boldsymbol{y}$, or $\boldsymbol{a}$ are taken from tensor of observed cumulants of $\boldsymbol{w}$. Thus, we can use the quadratic form of the contrast $\Upsilon_{1,4}$ in order to find eigenvectors of a 3×3 real symmetric matrix, as described in [2].

5 Computer Results

One consider a FIR complex mixture of length $L = 3$ of $N = 2$ QPSK white processes. The channels are paraunitary in order to preserve second-order whiteness and are constructed as explain in section 2. Thus, the 6 angles θ_i, ϕ_i, for $0 \leq i \leq L-1$, are drawn according to a uniform distribution in $[0, 2\pi)$ in order to generate paraunitary channels. For each randomly generated channel, block of noisy observations are generated according to

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

where $\boldsymbol{v}(n)$ is a white circular complex Gaussian noise with identity covariance matrix, and $s_i(n)$ are unit variance QPSK white sequences. Parameter ρ is introduced in order to control the Signal to Noise Ratio per bit (SNR), and defined as follows:

$$SNR_{db} = -20\log_{10}\rho$$

Figure 6 shows Symbol Error Rates (BER) for blocks of 500, and 1500

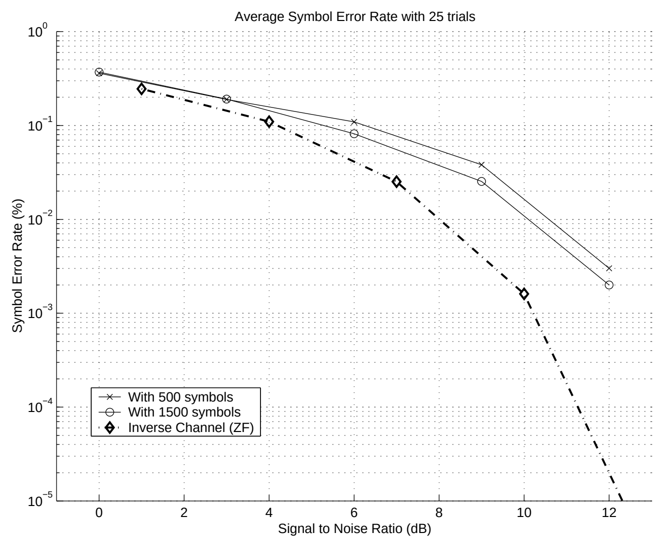

Fig. 6: Symbol Error Rate obtained when a 3-length equalizer is built from blocks of 500 or 1500 symbols.

symbols over 25 trials. The minimal resolution is $(1500 * 25)^{-1} = 2,6.10^{-5}$. Figure 6 presents median results of the 25 trials, *i.e.* average SER. As a basis for comparisons, average BER obtained with Zero-Forcing method is represented too. These curves demonstrate the good performances obtained for short data block, *i.e.* 500 symbols only.

6 Concluding remarks

Through this paper, we have presented a new model for paraunitary equalizer in order to respect the paraunitarity of the channel. Indeed, under this constraint but with this new model of equalizer, we have demonstrated that an algorithm could be performed. Then, from theoretical results of section 2, a numerical algorithm has been implemented in order to measure performances. Results obtained are very attractive since the algorithm works very well with a 3-length equalizer. Moreover, optimization tips have been proposed in section 4 in order to accurate precision and to reduce time calculus.

References

[1] P. COMON. Contrasts for multichannel blind deconvolution. *Signal Processing Letters*, 3(7):209–211, July 1996.

[2] P. COMON. From source separation to blind equalization. In *Int. Conf. on Image and Signal Processing (ICISP)*, pages 20–32, Agadir, Morocco, May 3-5 2001. invited plenary.

[3] P. COMON and E. MOREAU. Improved contrast dedicated to blind separation in communications. In *ICASSP*, pages 3453–3456, Munich, April 20-24 1997.

[4] 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*, pages 686–691, San Diego, Dec. 9-13 2001.

[5] D. DONOHO. On minimum entropy deconvolution. In Academic Press, editor, *in Applied time series analysis II*, pages 565–609, 1981.

[6] I. FIJALKOW, A. TOUZNI, and J. R. TREICHLER. Fractionally spaced equalization using CMA: Robustness to channel noise and lack of disparity. In *IEEE Trans. Sig. Proc.*, volume 45, pages 56–66, Jan. 1997.

[7] O. GRELLIER, P. COMON, B. MOURRAIN, and P. TREBUCHET. Analytical blind channel identification. *IEEE Trans. Signal Processing*, 50(9), September 2002.

[8] J. L. LACOUME, P. O. AMBLARD, and P. COMON. *Statistiques d'ordre supérieur pour le traitement du signal.* Collection Sciences de l'Ingénieur. Masson, 1997.

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

[10] O. SHALVI and E. WEINSTEIN. New criteria for blind deconvolution of nonminimum phase systems. *IEEE Trans. Inf. Theory*, 36(2):312–321, Mar. 1990.

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

[12] A. J. van der VEEN and A. PAULRAJ. An analytical constant modulud algorithm. In *IEEE Trans. Sig. Proc.*, volume 44, pages 1136–1155, May 1996.

[13] D. YELLIN and E. WEINSTEIN. Criteria for multichannel signal separation. In *IEEE Trans. Sig. Proc.*, volume 42, pages 2158–2168, Aug. 1994.