

Blind Paraunitary Equalization

Sylvie Icart, Pierre Comon,

Laboratoire I3S, CNRS UMR 6070, Université de Nice, 2000 route des Lucioles, BP 121, 06903 Sophia-Antipolis Cedex France

Ludwig Rota

IMRA Europe, Sophia-Antipolis

Abstract

In this paper a blind MIMO space-time equalizer is described, dedicated to convolutive mixtures when observations have been pre-whitened. Filters preserving space-time whiteness are paraunitary; a parameterization of such filters is proposed. Theoretical developments then lead to a numerical algorithm that sweeps all pairs of delayed outputs. This algorithm involves the solution of a polynomial system, whose coefficients depend on the output cumulants. Simulations and performance of the numerical algorithm are reported.

Key words: Blind deconvolution, Blind equalization, Multiple-Input/Multiple-Output (MIMO), High Order Statistics (HOS), Paraunitary filter.

1. Introduction

In actual digital communication systems, the equalization problem is solved using learning sequences. These sequences, known by transmitter and receiver, permit to estimate channel parameters. Nevertheless, in the future, learning sequences may be seen as reducing the throughput, because they occupy a non negligible space in transmitted sequences. Hence less actual information is transmitted, *i.e.* useful data rate is lower than system data rate. Next, in some situations, *e.g.* interception in electronic warfare, learning sequences exist but are not known to the receiver.

Blind separation methods for Multiple-Input Multiple-Output (MIMO) channels have raised an increasing interest for digital communications since they do not need learning sequences. Most blind MIMO equalization techniques use High Order Statistics (HOS) for separating signals

[1,2,3,4,5,6,7,8]. This can be implicit through constant modulus [9,10,11,12] or constant power [13] criteria. Purely deterministic approaches also exist and exploit either the finite alphabet property [14,15,16] or the presence of another diversity in addition to time and space [17,18].

Our main contribution consists of a block algorithm dedicated to blind MIMO equalization and exploiting HOS of observations [19]. The particularity of our method is that it is based on a factorization of paraunitary filters [20], as suggested in [21,22]. The paraunitary condition is not very restrictive since prewhitening can always be performed in a first stage (in a non unique manner) even if it is not always an obvious operation [23,22,24,25]. Moreover, the algorithm designed herein can be implemented either "on-line" or "off-line". More precisely, our algorithm can run iterations on a single given block until convergence, or on the contrary, run a new iteration after each symbol arrival. Convergence of on-line algorithms is known to be much longer in terms of number of symbols required (typically sev-

Email address: Sylvie.Icart@unice.fr (Sylvie Icart).

Article published in Signal Processing 89 (2009) 283–291

eral thousands of symbols).

Note that algorithms dedicated to paraunitary channels as in [26] or PAJOD [27], have already been proposed. Unfortunately, the paraunitary constraint was not accurately verified in [27], especially at low SNR, because the channel was parametrized with a semi-unitary matrix.

The paper is organized as follows. In the second section, model and hypotheses are presented. The parameterization of MIMO paraunitary filters is introduced in Section 3, and input-output cumulant relations are established. A contrast criterion is proposed in Section 4, for finding equalizer parameters and separating source signals. An iterative algorithm built from previous theoretical results is described in the fifth section. The core of the algorithm is a Jacobi-type iteration, in which all possible Givens rotations in the parameterization of the paraunitary filter are swept until convergence. Finally, the performance for various Signal-to-Noise Ratios is illustrated in the last section.

2. Model and notations

Throughout the paper, (\top) stands for transposition, $(^H)$ for conjugate transposition, $(^*)$ for complex conjugation, and $j = \sqrt{-1}$. Also denote by \mathbb{Z} the set of integers, by \mathbb{N} the subset of positive integers and by z^{-1} the time-delay operator. Vectors and matrices are denoted with bold lowercase and bold uppercase letters respectively. The entries of matrix \mathbf{G} are denoted G_{ij} , where subscript ij denotes the i -th row and the j -th column of \mathbf{G} .

Let us consider a digital communication system of N antennas and N receivers in a multipath environment. Let $\mathbf{s}(n) = (s_1(n), \dots, s_N(n))^T$ denote the N -dimensional source vector of baseband complex signals at time n and $\mathbf{w}(n) = (w_1(n), \dots, w_N(n))^T$ the N -dimensional observation vector. Let $\mathbf{C}[z]$ be the transfer function of the Linear Time Invariant (LTI) mixing channel and $\{\mathbf{C}(k), k = 0, \dots, K\}$ be the sequence of $N \times N$ matrices of the complex Finite Impulse Response (FIR) of the channel. We then have:

$$\mathbf{w}(n) = \sum_{k=0}^K \mathbf{C}(k)\mathbf{s}(n-k) \quad (1)$$

where K denotes the memory length of the channel. Also denote

$$\mathbf{C}[z] = \sum_{m=0}^K \mathbf{C}(m)z^{-m}. \quad (2)$$

The case of instantaneous mixtures is not considered in the present paper, since already addressed via pairwise processing by various authors since 1991. Thus, it will be assumed that $K > 0$.

The multichannel blind deconvolution problem consists of finding a LTI filter $\mathbf{H}[z]$, *i.e.* the *equalizer*, in order to retrieve the N input signals $s_i(n), i \in \{1, \dots, N\}, \forall n \in \mathbb{Z}$, solely from the observation of the outputs $\mathbf{w}(n)$ of the unknown LTI channel $\mathbf{C}[z]$.

Let $\hat{\mathbf{s}}(n) = (\hat{s}_1(n), \dots, \hat{s}_N(n))^T$ be the N -dimensional estimated source vector. This means, with the above notation:

$$\hat{\mathbf{s}}(n) = \sum_{l=0}^L \mathbf{H}(l)\mathbf{w}(n-l) \quad (3)$$

where L is the memory length of the equalizer, or, as a function of original sources themselves:

$$\hat{\mathbf{s}}(n) = \sum_{m=0}^{K+L} \mathbf{G}(m)\mathbf{s}(n-m), \quad (4)$$

where $\mathbf{G}[z]$ denotes the global filter $\mathbf{G}[z] = \mathbf{H}[z]\mathbf{C}[z]$ (see fig 1 below).

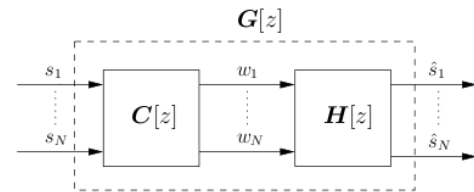


Fig. 1. Global system \mathbf{G} : sources s_i are filtered by channel $\mathbf{C}[z]$ and observations w_i are equalized by $\mathbf{H}[z]$.

Definition 1 Paraunitarity. A square polynomial matrix $\mathbf{H}[z] \in \mathbb{C}^{N \times N}$ is said to be paraunitary [21] if

$$\mathbf{H}^H[1/z^*]\mathbf{H}[z] = \mathbf{I}_N = \mathbf{H}[z]\mathbf{H}^H[1/z^*] \quad (5)$$

where $\mathbf{I}_N \in \mathbb{R}^{N \times N}$ is the identity matrix.

In this paper, we assume the following hypotheses:

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

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

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

do not depend on n . For definitions of cumulants, refer to [28] and references therein.

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

(H4) $\mathbf{C}[z]$, $\mathbf{H}[z]$, and hence $\mathbf{G}[z] = \mathbf{H}[z]\mathbf{C}[z]$ are all paraunitary. Thus we have the global relation:

$$\mathbf{H}[z]\mathbf{C}[z]\mathbf{C}^H[1/z^*]\mathbf{H}^H[1/z^*] = \mathbf{I}_N. \quad (6)$$

Remark 2 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 [23,22]. In other words, paraunitary filters are relevant after a space-time standardization of observations (second order white with unit covariance). The way space-time whitening is implemented is out of the scope of this paper, and it is assumed that observations $\mathbf{w}(n)$ are space-time white.

As is now well known, statistical independence of \hat{s}_i can only allow to blindly recover the sources up to a permutation matrix \mathbf{P} , and a diagonal delay filter $\mathbf{\Lambda}[z]$, so that $\mathbf{H}[z]\mathbf{C}[z] = \mathbf{\Lambda}[z]\mathbf{P}$.

3. Equalization

We introduce in this section the global parameterization of the paraunitary equalizer, which will allow in Section 5 an implementation in terms of Givens rotations.

3.1. Parameterization

From our observation model and from hypotheses described in Section 2, we have a first proposition:

Proposition 3 Let $\mathbf{H}[z] \in \mathbb{C}^{N \times N}$ be a FIR paraunitary filter of MacMillan degree ¹ λ_h . Then matrix $\mathbf{H}[z]$ can be factorized (non uniquely) as the product:

$$\mathbf{H}[z] = \mathbf{A}[z]\mathbf{Q}\mathbf{B}[z] \quad (7)$$

where $\mathbf{Q} \in \mathbb{C}^{N \times N}$ is unitary and $\mathbf{A}[z]$ and $\mathbf{B}[z]$ are FIR paraunitary filters of MacMillan degree ℓ_A and

¹ The MacMillan degree of a paraunitary filter is also the degree of its determinant [21, p. 737].

ℓ_B respectively, with $0 \leq \ell_A \leq \lambda_h$, $0 \leq \ell_B \leq \lambda_h$, $\ell_A + \ell_B = \lambda_h$.

PROOF. The proof is trivial since we can extend the paraunitary factorization of [29] to the N -dimensional complex case:

$$\mathbf{H}[z] = \mathbf{Q}_0\mathbf{Z}[z]\mathbf{Q}_1 \dots \mathbf{Z}[z]\mathbf{Q}_{\lambda_h-1}\mathbf{Z}[z]\mathbf{Q}_{\lambda_h} \quad (8)$$

where $\mathbf{Q}_p \in \mathbb{C}^{N \times N}$ are unitary for $p \in \{0, \dots, \lambda_h\}$, λ_h is the MacMillan degree of $\mathbf{H}[z]$, and $\mathbf{Z}[z]$ denotes the $N \times N$ diagonal matrix

$$\mathbf{Z}[z] = \begin{bmatrix} \mathbf{I}_{N-1} & 0 \\ 0 & z^{-1} \end{bmatrix}. \quad (9)$$

Thus, for $0 < \ell_A \leq p$, we have

$$\mathbf{A}[z] = \mathbf{Q}_0\mathbf{Z}[z] \dots \mathbf{Q}_{p-1}\mathbf{Z}[z] \quad (10)$$

and for $0 < \ell_B \leq \lambda_h - p$, we have

$$\mathbf{B}[z] = \mathbf{Z}[z]\mathbf{Q}_{p+1} \dots \mathbf{Z}[z]\mathbf{Q}_{\lambda_h}. \quad (11)$$

When $\ell_A = 0$ (respectively $\ell_B = 0$), we can replace $\mathbf{A}[z]$ (respectively $\mathbf{B}[z]$) by \mathbf{I}_N . Then, for a fixed p , any paraunitary filter of MacMillan degree λ_h can be factorized like in (7).

3.2. MIMO relations

In the remaining, we assume the following notation for fourth order cumulants, e.g. cumulants of vector \mathbf{w} ,

$$\Gamma_{eg, fh}^{\mathbf{w}}(\boldsymbol{\nu}) = \text{Cum}[w_e(n - \nu_1), w_f^*(n - \nu_2), w_g(n - \nu_3), w_h^*(n - \nu_4)]$$

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

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

$$\hat{s}_i(n) = \sum_{m=0}^{\ell_A} \sum_{r=1}^N \sum_{q=1}^N A_{iq}(m) Q_{qr} x_r(n-m), \quad (13)$$

and

$$x_r(n-m) = \sum_{l=0}^{\ell_B} \sum_{s=1}^N B_{rs}(l) w_s(n-m-l). \quad (14)$$

From (13) and thanks to the multilinearity property of cumulants, we can express the input-output relations between cumulants of input \mathbf{x} and output $\hat{\mathbf{s}}$,

$$\Gamma_{ik, jl}^{\hat{\mathbf{s}}} = \sum_{abcd} \sum_{\boldsymbol{\tau}} \sum_{qrst} A_{iq}(\tau_1) A_{jr}^*(\tau_2) A_{ks}(\tau_3) A_{lt}^*(\tau_4) Q_{qa} Q_{rb}^* Q_{sc} Q_{td}^*$$

with $\boldsymbol{\tau} = (\tau_1, \tau_2, \tau_3, \tau_4)$. Since $\hat{\mathbf{s}}$ is the output of filter $\mathbf{A}[z]\mathbf{Q}$, the range of each τ_i is $[0, \dots, \ell_A]$ and all indices $a, b, c, d, i, j, k, l, q, r, s, t$ take their values in $\{1, \dots, N\}$. Then, we can write a similar relation between cumulants of input \mathbf{w} and output \mathbf{x} ,

$$\Gamma_{ac,bd}^{\mathbf{x}}(\boldsymbol{\tau}) = \sum_{\boldsymbol{\rho}} \sum_{efgh} B_{ae}(\rho_1) B_{bf}^*(\rho_2) B_{cg}(\rho_3) B_{dh}^*(\rho_4) \Gamma_{eg,fd}^{\mathbf{w}}(\boldsymbol{\tau})$$

with $\boldsymbol{\rho} = (\rho_1, \rho_2, \rho_3, \rho_4)$. Here, the range of each ρ_i is $[0, \dots, \ell_B]$ since inputs are filtered by $\mathbf{B}[z]$.

3.3. Construction of \mathbf{Q}_p

Matrices \mathbf{Q}_p involved in the paraunitary factorization (8) are unitary. It is well known that any $N \times N$ unitary matrix can be written as a product of $M = N(N-1)/2$ Givens rotations [30, p. 215] $U(i, j, \theta, \phi)$ with $1 \leq i < j \leq N$ and

$$U(i, j, \theta, \phi) = \begin{bmatrix} \mathbf{I}_{n_1} & \vdots & \mathbf{0} & \vdots & \mathbf{0} \\ \cdots & \cos \theta & \cdots & \sin \theta e^{j\phi} & \cdots \\ \mathbf{0} & \vdots & \mathbf{I}_{n_2} & \vdots & \mathbf{0} \\ \cdots & -\sin \theta e^{-j\phi} & \cdots & \cos \theta & \cdots \\ \mathbf{0} & \vdots & \mathbf{0} & \vdots & \mathbf{I}_{n_3} \end{bmatrix}$$

where $n_1 + n_2 + n_3 = N - 2$. These unitary rotations allow to develop a *Jacobi*-type iteration algorithm for equalizer (7).

Since we use Givens matrices, we can make 2 remarks.

Remark 4 For $N = 2$, unitary matrix \mathbf{Q}_p consists of a single rotation

$$\mathbf{Q}_p = \begin{bmatrix} \cos \theta_p & \sin \theta_p e^{j\phi_p} \\ -\sin \theta_p e^{-j\phi_p} & \cos \theta_p \end{bmatrix}. \quad (17)$$

Then, only one pair of angles (θ_p, ϕ_p) is needed for each \mathbf{Q}_p . Hence, when more than two signals are observed, it is also possible to use a deflation-type algorithm [31].

Remark 5 Proposition 3 tells that $\lambda_h + 1$ unitary matrices are used to obtain a complete factorization of the equalizer. For instance, the 2×2 length-1 paraunitary filter defined by $\text{diag}(\mathbf{H}[z]) = [-z^{-1}, z^{-1}]$ and 0 elsewhere, needs $\lambda_h = \text{deg}[\det(\mathbf{H}[z])] = 2$ diagonal matrices $\mathbf{Z}[z]$. Indeed, $\mathbf{H}[z]$ has the following factorization

$$\mathbf{H}[z] = \mathbf{Q}_2 \cdot \mathbf{Z}[z] \cdot \mathbf{Q}_1 \cdot \mathbf{Z}[z] \cdot e^{j\frac{\pi}{2}} \quad (18)$$

where $\mathbf{Q}_2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ and $\mathbf{Q}_1 = \begin{bmatrix} 0 & j \\ j & 0 \end{bmatrix}$, i.e. $(\theta_2 = \frac{\pi}{2}, \phi_2 = 0)$ and $(\theta_1 = \frac{\pi}{2}, \phi_1 = \frac{\pi}{2})$.

The exponential is part of the unknown matrix $\mathbf{\Lambda}[z]\mathbf{P}$, an indetermination inherent in the problem (7) (Proposition (16) in Section 2). It is also true for $L \geq 1$ since we can group all exponentials in a single matrix $\mathbf{\Lambda}[z]\mathbf{P}$. Thus, reflections, e.g. $\text{diag}(\mathbf{Q}) = [1, -1]$, are covered by the parameterization of Proposition 3, thanks to the inherent indeterminacy. The parameterization (7) is thus complete.

4. Optimization criterion

Let us start this section with two key definitions. The first definition deals with the set of filters that do not affect the independence of sources, a key point of blind source separation methods. The second definition describes the main properties of a *contrast*, i.e. an optimization criterion for separating signals.

Definition 6 Trivial filters. The set \mathcal{S} of source J processes is characterized by assumptions, such as **H1**. 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 $\mathbf{\Lambda}[z]\mathbf{P}$, where \mathbf{P} is a permutation matrix, and $\mathbf{\Lambda}[z]$ a diagonal filter, do not affect mutual independence between components of $\mathbf{s}(n)$. If in addition $\mathbf{s}(n)$ is an i.i.d. non Gaussian process, $\mathbf{\Lambda}[z]$ should contain only pure delays, integer multiples of the sampling period, and fixed complex factors; in other words, the entries of $\mathbf{\Lambda}[z]$ are of the form γz^k , with $k \in \mathbb{Z}$ and $|\gamma| = 1$.

Definition 7 Contrast. Let \mathcal{H} be a set of filters, and denote $\mathcal{H} \cdot \mathcal{S}$ the set of standardized linear processes obtained by operation of filters of \mathcal{H} on processes of \mathcal{S} . Denote \mathbf{I} the identity matrix. An optimization criterion, $\Upsilon(\mathbf{H}; \mathbf{w})$, will be referred to as a contrast for $\mathbf{H} \in \mathcal{H}, \mathbf{w} \in \mathcal{H} \cdot \mathcal{S}$, if it satisfies the three properties below [2]:

- **Invariance:** The contrast should not change within the set of acceptable solutions, which means that $\forall \mathbf{w} \in \mathcal{H} \cdot \mathcal{S}, \forall \mathbf{H} \in \mathcal{T}$ then $\Upsilon(\mathbf{H}; \mathbf{w}) = \Upsilon(\mathbf{I}; \mathbf{w})$.
- **Domination:** If sources are already separated, any filter should decrease the contrast i.e. $\forall \mathbf{w} \in \mathcal{S}, \forall \mathbf{H} \in \mathcal{H}$, then $\Upsilon(\mathbf{H}; \mathbf{w}) \leq \Upsilon(\mathbf{I}; \mathbf{w})$.

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

Now, let us focus on the contrast used to carry out the equalization of the system. We can separate signals blindly by using an approach based on High Order Statistics. Indeed, since fourth-order cross-cumulants are a good measure of statistical independence, we can build a contrast based on these values [2]:

Proposition 8 *Consider sources having kurtosis of the same sign ϵ , $\epsilon = \pm 1$. Then, the source separation can be performed solely from channel outputs by maximizing the following contrast*

$$\Upsilon_{1,4} = \epsilon \sum_{i=1}^N \Gamma_{ii,ii}^{\hat{\mathbf{s}}} \quad (19)$$

where cumulants of $\hat{\mathbf{s}}$ are defined by

$$\Gamma_{ii,ii}^{\hat{\mathbf{s}}} = \sum_{abcd} \sum_{\tau} \sum_{qrst} A_{iq}(\tau_1) A_{ir}^*(\tau_2) A_{is}(\tau_3) A_{it}^*(\tau_4) Q_{qa} Q_{rb}^* Q_{sc} Q_{td}^* \quad (20)$$

Hence, in order to estimate $\mathbf{H}[z]$, the optimization problem can be rewritten as:

$$\mathbf{H} = \underset{\mathcal{Q}}{\text{Arg max}} \Upsilon_{1,4} \quad (21)$$

where \mathcal{Q} stands for the set of all \mathbf{Q}_p .

It has been proved in [2] that $\Upsilon_{1,4}$ is a contrast. We maximize $\Upsilon_{1,4}$ with respect to each pair (θ, ϕ) in turn. The sequence of values of $\Upsilon_{1,4}$ obtained this way is monotonically increasing. Since it is also bounded above, it converges. Figure 2) is an example of such a behaviour.

5. A Jacobi-type algorithm

In this section we detail an iteration of the algorithm for one pair (θ, ϕ) of Givens rotation. Then, as in *Jacobi sweeping* algorithms for matrices [30], we sweep $M = N(N - 1)/2$ pairs for each unitary matrix \mathbf{Q}_p .

5.1. Processing a pair of outputs

For the sake of clarity, we drop index p . We have to find all pairs (θ, ϕ) that maximize (19), the other pairs being fixed. To reach this goal, (20) is simplified firstly by expanding it, and secondly by collecting terms involving θ or ϕ . In this manner, we obtain

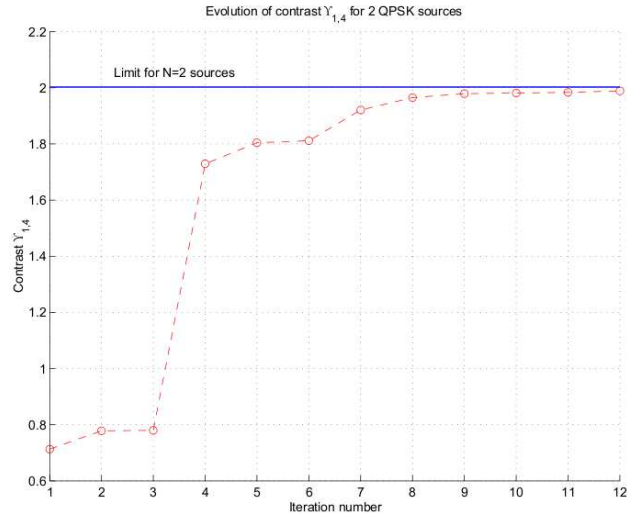


Fig. 2. Typical evolution of contrast $\Upsilon_{1,4}$ for 25dB of SNR and length-3 equalizer.

the output cumulants of $\hat{\mathbf{s}}$ as a linear combination of terms of the form $\cos^\alpha \theta \sin^{4-\alpha} \theta e^{j\eta\phi}$. More precisely:

$$\begin{aligned} \Gamma_{ii,ii}^{\hat{\mathbf{s}}} = & \mathcal{K}_{(4)}^{(0)} \cos^4 \theta + \mathcal{K}_{(3)}^{(1)} \cos^3 \theta \sin \theta e^{j\phi} + \mathcal{K}_{(3)}^{(-1)} \cos^3 \theta \sin \theta e^{-j\phi} + \\ & + \mathcal{K}_{(2)}^{(2)} \cos^2 \theta \sin^2 \theta e^{2j\phi} + \mathcal{K}_{(2)}^{(-2)} \cos^2 \theta \sin^2 \theta e^{-2j\phi} + \mathcal{K}_{(1)}^{(1)} \cos \\ & + \mathcal{K}_{(1)}^{(-1)} \cos \theta \sin^3 \theta e^{-j\phi} + \mathcal{K}_{(1)}^{(3)} \cos \theta \sin^3 \theta e^{3j\phi} + \mathcal{K}_{(1)}^{(-3)} \cos \\ & + \mathcal{K}_{(0)}^{(0)} \sin^4 \theta + \mathcal{K}_{(0)}^{(4)} \sin^4 \theta e^{4j\phi} + \mathcal{K}_{(0)}^{(-4)} \sin^4 \theta e^{-4j\phi} \\ & + \mathcal{K}_{(0)}^{(2)} \sin^4 \theta e^{2j\phi} + \mathcal{K}_{(0)}^{(-2)} \sin^4 \theta e^{-2j\phi}. \end{aligned}$$

where

$$\mathcal{K}_{(\alpha)}^{(\eta)} e^{j\eta\phi} = \sum_{\tau} A_{iq}(\tau_1) A_{ir}^*(\tau_2) A_{is}(\tau_3) A_{it}^*(\tau_4) \Gamma_{ac,bd}^{\mathbf{x}}(\tau). \quad (23)$$

According to (17), (20) and (23), each coefficient $\mathcal{K}_{(\alpha)}^{(\eta)}$ depends on indices a, b, c, d and q, r, s, t and as the product $Q_{qa} Q_{rb}^* Q_{sc} Q_{td}^*$ gives an expression of the form $\cos^\alpha \theta \sin^{4-\alpha} \theta e^{j\eta\phi}$ with $\alpha \in \{0, \dots, 4\}$ and $\eta \in \{\pm 1, \pm 2, \pm 3, \pm 4\}$; let $\eta = 2\beta + \alpha - 4$, we get

$$\Gamma_{ii,ii}^{\hat{\mathbf{s}}} = \sum_{\alpha=0}^4 \sum_{\beta=0}^{4-\alpha} \mathcal{K}_{(\alpha)}^{(2\beta+\alpha-4)} \cos^\alpha \theta \sin^{4-\alpha} \theta e^{j(2\beta+\alpha-4)\phi}. \quad (24)$$

Remark 9 *Computing each $\Gamma_{ii,ii}^{\hat{\mathbf{s}}}$ without simplifications require a total of $(L + 1)^4 N^8$ loops (equation (20)). Thus, readers can find an implementation trick in Appendix A for reducing this number of loops to $((L + 1)N)^4$. Moreover, the method detailed in Appendix A uses less memory for storing matrices. Only the cumulant tensor $\Gamma_{ac,bd}^{\mathbf{x}}(\tau)$, which*

is computed at the beginning of each loop, needs a large memory, i.e. up to $N^4(L+1)^4$ values to store.

Next, we 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 $\cos \theta = \frac{1-u^2}{\sqrt{1+u^2}}$, $\sin \theta = \frac{u}{\sqrt{1+u^2}}$ with $u = \tan \theta$. Thus, the numerator and the denominator of the polynomial obtained are both of degree 8 in variable t and 4 in variable u . For maximizing contrast (19), we have to find roots of polynomial system (25), i.e. stationary points of $\Upsilon_{1,4}$, obtained by canceling the partial derivatives:

$$\left. \begin{aligned} \Phi_1(u, t) &\stackrel{\text{def}}{=} \frac{\partial \Upsilon_{1,4}}{\partial u} \\ \Phi_2(u, t) &\stackrel{\text{def}}{=} \frac{\partial \Upsilon_{1,4}}{\partial t} \end{aligned} \right\} \quad (25)$$

After some simplifications, e.g. obvious root $u = 0$ in $\Phi_2(u, t)$, polynomials obtained have different global degrees: degree 12 for $\Phi_1(u, t)$ (leading monomial is $t^8 u^4$), and degree 11 for $\Phi_2(u, t)$ (leading monomial is $t^8 u^3$). Then we consider only variable u for $\Phi_1(u, t)$ and $\Phi_2(u, t)$, and we collect terms of same degree in u . Hence, system (25) can be rewritten

$$\left. \begin{aligned} \Phi_1(u, t) &= \sum_{k=0}^4 \chi_{4-k}(t) u^k \\ \Phi_2(u, t) &= \sum_{k=0}^3 \xi_{3-k}(t) u^k \end{aligned} \right\} \quad (26)$$

The polynomial system (26) admits a solution if and only if the resultant, i.e. determinant of a Sylvester matrix, is null [32]. Thus, we obtain the determinant to solve

$$\begin{vmatrix} \xi_0(t) & 0 & 0 & 0 & \chi_0(t) & 0 & 0 \\ \xi_1(t) & \xi_0(t) & 0 & 0 & \chi_1(t) & \chi_0(t) & 0 \\ \xi_2(t) & \xi_1(t) & \xi_0(t) & 0 & \chi_2(t) & \chi_1(t) & \chi_0(t) \\ \xi_3(t) & \xi_2(t) & \xi_1(t) & \xi_0(t) & \chi_3(t) & \chi_2(t) & \chi_1(t) \\ 0 & \xi_3(t) & \xi_2(t) & \xi_1(t) & \chi_4(t) & \chi_3(t) & \chi_2(t) \\ 0 & 0 & \xi_3(t) & \xi_2(t) & 0 & \chi_4(t) & \chi_3(t) \\ 0 & 0 & 0 & \xi_3(t) & 0 & 0 & \chi_4(t) \end{vmatrix} = 0. \quad (27)$$

It turns out that this polynomial generally admits no more than 16 real roots. When all real roots are found, we have to plug them back in (22) for selecting the one leading to the *absolute* maximum of (21). It is not that hard to compute since powerful algorithms exist that can rapidly find roots of univariate polynomials. The complete algorithm can be summarized as follows.

compute $\Gamma_{eg, fh}^w(\tau + \rho)$

set $M := N(N-1)/2$ and $T := \lceil \sqrt{N\lambda_h} \rceil + 1$

set $k := 0..T$ and $p := 0..\lambda_h$

initialize $\mathbf{Q}_p^{(k)}$ to identity

(i.e all (θ, ϕ) to zero) for all k and p

for each k **do**

for each p **do**

$\mathbf{A}_p^{(k)}[z] = \mathbf{Q}_0^{(k)} \mathbf{Z}[z] \dots \mathbf{Q}_{p-1}^{(k)} \mathbf{Z}[z]$

$\mathbf{B}_p^{(k)}[z] = \mathbf{Z}[z] \mathbf{Q}_{p+1}^{(k-1)} \dots \mathbf{Z}[z] \mathbf{Q}_{\lambda_h}^{(k-1)}$

compute $\Gamma_{ac, bd}^{\mathbf{x}^{(k)}}(\tau)$ (with $x^{(k)}[z] = \mathbf{B}_p^{(k)}[z] \mathbf{w}[z]$)

search for $\mathbf{Q}_p^{(k)}$ maximizing $\Upsilon_{1,4}$

(one obtains M pairs of angles (θ, ϕ) yielding $\mathbf{Q}_p^{(k)}$)

return $\mathbf{H}[z] = \mathbf{Q}_0^{(T)} \mathbf{Z}[z] \dots \mathbf{Q}_p^{(T)} \mathbf{Z}[z] \mathbf{Q}_{p+1}^{(T)} \dots \mathbf{Z}[z] \mathbf{Q}_{\lambda_h}^{(T)}$

5.2. Sweeping the pairs

Let us comment some points of this algorithm:

- the resulting tensor $\Gamma_{ac, bd}^{\mathbf{x}}(\tau)$ is composed of $N^2(L - \ell_B + 1)^2$ matrices each of size $N(L - \ell_B + 1) \times N(L - \ell_B + 1)$. Hence, the algorithm needs to store $N^4(L - \ell_B + 1)^4$ entries in memory. Simplifications are generally impossible since we need exact values for computing roots of polynomial system (26),
- 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 still null (set at stage **initialize**). Hence, when loop on each k is repeated several times, angles are better estimated. This procedure reminds the well-known *Jacobi sweeping* widely used for eigenvalue computation.

Remark 10 In the case of a slowly time varying linear channel, the initialization step **initialize** can be modified since angles previously found give a better approximation than null values.

6. Computer Results

For large data blocks, the performance of the algorithm described in Section 5 comes close to that of the MMSE equalizer, and the larger the block length, the larger the accuracy.

Algorithm has been tested on mixtures of length $K = 3$. We have generated 100 random channels. The equalizer has been determined for each channel with data blocks of 256 symbols or 512 symbols. Then, the equalizer has been tested with two different realizations of 5000 symbols each in order to compute the Symbol Error Rate (SER). Thus, the

minimal resolution is $(2 * 5000 * 100)^{-1} = 10^{-6}$. Figure 3 shows median results of the 100 trials for QPSK signals, *i.e.* median is an estimate of average SER. This figure proves that algorithm works well on short data length since from only 256 symbols and with a noise of 16dB, the median SER is below the minimal resolution.

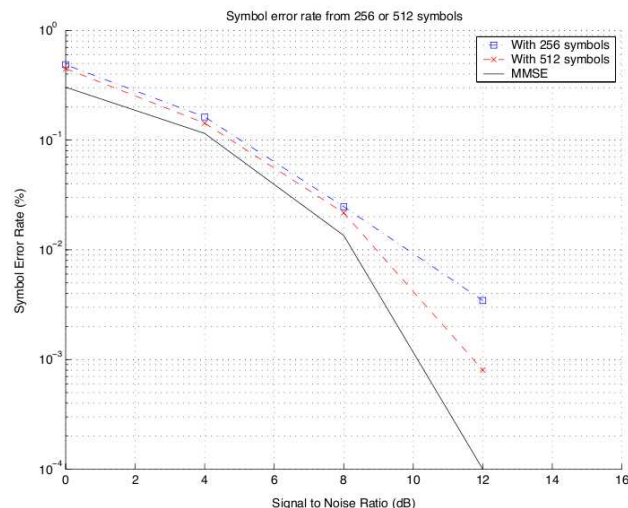


Fig. 3. Symbol Error Rate when a length-3 equalizer is built from blocks of 256 or 512 symbols. For both lengths, the median of the SER is below 10^{-6} after 12dB.

For each trial, the MMSE solution is computed. Another basis for comparison is the distance (Frobenius norm) between the equalizer found and the Zero-Forcing (ZF) solution. In fact, the ZF solution for MIMO convolutive mixtures is simply the inverse of the channel. Then, when $K = L$, the global transfer function with the ZF equalizer $\mathbf{H}_{ZF}[z]$ is

$$\mathbf{G}[z] = \sum_n \sum_m \mathbf{C}(m) \mathbf{H}_{ZF}(n) z^{-(n-m)} = \mathbf{I} z^{-L}. \quad (28)$$

So, we just have to compare the global transfer function $\mathbf{G}[z]$ obtained with the paraunitary equalizer, to the identity matrix up to a multiplicative matrix $\mathbf{D}[z]$ of the form $\mathbf{\Lambda}[z] \mathbf{P}$, where $\mathbf{\Lambda}[z]$ is diagonal and \mathbf{P} a permutation. In order to do this, we proceed by determining the best matrix $\mathbf{D}[z]$ as follows. Once $\mathbf{G}[z]$ is estimated, we store in an $N \times N(2L - 1)$ array \mathbb{G} by merely stacking the matrices one after the other. Next, we search for columns containing the entry of largest modulus in each row of \mathbb{G} . By comparing column indices, which have to be different modulo N , we find matrix \mathbf{P} and $\mathbf{\Lambda}[z]$. Next, we normalize each row by its entry of largest modulus. The matrix obtained this way is denoted $\bar{\mathbb{G}}$. Then,

we replace the N previous entries of $\bar{\mathbb{G}}$ by zeros. This is the same as subtracting \mathbb{D} to $\bar{\mathbb{G}}$, where \mathbb{D} is made from $\mathbf{D}[z]$ with the same size as \mathbb{G} . Finally, we compute the Frobenius norm of the resulting matrix, *i.e.* $\|\bar{\mathbb{G}} - \mathbb{D}\|$. This value is the distance between the paraunitary equalizer found and the ZF solution. Medians of distances are reported in figure 4. We note that between 0 and 8dB and for 512 symbols, the distance decreases rapidly from 1 to 0.25. Then, after 8dB, the distance remains approximately constant.

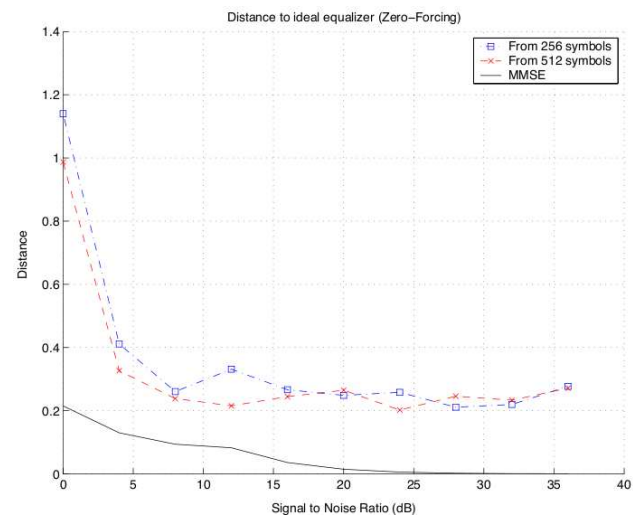


Fig. 4. Distance to the Zero-Forcing equalizer when a length-3 equalizer is built from blocks of 256 or 512 symbols. Solid line: distance between ZF and MSE.

7. Concluding remarks

The Blind MIMO Equalization problem can be reduced to a simpler one via a prior space-time second-order whitening, namely the equalization of a paraunitary channel. Thanks to this statement, we have proposed a parameterization of paraunitary equalizers, in order to carry out blind source separation, *i.e.* deconvolution without a learning sequence. Based on this parameterization, a numerical algorithm has been devised, which iteratively maximizes a contrast through successive sweeps of plane rotations, and eventually equalizes the channel. Moreover, we have demonstrated that the computation of the blind equalizer can be completed within a polynomial complexity, useful in burst-mode transmissions. Performances presented in the last section report an average SER of 0.35% at 12dB of SNR, with data length of 256 symbols only. Other issues currently being addressed include robustness to chan-

nel length misadjustment, which are of prime importance as pointed out in [33,34], and space-time whitening implementations.

Appendix A. Complexity savings

In order to reduce the complexity of the algorithm, it is suggested to store each \mathcal{K} in a matrix $\mathbf{T} \in \mathbb{C}^{N^4 \times N^4}$. This matrix is defined by

$$\mathbf{T} \stackrel{\text{def}}{=} \sum_{\tau, i} \mathbf{f}_i(\boldsymbol{\tau}) \mathbf{c}(\boldsymbol{\tau})^\top. \quad (\text{A.1})$$

where, for $N = 2$,

$$\mathbf{f}_i(\boldsymbol{\tau}) = \begin{bmatrix} f_i^{1111}(\tau_1, \tau_2, \tau_3, \tau_4) \\ f_i^{1112}(\tau_1, \tau_2, \tau_3, \tau_4) \\ f_i^{1121}(\tau_1, \tau_2, \tau_3, \tau_4) \\ \vdots \\ f_i^{2222}(\tau_1, \tau_2, \tau_3, \tau_4) \end{bmatrix} \quad (\text{A.2})$$

with $f_i^{qrst}(\tau_1, \tau_2, \tau_3, \tau_4) = A_{iq}(\tau_1)A_{ir}^*(\tau_2)A_{is}(\tau_3)A_{it}^*(\tau_4)$, and

$$\mathbf{c}(\boldsymbol{\tau}) = \begin{bmatrix} \Gamma_{11,11}^{\mathbf{x}}(\tau_1, \tau_2, \tau_3, \tau_4) \\ \Gamma_{11,12}^{\mathbf{x}}(\tau_1, \tau_2, \tau_3, \tau_4) \\ \Gamma_{12,11}^{\mathbf{x}}(\tau_1, \tau_2, \tau_3, \tau_4) \\ \vdots \\ \Gamma_{22,22}^{\mathbf{x}}(\tau_1, \tau_2, \tau_3, \tau_4) \end{bmatrix}. \quad (\text{A.3})$$

For only 2 observations, we have $\mathbf{T} \in \mathbb{C}^{16 \times 16}$. Hence we obtain the $N^8 = 256$ possible values of (20) by summing L^4 matrices of size 16×16 . Then, we just have to define each \mathcal{K} by summing entries of \mathbf{T} . Moreover, we see that elements of $\mathbf{f}_i(\boldsymbol{\tau})$ and $\mathbf{c}(\boldsymbol{\tau})$ have the same ordering, *i.e.* $q \equiv a, r \equiv b, s \equiv c, t \equiv d$. Hence, both vectors can be constructed thanks to the same loops. So, we reduce the number of loops by N^4 , and hence the computation time.

We immediately see another simplification if we include indices $\tau_1, \tau_2, \tau_3, \tau_4$ in vectors \mathbf{f}_i and \mathbf{c} , *e.g.* by Kronecker product. In this case, both vectors are of length $(N(L+1))^4$ and then \mathbf{T} is $(N(L+1))^4 \times (N(L+1))^4$. So, we only have $(N(L+1))^4$ loops for building vectors \mathbf{f}_i and \mathbf{c} , and then we have to make only one product $\mathbf{f}_i \mathbf{c}^\top$ for matrix \mathbf{T} . Nevertheless, the memory capacity is then the drawback of this

method since more memory is used for storing matrix \mathbf{T} . As a consequence, this last simplification is probably interesting for small values of L .

References

- [1] D. DONOHO, "On minimum entropy deconvolution," in *Applied time series analysis II*, A. Press, Ed., 1981, pp. 565–609.
- [2] P. COMON, "Contrasts for Multichannel Blind Deconvolution," *IEEE Signal Processing Letters*, vol. 3, no. 7, pp. 209–211, July 1996.
- [3] J. K. TUGNAIT, "Identification and Deconvolution of Multichannel NonGaussian Processes using Higher Order Statistics and Inverse Filter Criteria," *IEEE Trans. on Signal Processing*, vol. 45, pp. 658–672, Mar. 1997.
- [4] E. MOREAU and J. C. PESQUET, "Generalized contrasts for multichannel blind deconvolution of linear systems," *IEEE Signal Processing Letters*, vol. 4, no. 6, pp. 182–183, June 1997.
- [5] J. K. TUGNAIT, "On Blind Separation of Convolutional Mixtures of Independent Linear Signals in Unknown Additive Noise," *IEEE Trans. on Signal Processing*, vol. 46, no. 11, pp. 3117–3123, 1998.
- [6] Y. INOUE and R. LIU, "Direct Blind Deconvolution of Multiuser-Multichannel System Driven by Temporally White Source Signals," in *Proc. of ICASSP*, vol. 5, Phoenix, Arizona, USA, 1999, pp. 2889–2892.
- [7] Z. DING and T. NGUYEN, "Stationary Points of a Kurtosis Maximization Algorithm for Blind Signal Separation and Antenna Beamforming," *IEEE Trans. on Signal Processing*, vol. 48, no. 6, pp. 1587–1596, June 2000.
- [8] N. THIRION-MOREAU and E. MOREAU, "Generalized criteria for blind multivariate signal equalization," *IEEE Sig. Proc. Letters*, vol. 9, no. 2, pp. 72–74, Feb. 2002.
- [9] I. FIJALKOW, A. TOUZNI, and J. R. TREICHLER, "Fractionally spaced equalization using CMA: Robustness to channel noise and lack of disparity," *IEEE Trans. on Signal Processing*, vol. 45, no. 1, pp. 56–66, Jan. 1997.
- [10] A. J. V. der VEEN and A. PAULRAJ, "An Analytical Constant Modulus Algorithm," *IEEE Trans. on Signal Processing*, vol. 44, no. 5, pp. 1136–1155, May 1996.
- [11] C. B. PAPADIAS, "Globally Convergent Blind Source Separation Based on a Multiuser Kurtosis Maximization Criterion," *IEEE Trans. on Signal Processing*, vol. 48, no. 12, pp. 3508–3519, December 2000.
- [12] A. TOUZNI, I. FILJAKOW, M. G. LARIMORE, and J. R. TREICHLER, "A Globally Convergent Approach for Blind MIMO Adaptive Deconvolution," *IEEE Trans. on Signal Processing*, vol. 49, no. 6, pp. 1166–1178, June 2001.
- [13] O. GRELLIER, P. COMON, B. MOURRAIN, and P. TREBUCHET, "Analytical Blind Channel Identification," *IEEE Trans. on Signal Processing*, vol. 50, no. 9, Sept. 2002.

- [14] D. YELLIN and B. PORAT, "Blind identification of FIR systems excited by discrete-alphabet inputs," *IEEE Trans. Signal Processing*, vol. 41, no. 3, pp. 1331–1339, 1993.
- [15] S. TALWAR, M. VIBERG, and A. PAULRAJ, "Blind estimation of multiple co-channel digital signals arriving at an antenna array: Part I, algorithms," *IEEE Trans. on Signal Processing*, vol. 44, no. 5, pp. 1184–1197, May 1996.
- [16] P. COMON, "Contrasts, independent component analysis, and blind deconvolution," *Int. Journal Adapt. Control Sig. Proc.*, vol. 18, no. 3, pp. 225–243, Apr. 2004, special issue.
- [17] N. D. SIDIROPOULOS, G. B. GIANNAKIS, and R. BRO, "Blind PARAFAC receivers for DS-CDMA systems," *IEEE Trans. on Signal Processing*, vol. 48, no. 3, pp. 810–823, Mar. 2000.
- [18] D. NION and L. D. LATHAUWER, "A block factor analysis based receiver for blind multi-user access in wireless communications," in *Proc. of ICASSP*, Toulouse, France, May 14-19 2006, pp. 825–828.
- [19] L. ROTA, P. COMON, and S. ICART, "Blind Equalization of MIMO Channels," in *Proc. of Signal Processing Advances in Wireless Communications*, Rome, Italy, June 15-18 2003.
- [20] —, "Blind MIMO Equalization under paraunitary constraint," Laboratoire I3S, I3S Report RR-2002-35-FR, Sept. 2002.
- [21] P. P. VAIDYANATHAN, *Multirate Systems and Filter Banks*. London: Prentice-Hall, 1993.
- [22] Y. INOUE and R. LIU, "A System-Theoretic Foundation for Blind Equalization of an FIR MIMO Channel System," *IEEE Trans. on Circuits and Systems*, vol. 49, no. 4, pp. 425–436, April 2002.
- [23] A. H. SAYED and T. KAILATH, "A Survey of Spectral Factorization Methods," *Numer. Linear Algebra Appl.*, vol. 8, pp. 467–496, 2001.
- [24] L. D. LATHAUWER, B. D. MOOR, and J. VANDEWALLE, "An Algebraic Approach to Blind MIMO Identification," in *Proc. 2nd Int. Workshop on Independent Component Analysis and Blind Signal Separation (ICA 2000)*, Helsinki, Finland, June 19-22 2000, pp. 211–214.
- [25] —, "An Algebraic Approach to the Blind Identification of Paraunitary Filters," in *Proc. of IEEE Wireless Communications and Networking Conference*, Chicago, USA, September 23-28 2000.
- [26] E. MOREAU, "A block algorithm for blind signal deconvolution," in *Proc. of SPAWC'97*, Paris, France, Apr. 1997, pp. 93–96.
- [27] P. COMON, E. MOREAU, and L. ROTA, "Blind Separation of Convulsive Mixtures: a Contrast-Based Joint Diagonalization Approach," in *Proc. of Int. Conf. Independent Component Analysis*, San Diego, Dec. 9-13 2001, pp. 686–691.
- [28] P. McCULLAGH, *Tensor Methods in Statistics*, ser. Monographs on Statistics and Applied Probability. Chapman and Hall, 1987.
- [29] X. GAO, T. Q. NGUYEN, and G. STRANG, "On Factorization of M -Channel Paraunitary Filterbanks," *IEEE Trans. on Signal Processing*, vol. 49, no. 7, pp. 1433–1446, July 2001.
- [30] G. H. GOLUB and C. F. V. LOAN, *Matrix Computation*, 3rd ed. The John Hopkins University Press, 1996.
- [31] N. DELFOSSE and P. LOUBATON, "Adaptive blind separation of independent Sources: a deflation approach," *Signal Processing*, vol. 45, pp. 59–83, 1995.
- [32] D. COX, J. LITTLE, and D. O'SHEA, *Using Algebraic Geometry*. Springer Verlag, New York, 1998.
- [33] Y. Y. T. ACAR and A. PETROPULU, "Blind MIMO system estimation based on PARAFAC decomposition of tensors formed based on the system output," *IEEE Trans. Signal Processing*, vol. 54, no. 11, pp. 4156–4168, 2006.
- [34] B. CHEN and A. P. PETROPULU, "Frequency domain blind MIMO system identification based on second- and higher-order statistics," *IEEE Trans. Signal Processing*, vol. 49, no. 8, pp. 1677–1688, 2001.