Blind Separation of Independent Sources from Convolutive Mixtures

Pierre Comon, Ludwig Rota

Blind Separation of Independent Sources from Convolutive Mixtures

P. Comon, and L. Rota

***Abstract*—The problem of separating blindly independent sources from a convolutive mixture cannot be addressed in its widest generality without resorting to statistics of order higher than two. The core of the problem is in fact to identify the paraunitary part of the mixture, which is addressed in this paper. With this goal, a family of statistical contrast is first defined. Then it is shown that the problem reduces to a Partial Approximate Joint Diagonalization (PAJOD) of several cumulant matrices. Then, a numerical algorithm is devised, which works block-wise, and sweeps all the output pairs. Computer simulations show the good behavior of the algorithm in terms of Symbol Error Rates, even on very short data blocks.**

keywords Blind Source Separation, Blind Equalization, Statistical Contrasts

I. Introduction

When channel inputs are not observed, equalization or identification is referred to as blind, as opposed to pilot-aided techniques. They do not need the eye to be open, hence their name. Pilot sequences are difficult to fully exploit when channels are fast varying, or have long impulse responses. In addition, the present tendency is to reduce the length of pilot sequences, in order to increase the throughput, among others.

Even if Multiple Input Multiple Output (MIMO) equalization can be sometimes carried out with the help of second-order statistics, in particular cases exploiting either the source color [1] or their discrete character [12], it generally requires the use of High-Order Statistics (HOS) [20] [14], at least in a final stage [18] [4]. See [13, ch.5] for further references. It is believed that HOS require "very long convergence times"; this belief is actually often due to an on-line (time recursive) implementation. In fact, on-line blind equalization algorithms require long data blocks to converge (typically from 10,000 to 100,000 symbols); therefore, it is of interest to devise off-line algorithms able to converge much faster (typically 500 symbols), in order to cope with channels with shorter stationarity durations. Therefore, it is exclusively focussed on such *block* algorithms in the present paper.

The case of static mixtures (as opposed to convolutive) has also retained a lot of attention because its simpler form allows a deeper treatment. This special instance will not be addressed here; we refer to [13] for various aspects of that question.

On-line algorithms generally suffer from a number of drawbacks: they cumulate the convergence times of the optimization algorithm, and the estimation of moments. They are also very sensitive to initialization, and may lead to local extrema, and consequently spurious solutions. On the other hand, block algorithms enjoy some advantages. They fully exploit the data (in the sense of a better weighting), are well matched to TDMA transmission formats, and allow the design of analytical algorithms. Their main drawback has been for a long time their excessive computational complexity. But this bottleneck seems to be today much less critical. Let us insist eventually on one striking property of MIMO systems : Finite Impulse Response (FIR) filters can have a FIR inverse [15], which is of course impossible in the SISO case.

The authors are with the I3S Laboratory, CNRS, Sophia-Antipolis, France

The paper is organized as follows. In section II, assumptions and notation are stated. The HOS-based contrast criterion is defined in section III. The block algorithm, aiming at reaching acceptable performances on very short data records (*e.g.*, 500 symbols), is then described in section IV. Extensive computer experiments are eventually concisely reported in section V.

II. System Model

Consider the following linear time-invariant (LTI) invertible system:

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

where $\boldsymbol{s}(n)$ denotes the N-dimensional source vector, $\boldsymbol{x}(n)$ the N-dimensional observation, and $\{\boldsymbol{F}\} = \{\boldsymbol{F}(n), n \in \mathbb{Z}\}$ denotes the $N \times N$ channel impulse response matrix sequence.

For convenience, vectors and matrices are denoted with bold lowercase and bold uppercase letters, respectively. For examples, $\boldsymbol{I}$ denotes the identity matrix. Throughout the paper, $(^{\mathsf{T}})$ stands for transposition, $(^{\mathsf{H}})$ for conjugate transposition, end $(^*)$ for complex conjugation. Also denote by $\mathbb{Z}$ the set of integers, by $\mathbb{N}$ the subset of positive integers, and by $\breve{\boldsymbol{G}}(z)$ the Z-transform of the time sequence $\boldsymbol{G}(n)$: $\breve{\boldsymbol{G}}(z) = \sum_{-\infty}^{\infty} \boldsymbol{G}(k) z^{-k}$.

The MIMO equalization problem consists of finding a filter $\{\boldsymbol{H}\} = \{\boldsymbol{H}(n), n \in \mathbb{Z}\}$ from the sole observation of the channel outputs $\boldsymbol{x}(n)$. Thus the outputs $\boldsymbol{y}(n)$ of the equalizer are estimations of the inputs $\boldsymbol{s}(n)$.

The following hypothesis are assumed :

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

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

$$\mathsf{C}_p^q[s_i] = \mathsf{Cum}[\underbrace{s_i(n), \dots, s_i(n)}_{p \text{ terms}}, \underbrace{s_i^*(n), \dots, s_i^*(n)}_{q=r-p \text{ terms}}] \tag{2}$$

do not depend on n; for definitions of cumulants, refer to [17] and references therein.

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

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

$$\breve{\boldsymbol{G}}(z)\breve{\boldsymbol{G}}^{H}(1/z^*) = \boldsymbol{I} \quad (3)$$

where $\boldsymbol{I}$ denotes the $N \times N$ identity matrix; in other words, $\breve{\boldsymbol{F}}(z)$ and $\breve{\boldsymbol{G}}(z)$ are *paraunitary*, and hence $\breve{\boldsymbol{H}}(z)$.

The interest in using cumulants is that cross cumulants of independent random variables cancel (whereas moments do not necessarily do) [16]. In addition, as pointed out in introduction, HOS are mandatory to restore identifiability.

Remark 1. More generally, if sources are not i.i.d. but are still linear processes, our approach of this problem holds valid. It suffices to assume **H1** in a first stage in order to equalize the channel, and to extract the original sources in a second stage by linear regression between each equalizer output and the observations. In fact the equalizer outputs are the driving processes of the sources.

Remark 2. Hypothesis **H4** is not restrictive. Indeed, one can always whiten the observations, by using a filter that factorizes the second-order power spectrum. This whitening filter is not unique, and one can merely choose it to be minimum phase.

III. Contrasts

The results stated in this section show how contrast-based blind MIMO equalization can be posed in terms of a Partial Joint Approximate Diagonalization (PAJOD) of a set of cumulant matrices. This may be very convenient from the numerical point of view, since we are more familiar with the manipulation of matrices than that of tensors. Proposition 1 defines the contrast optimization criterion, and Proposition 2 proves that the maximization problem deflates into a joint matrix diagonalization. Proposition 3 allows to choose the subset of matrices to be diagonalized. For the sake of clarity, we shall subsequently consider only cumulants of order $r = 4$, but principles hold for orders 3 and higher.

A. Definitions

To start with, denote the following cumulant:

$$\mathsf{C}_2^{2,\boldsymbol{y}}[i,\boldsymbol{j},\boldsymbol{\ell}] = \mathsf{Cum}[y_i(n), y_i(n)^*, y_{j_1}(n-\ell_1), y_{j_2}(n-\ell_2)^*] \quad (4)$$

where $\boldsymbol{j} = (j_1, j_2)$ and $\boldsymbol{\ell} = (\ell_1, \ell_2)$. Also define $\mathsf{J}=\{1,2,\ldots,N\}^2$, and L a subset of $\mathbb{Z}^2$; unless otherwise specified, $\mathsf{L} = \mathbb{Z}^2$. The delays ℓ_1 and ℓ_2 are introduced in the above definition, because they will be necessary to devise an optimization criterion (Proposition 1) that will be reducible to matrix diagonalization (Proposition 2). This reduction, together with the adequate numerical algorithm, constitues the core of our contribution.

Next, define the following terms:

- **Trivial filters.** Clearly, the blind equalization problem we have stated contains inherent indeterminacies. In fact, the set $\mathcal{S}$ of source 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 $\boldsymbol{\Lambda}(z)\cdot\boldsymbol{P}$, 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 $\lambda z^k, k \in \mathbb{Z}$.
- **Contrasts.** 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 [5] :

P1. *Invariance:* The contrast should not change within the set of acceptable solutions, which means that $\Upsilon(\boldsymbol{H};\boldsymbol{x}) = \Upsilon(\boldsymbol{I};\boldsymbol{x}), \forall \boldsymbol{H}\in\mathcal{T}, \forall\boldsymbol{x}\in\mathcal{H}\cdot S$.

P2. *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 $\Upsilon(\boldsymbol{H};\boldsymbol{x}) \le \Upsilon(\boldsymbol{I};\boldsymbol{x})$.

P3. *Discrimination:* The maximum contrast should be reached only for filters linked to each other via trivial filters: $\forall\boldsymbol{x}\in\mathcal{S}, \Upsilon(\boldsymbol{H};\boldsymbol{x}) = \Upsilon(\boldsymbol{I};\boldsymbol{x}) \Rightarrow \boldsymbol{H}\in\mathcal{T}$.

In the remaining, and in accordance with assumptions **H1** through **H4**, $\mathcal{H}$ will denote the set of paraunitary filters, and $\mathcal{S}$ the set of i.i.d. processes with mutually independent components. As a consequence, $\mathcal{H}\cdot S$ is the set of standardized linear processes (*i.e.*, second-order white with unit covariance). Lastly; trivial filters of $\mathcal{T}$ are of the form $\boldsymbol{\Lambda}(z)\cdot\boldsymbol{P}$, where $\boldsymbol{P}$ is a permutation, and $\boldsymbol{\Lambda}(z)$ a diagonal filter, whose entries are of the form λz^k, with $k\in\mathbb{Z}$ and $|\lambda| = 1$.

B. Particular contrast proposed

We are now in a position to prove the proposition below:

Proposition 1 *The functional*

$$\mathcal{J}_2^2(\boldsymbol{H};\boldsymbol{x}) = \sum_{i=1}^{\mathrm{N}}\sum_{\boldsymbol{j}\in\mathsf{J}}\sum_{\boldsymbol{\ell}\in\mathsf{L}} |\mathsf{C}_2^{2,\boldsymbol{y}}[i,\boldsymbol{j},\boldsymbol{\ell}]|^2 \quad (5)$$

is a contrast when observations $\boldsymbol{x}(n)$, and hence the outputs $\boldsymbol{y}(n)$ of the paraunitary equalizer, are standardized.

Let us first comment this criterion. As in the static case [2] [10], the idea is to contract the fourth order cumulant tensor on two indices in order to get a set of matrices to be diagonalized. Because the mixture is convolutive, the contraction should also apply on all delays associated with the contracted indices. Other contrasts of the same family can be defined [7], but will not be discussed here. Let us now turn to the proof.

Proof. Let us then prove proposition 1. The input-output relations of the global system is

$$y_i(n) = \sum_{q,m} G_{iq}(m)s_q(n-m). \quad (6)$$

Thus, using the multilinearity of cumulants and the definition of $\mathcal{J}_2^2$, we get:

$$\begin{aligned}\mathcal{J}_2^2 = & \sum_i \sum_{j_1,j_2} \sum_{\ell_1,\ell_2} \Big| \sum_{q,m} \sum_{q',m'} \sum_{k_1,p_1} \sum_{k_2,p_2} G_{iq}(m) \\ & G^*_{iq'}(m') G_{j_1k_1}(p_1) G^*_{j_2k_2}(p_2) \cdot \\ & \mathsf{C}um[s_q(n-m), s^*_{q'}(n-m'), \\ & s_{k_1}(n-\ell_1-p_1), s^*_{k_2}(n-\ell_2-p_2)] \Big|^2 \end{aligned} \quad (7)$$

Since $s_i(n)$ are i.i.d. (assumption **H1**), the only non-zero cumulants are obtained for $m = m' = \ell_1 + p_1 = \ell_2 + p_2$. Next, since $s_i(n)$ are mutually independent, non-zero terms also need that $q = q' = k_1 = k_2$. Deleting the null terms, and expanding the squared modulus yields:

$$\begin{aligned}\mathcal{J}_2^2 = & \sum_i \sum_{j_1,j_2} \sum_{\ell_1,\ell_2} \sum_{q,m} \sum_{q',m'} G^2_{iq}(m) G^{2*}_{iq'}(m') \\ & G_{j_1q}(m-\ell_1) G^*_{j_1q'}(m'-\ell_1) G_{j_2q}(m-\ell_2) \\ & G^*_{j_2q'}(m'-\ell_2) \mathsf{C}_2^2[s_q] \mathsf{C}_2^{2*}[s_{q'}] \end{aligned} \quad (8)$$

Yet, since $\boldsymbol{G} \in \mathcal{H} \cdot S$ is standardized, it satisfies (3), and in particular, its columns are orthogonal and of unit modulus [5], which means:

$$\sum_{j,\ell} G_{jq}(k-\ell) G^*_{jq'}(k'-\ell) = \delta_{qq'} \delta_{kk'} \quad (9)$$

Applying this property to the pairs of indices (j_1, ℓ_1) and (j_2, ℓ_2), we get:

$$\mathcal{J}_2^2 = \sum_i \sum_{q,m} \left| G^2_{iq}(m) \right|^2 \left| \mathsf{C}_2^{2*}[s_q] \right|^2 \quad (10)$$

Last from (9), we have in particular [5] [19]: $\sum_{k,i} |G_{ij}(k)|^4 \leq 1$ which eventually yields $\mathcal{J}_2^2 \leq \sum_i \left| \mathsf{C}_2^2[s_i] \right|^2$ which proves that $\mathcal{J}_2^2(\boldsymbol{H}; \boldsymbol{x}) \leq \mathcal{J}_2^2(\boldsymbol{I}; \boldsymbol{x})$, for any $\boldsymbol{G} \in \mathcal{H}$ and $\boldsymbol{s} \in \mathcal{S}$. Equality holds if and only if $\sum_{k,i} |G_{ij}(k)|^4 = 1$, which is possible only for trivial filters. □

Now denote the cumulant tensor of observations :

$$\begin{aligned} T_{\boldsymbol{a},\boldsymbol{b}}(\boldsymbol{\alpha}, \boldsymbol{\beta}) = & \mathsf{Cum}[x_{a_1}(n-\alpha_1), \\ & x^*_{a_2}(n-\alpha_2), x_{b_1}(n-\beta_1), x^*_{b_2}(n-\beta_2)] \end{aligned} \quad (11)$$

where $\boldsymbol{a}, \boldsymbol{\alpha}, \boldsymbol{b}, \boldsymbol{\beta}$ are vectors of size 2. The entries of $\boldsymbol{a}$ and $\boldsymbol{b}$ belong to $1, \ldots, N$, by construction.

Consider a FIR equalizer $\{\boldsymbol{H}(n), 0 \leq n \leq L-1\}$, and store the whole impulse response in the block matrix below:

$$\mathbb{H} = [\boldsymbol{H}(0), \boldsymbol{H}(1), \ldots \boldsymbol{H}(L-1)]. \quad (12)$$

The range of variation of $\boldsymbol{\beta}$ is left unspecified for the moment, whereas that of $\boldsymbol{\alpha}$ is set to $\{0, 1, \ldots, L-1\}^2$. The reasons for this choice will become clear in the proof of proposition 2.

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

$$M_{\eta\mu}(\boldsymbol{b}, \boldsymbol{\beta}) = T_{\boldsymbol{a},\boldsymbol{b}}(\boldsymbol{\alpha}, \boldsymbol{\beta}), \quad (13)$$

with $\eta = \alpha_1 N + a_1$, $\mu = \alpha_2 N + a_2$. In short, we shall denote this matrix storage by $\boldsymbol{M}(\boldsymbol{b}, \boldsymbol{\beta})$ in the sequel. Define $\|Diag\{\boldsymbol{A}\}\|^2 = \sum_i |A_{ii}|^2$, we have the following [9]:

Proposition 2. *The contrast $\mathcal{J}_2^2$ can be rewritten as PAJOD criterion of a set of $NL \times NL$ matrices:*

$$\mathcal{J}_2^2(\boldsymbol{H}; \boldsymbol{x}) = \sum_{\boldsymbol{b}} \sum_{\boldsymbol{\gamma}} \left\| Diag\{\mathbb{H} \boldsymbol{M}(\boldsymbol{b}, \boldsymbol{\gamma}) \mathbb{H}^{\mathsf{H}}\} \right\|^2 \quad (14)$$

with

$$\begin{aligned} M_{\eta,\mu}(\boldsymbol{b}, \boldsymbol{\gamma}) = T_{\boldsymbol{a},\boldsymbol{b}}(\boldsymbol{\alpha}, \boldsymbol{\gamma}) = & \mathsf{Cum}[x_{a_1}(n-\alpha_1), \\ & x^*_{a_2}(n-\alpha_2), x_{b_1}(n-\gamma_1), x^*_{b_2}(n-\gamma_2)] \end{aligned} \quad (15)$$

where $\mathbb{H}$ is $N \times NL$ semi-unitary, i.e, satisfies $\mathbb{H}\mathbb{H}^{\mathsf{H}} = \boldsymbol{I}$, and $\boldsymbol{M}(\boldsymbol{b}, \boldsymbol{\gamma})$ is defined as in (13). Here, $\boldsymbol{b}$ varies in $\{1, \ldots, N\}^2$, and $\boldsymbol{\gamma}$ in $\mathbb{Z}^2$.

Proof. The relations between equalizer inputs and outputs can be written as:

$$\begin{aligned} \mathsf{C}_2^{2,\boldsymbol{y}}[i, \boldsymbol{j}, \boldsymbol{\ell}] = & \sum_{\boldsymbol{a},\boldsymbol{b}} \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}} H_{ia_1}(\alpha_1) H^*_{ia_2}(\alpha_2) H_{j_1b_1}(\beta_1) \\ & H^*_{j_2b_2}(\beta_2) T_{\boldsymbol{a},\boldsymbol{b}}(\boldsymbol{\alpha}, \boldsymbol{\beta} + \boldsymbol{\ell}). \end{aligned} \quad (16)$$

Yet the paraunitary condition **H4** on $\breve{\boldsymbol{G}}(z)$ yields that $\breve{\boldsymbol{H}}(z)$ is itself paraunitary, which yields the same orthogonality property as (9):

$$\sum_{j\ell} H^*_{jr}(\tau + \ell) H_{jr'}(\tau' + \ell) = \delta_{rr'} \delta_{\tau\tau'}. \quad (17)$$

Thus, taking the square modulus of (16), making the change of variables $\gamma_k = \beta_k + \ell_k$, and eliminating the unuseful indices leads to

$$\begin{aligned} \mathcal{J}_2^2 = & \sum_{i\boldsymbol{a}\boldsymbol{a'}\boldsymbol{\alpha}\boldsymbol{\alpha'}\boldsymbol{b}\boldsymbol{b'}\boldsymbol{\gamma}\boldsymbol{\gamma'}} H_{ia_1}(\alpha_1) H_{ia_2}(\alpha_2) H^*_{ia'_1}(\alpha'_1) \\ & H^*_{ia'_2}(\alpha'_2) T_{\boldsymbol{a},\boldsymbol{b}}(\boldsymbol{\alpha}, \boldsymbol{\gamma}) T^*_{\boldsymbol{a'},\boldsymbol{b'}}(\boldsymbol{\alpha'}, \boldsymbol{\gamma'}) \\ & \cdot \delta(\boldsymbol{b} - \boldsymbol{b'}) \delta(\boldsymbol{\gamma} - \boldsymbol{\gamma'}), \end{aligned} \quad (18)$$

which can be rearranged into

$$\mathcal{J}_2^2(\boldsymbol{H}; \boldsymbol{x}) = \sum_{i\boldsymbol{b}\boldsymbol{\gamma}} |H_{ia_1}(\alpha_1) H_{ia_2}(\alpha_2) T_{\boldsymbol{a},\boldsymbol{b}}(\boldsymbol{\alpha}, \boldsymbol{\gamma})|^2 . \quad (19)$$

Lastly, grouping indices a_j and α_j together in a single index $\boldsymbol{p}_j$, one can remark that the L matrices $\boldsymbol{H}(\alpha)$ can be stored in the $N \times NL$ matrix $\mathbb{H}$, defined in (12), with full compatibility with (13), so that eventually $\mathcal{J}_2^2 = \sum_{i\boldsymbol{b}\boldsymbol{\gamma}} |\sum_{\boldsymbol{p}_1\boldsymbol{p}_2} \mathbb{H}_{i\boldsymbol{p}_1} \mathbb{H}_{i\boldsymbol{p}_2} M_{\boldsymbol{p}_1\boldsymbol{p}_2}(\boldsymbol{b}, \boldsymbol{\gamma})|^2$. Here the paraunitary property of $\boldsymbol{H}(\tau)$ implies that $\mathbb{H}\mathbb{H}^{\mathsf{H}} = \boldsymbol{I}$. □

Remark 3. The para-unitarity of $\breve{\boldsymbol{H}}(z)$ implies that $\mathbb{H}$ is semi-unitary, but the reverse is not true. In other words, only part of the information is exploited.

Remark 4. The criterion $\mathcal{J}_2^2$ differs from that proposed in [2] in several respects: (i) the matrices $\boldsymbol{M}(b, \gamma)$ are built completely differently, because of the convolutive model, (ii) the matrix sought for is not square unitary but rectangular, which involves quite different calculations, as will be subsequently seen.

Proposition 3. *If the equalizer is of finite length L, and the channel of finite length M, then contrasts $\mathcal{J}_2^2$, defined in proposition 2, can be rewritten as PAJOD criteria of a finite set of at most $(2M + L - 2)^2 N^2$ matrices, where $\mathbb{H}$ is semi-unitary, $\boldsymbol{b}$ varies in $\{1, \ldots, N\}^2$, and $\boldsymbol{\gamma}$ in $\{-M+1, \ldots, M+L-2\}^2$.*

Lemma 5 *If channel and equalizer are both of finite length M and L, respectively, then the cumulant tensor $\boldsymbol{T} = \{T_{\boldsymbol{a},\boldsymbol{b}}(\boldsymbol{\alpha}, \boldsymbol{\gamma})\}$, is null whenever an entry γ_k of $\boldsymbol{\gamma}$ falls outside the interval $\{-M+1, \ldots, M+L-2\}$.*

Proof. In fact, proposition 2 still apply. Consider proposition 2 for instance ($q = 2$), and let's prove the lemma. From definition (11) and input-output channel equations $x_i(n) = \sum_{qm} F_{iq}(m) s_q(n-m)$, we get by multi-linearity of cumulants:

$$\begin{aligned} T_{\boldsymbol{a},\boldsymbol{b}}(\boldsymbol{\alpha}, \boldsymbol{\gamma}) = \sum_{i,j=0}^{M-1} \sum_{\boldsymbol{\ell}} \sum_{u,v=0}^{N} \sum_{\boldsymbol{w} \in \mathbf{J}} F_{a_1 u}(i) F^*_{a_2 v}(j) \\ F_{b_1 w_1}(\ell_1) F^*_{b_2 w_2}(\ell_2) \\ \mathsf{Cum}[s_u(t - \alpha_1 - i), s^*_v(t - \alpha_2 - j), \\ s_{w_1}(t - \gamma_1 - \ell_1), s^*_{w_2}(t - \gamma_2 - \ell_2)], \end{aligned} \tag{20}$$

with $\boldsymbol{\ell} \in \{0, \ldots, M-1\}^2$. Yet, from **H1**, $s_u(n)$ are i.i.d. processes, and the expression is null unless $\alpha_1 + i = \alpha_2 + j = \gamma_1 + \ell_1 = \gamma_2 + \ell_2$. Next, from **H1**, $s_u(n)$ are mutually independent, so that the expression is also null unless $u = v = w_1 = w_2$. this yields

$$\begin{aligned} T_{\boldsymbol{a},\boldsymbol{b}}(\boldsymbol{\alpha}, \boldsymbol{\gamma}) = \sum_{i=0}^{M-1} \sum_{u=0}^{N} F_{a_1 u}(i) F^*_{a_2 u}(i + \alpha_1 - \alpha_2) \\ F_{b_1 u}(i + \alpha_1 - \gamma_1) F^*_{b_2 u}(i + \alpha_1 - \gamma_2) \mathsf{C}_2^2[s_u] \end{aligned} \tag{21}$$

since the support of $\boldsymbol{F}(\cdot)$ is $\{0, 1, \ldots, M-1\}$, the above quantity is null outside the intervals $0 \leq i + \alpha_1 - \gamma_k \leq M-1, \forall k, 1 \leq k \leq 2$. The fact that $0 \leq \alpha_1 \leq L-1$ proves eventually the lemma. The proposition 3 then directly follows.

Remark 5. In practice, it is sufficient to vary the entries γ_k in the central third of the set $\{-M+1, \ldots, M+L-2\}$, namely $\{0, 1, \ldots, L-1\}$. This choice may be suboptimal, and could be improved.

IV. Numerical algorithms

The goal of this section is to demonstrate that the computation of the equalizer can be carried out within a limited (polynomial) number of operations. From now on, we shall assume that (i) the channel length M is known, (ii) the equalizer has the same length $L = M$, and (iii) $\mathsf{L} = \{0, 1, \ldots, L-1\}^2$.

The propositions of the previous section teach us that a semi-unitary matrix, $\mathbb{H}$, of size $N \times NL$, must be found, which should diagonalize approximately and jointly the set of $N^2 L^2$ matrices, $\boldsymbol{M}(b_1, b_2, \gamma_1, \gamma_2)$. Each of these matrices is of size $NL \times NL$. The goal is to maximize the sum of the squared moduli of the N first diagonal entries of the $N^2 L^2$ matrices as shown in figure 2.

A. Jacobi sweeping

In order to reach this goal, one looks for a $NL \times NL$ unitary matrix, $\mathbb{V}$, whose leading $N \times NL$ submatrix (the first N rows) will yield matrix $\mathbb{H}$. This unitary matrix can be built by accumulating Givens rotations, as proposed in the Jacobi algorithm [11] :

$$\mathbb{V} = \prod_{1 \leq i < j \leq NL} \Theta[i,j]^{\mathsf{H}}, \tag{22}$$

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

$$\begin{aligned} \Theta_{ii}[i,j] = \Theta_{jj}[i,j] = \cos(\theta[i,j]) \\ \text{and } \Theta_{ji}[i,j] = -\Theta_{ij}[i,j]^* = \sin(\theta[i,j]) e^{\jmath \psi[i,j]} \end{aligned}$$

with $\jmath = \sqrt{-1}$. This rotation can indeed always be imposed to have a real cosine [11]. The cosine, c, and the sine, s, must be determined so as to maximize, successively for every pair $[i,j]$:

$$\mathcal{J}_2^2 = \sum_{\boldsymbol{b},\boldsymbol{\beta}} \sum_{k=1}^{N} \left| \sum_{\eta,\mu=1}^{NL} \Theta^*_{\eta k}[i,j] \Theta_{\mu k}[i,j] \, M_{\eta\mu}(\boldsymbol{b}, \boldsymbol{\beta}) \right|^2. \tag{23}$$

Put in other words, a PAJOD of a set of matrices $\boldsymbol{M}(\boldsymbol{b}, \boldsymbol{\beta})$ is performed, which means that the $NL \times NL$ matrix

$$\mathbb{V} \boldsymbol{M}(\boldsymbol{b}, \boldsymbol{\beta}) \mathbb{V}^{\mathsf{H}} \tag{24}$$

has an approximatively diagonal $N \times N$ leading block.

B. Processing every pair

Indices $[i,j]$ do not need to describe all possible pairs from the set $\{1, \ldots, NL\}^2$. In fact, since $k \leq N$ in criterion (23), plane rotations $\Theta[i,j]$ will have no effect if $i > N$ and $j > i$. Therefore, it suffices to consider rotations for which $i \leq N$, since $j > i$ by construction.

As a consequence, two cases must be distinguished, depending on the fact that $j \leq N$ or not. In the two cases, we have to find the roots of polynomials (stationary points of a contrast, a rational function in the unknown). Denote in this section $c = \cos(\theta[i,j])$ and $s = \sin(\theta[i,j]) e^{\jmath \psi[i,j]}$:

$$\Theta[i,j] = \begin{pmatrix} c & -s^* \\ s & c \end{pmatrix} \tag{25}$$

and drop provisionally $(\boldsymbol{b}, \boldsymbol{\beta})$ in $\boldsymbol{M}(\boldsymbol{b}, \boldsymbol{\beta})$ for the sake of convenience.

- **Case** $j \leq N$: One maximizes the sum of the 2 diagonal terms on which one has some action. For $\mathcal{J}_2^2$, this is a classical expression [3]:

$$\mathcal{J}_2^2 = \sum_{\boldsymbol{b},\boldsymbol{\beta}} \left| c^2 M_{ii} + c s^* M_{ji} + c s M_{ij} + s s^* M_{jj} \right|^2$$

$$+\left|ss^*M_{ii} - cs^*M_{ji} - csM_{ij} + c^2M_{jj}\right|^2, \qquad (26)$$

- **Case** $j > N$: here only the first diagonal term should be maximized, so that:

$$\mathcal{J}_2^2 = \sum_{\boldsymbol{b},\boldsymbol{\beta}} \left|c^2M_{ii} + cs^*M_{ji} + csM_{ij} + ss^*M_{jj}\right|^2 \quad (27)$$

with appropriate definitions of matrices $\boldsymbol{M}(\boldsymbol{b},\boldsymbol{\beta})$.

C. Complex framework

One considers in this section complex data, channel, and equalizer. In this framework, stationary points are defined by two polynomial equations in two (real) variables, which makes the solution a little more complicated than real framework, described in [8]. In the first case ($j \leq N$), with the help of a change of variables, this rooting can be converted into the solving of two trinomials of degree 2, as in [3]. This transformation is not possible in the second case ($j > N$), and the rooting of the eighth global degree polynomial is mandatory.

We consider a set of 2×2 sub-matrices, say $\boldsymbol{M}(k)$, and a plane rotation θ, that we decide to parameterize by the tangent of its angle, ρ and its complex phase, ψ:

$$\boldsymbol{M}(k) = \begin{pmatrix} \alpha_k & \beta_k \\ \gamma_k & \delta_k \end{pmatrix} \qquad (28)$$

and

$$\Theta = \frac{1}{\sqrt{1+\rho^2}} \begin{pmatrix} 1 & -\rho e^{-\jmath\psi} \\ \rho e^{\jmath\psi} & 1 \end{pmatrix} \qquad (29)$$

with $\rho = \tan\theta$.

The transformed matrices are expressed as $\Theta^H \boldsymbol{M}(k)\Theta$. Define Φ_1 (resp. Φ_2) as the sum of the squared moduli of the first (resp. second) diagonal entries of all transformed matrices. Then we have:

$$\Phi_1 = \frac{1}{(1+\rho^2)^2}\sum_k \left|\alpha_k + \rho e^{-\jmath\psi}\gamma_k + \rho e^{\jmath\psi}\beta_k + \rho^2\delta_k\right|^2 \qquad (30)$$

$$\Phi_2 = \frac{1}{(1+\rho^2)^2}\sum_k \left|\delta_k - \rho e^{-\jmath\psi}\beta_k - \rho e^{\jmath\psi}\gamma_k + \rho^2\alpha_k\right|^2 \qquad (31)$$

Of course, by construction, $\Phi_2(\rho,\psi) = \Phi_1(\frac{1}{\rho}, -\psi+\pi)$.

Case $j > N$: here, the unknowns ρ and ψ should be found so as to maximize Φ_1. For this purpose, the variable $t = \tan\psi/2$ is introduced. Then, $(1+\rho^2)^2(1+t^2)^2\Phi_1$ is a polynomial in t and ρ. Stationary values in ρ and t exactly cancel both the polynomials below:

$$\left.\begin{array}{l} P(\rho,t) = (1+\rho^2)^3(1+t^2)^2\dfrac{\partial\Phi_1}{\partial\rho} \\[2ex] Q(\rho,t) = (1+\rho^2)^2(1+t^2)^3\dfrac{\partial\Phi_1}{\partial t} \end{array}\right\} \qquad (32)$$

$P(\rho,t)$ contains 22 monomials, whose leading one is ρ^4t^4, whereas $Q(\rho,t)$ contains 13 monomials, whose leading one is ρ^2t^4. We note that the second one is much simpler, and that is of degree 2 in ρ.

Considered as polynomials in ρ, P and Q admit a common solution if and only if their resultant (determinant of a Sylvester matrix) is null, which yields:

$$\begin{vmatrix} Q_4 & 0 & P_2 & 0 & 0 & 0 \\ Q_3 & Q_4 & P_1 & P_2 & 0 & 0 \\ Q_2 & Q_3 & P_0 & P_1 & P_2 & 0 \\ Q_1 & Q_2 & 0 & P_0 & P_1 & P_2 \\ Q_0 & Q_1 & 0 & 0 & P_0 & P_1 \\ 0 & Q_0 & 0 & 0 & 0 & P_0 \end{vmatrix} \qquad (33)$$

where $Q_i(t)$ (resp. $P_i(t)$) denote the coefficients of $\rho^i, 0 \leq i \leq 4$ in $Q(\rho,t)$ (resp. of $\rho^j, 0 \leq j \leq 2$ in $P(\rho,t)$). This determinant is a polynomial in t only, and its roots contain all the roots of system (32). It turns out that this polynomial is of degree 24, and that it generally admits no more than 8 real roots, which is consistent with $B\acute{e}zout$ theorem, stating that maximal number of solutions should be 4^2. Plugging back these real roots in $Q(\rho,t)$ allows to compute two candidates for ρ associated with each candidate for t. The best solution (ρ,t) (*i.e.* leading to the global maximum) is then selected by computing the value of the rational function $\Phi_1(\rho,t)$.

Case $j \leq N$: now, the optimization criterion is $\mathcal{J} = \Phi_1 + \Phi_2$. Because of symmetries, this criterion is much simpler to maximize [3] [6]. In fact, define

$$\begin{pmatrix} a_k & b_k \\ c_k & d_k \end{pmatrix} = \Theta^H\boldsymbol{M}(k)\Theta. \qquad (34)$$

Then, one can first notice that

$$\begin{aligned} \mathcal{J}_2^2 &= \sum_k |a_k|^2 + |d_k|^2 \\ &= \frac{1}{2}\sum_k\{|a_k - d_k|^2 + |a_k + d_k|^2\} \end{aligned} \qquad (35)$$

and next, that $a_k + d_k = \alpha_k + \delta_k$, which is thus constant with respect to Θ. The maximization of $\mathcal{J}_2^2$ is consequently equivalent to that of $\sum_k |a_k - d_k|^2$.
Yet, if $\rho = \tan\theta$, one can check out that

$$\begin{aligned} a_k - d_k = \; & (\alpha_k + \delta_k)\cos\theta \\ & + (\beta_k + \gamma_k)\sin\theta\cos\psi \\ & + \jmath(\beta_k - \gamma_k)\sin\theta\sin\psi \end{aligned} \qquad (36)$$

Then, it is easy to show that $\mathcal{J}_2^2$ can be expressed as a quadratic form:

$$\boldsymbol{w}^T\Re[\boldsymbol{B}\boldsymbol{B}^H]\boldsymbol{w} + constant, \qquad (37)$$

where

$$\boldsymbol{w} = [\cos 2\theta, \sin 2\theta\cos\psi, \sin 2\theta\sin\psi]^T, \qquad (38)$$

and where the k-th column of $\boldsymbol{B}$ is:

$$\boldsymbol{B}_k = [\alpha_k + \delta_k, \beta_k + \gamma_k, \jmath(\beta_k - \gamma_k)]^T. \qquad (39)$$

As a consequence, finding the maxima of $\mathcal{J}_2^2$ amounts to maximizing a real quadratic form in 3 variables.

It has been possible to arrange criterion $\mathcal{J}_2^2$ in a quadratic form because some terms in Φ_1 and Φ_2 have cancelled each other, in particular those involving: $\sin^2\theta, \cos\theta, \sin\theta\sin\psi$, and $\sin\theta\sin\psi$, which are not present in (37).
Space is lacking to give the exact analytical expressions of the solutions $\theta[i,j]$ and $\psi[i,j]$; see [7] for further a details. Once the plane rotation is obtained, it is applied to the set of cumulant matrices as $\Theta^{\mathsf{H}}\boldsymbol{M}(k)\Theta$ for criteria $\mathcal{J}_2^2$.

V. Computer results

One considers a Finite Impulse Response (FIR) complex mixture of length $L = 5$ of $N = 2$ QPSK white processes. Thus, there are $N^2L^2 = 100$ square matrices, each of size $NL = 10$, and the goal is to jointly and approximately diagonalize their 2×2 leading matrix by congruent transform. With this goal, a 10×10 unitary matrix, $\mathbb{V}$ is estimated. Matrix $\mathbb{H}$ corresponds to the first two rows of $\mathbb{V}$. The channel is paraunitary, to preserve second-order whiteness as explained is section 2. According to the general decomposition of paraunitary matrices [21], the channel has been generated as follows:

$$\breve{\boldsymbol{F}}(z) = \boldsymbol{R}(\phi_0) \cdot \prod_{m=1}^{4} \left(\boldsymbol{Z}(z)\boldsymbol{R}(\phi_m)\right) \tag{40}$$

where

$$\boldsymbol{Z}(z) = \begin{pmatrix} 1 & 0 \\ 0 & z^{-1} \end{pmatrix} \tag{41}$$

and

$$\boldsymbol{R}(\phi) = \begin{pmatrix} \cos\phi & -\sin\phi e^{-\jmath\psi} \\ \sin\phi e^{\jmath\psi} & \cos\phi \end{pmatrix} \tag{42}$$

Because of the 10 free parameters above,we have some control on the location of zeros of the 4 length-5 SISO channels. In this section, the 10 angles $\phi_i, \psi_i, 1 \le i \le 3$, are drawn according to a uniform distribution in $[0, 2\pi)$ in order to generate paraunitary random channels. For each randomly generated channel, blocks of noisy observations are generated according to:

$$\boldsymbol{x}(n) = \sum_{k=0}^{2} \boldsymbol{F}(k)\boldsymbol{s}(n-k) + \rho\boldsymbol{w}(n) \tag{43}$$

where $\boldsymbol{w}(n)$ is a white circular complex Gaussian noise with identity covariance, and $s_i(n)$ are unit covariance QPSK white sequences. Parameters ρ is introduced in order to control the Signal to Noise Ratio per bit (SNR), that we may define as follows:

$$SNR_{dB} = \frac{E_b}{N_0} = -20\log_{10}\rho \tag{44}$$

In fact, signal and noise parts are both standardized (*i.e.* second-order space-time white).

When evaluating performances of blind MIMO equalizers, a difficulty to overcome stems from inherent indeterminacies. In fact, equalizer $\breve{\boldsymbol{H}}(z)$, and hence global filter $\breve{\boldsymbol{G}}(z)$, can be estimated only up to a multiplicative matrix of the form $\boldsymbol{D}(z) = \boldsymbol{\Lambda}(z)\boldsymbol{P}$, as defined in section 3.1. Let us store the global impulse response $\boldsymbol{G}(n)$ in a $N \times N(2L-1)$ array $\mathbb{G}$. then finding the best matrix $\boldsymbol{D}(z)$ amounts to searching every row of $\mathbb{G}$ for the entry of largest modulus, under the constraint that their column index are different modulo N. This fixes delay and permutation indeterminacies. The phase delay is easier to fix because the alphabet is known: it suffices to compare to 1 the output raised to the fourth power. In other words, we calculate the error rate of $N!N(L+M-1)$ potential estimators, and chose the best. For $N = 2$ and $L = 5$, we have thus explore 18 different cases (9 possible delays for each row), for each of the 2 permutations.

Results are reported in figure 3 for blocks of 500 an 1000 symbols, as a function of SNR. 45 trials have been run. For every trial, two blocks of data have been independently generated, of length 500 (or 1000) and 5000, respectively. Once the equalizer has been calculated from the the whole block of length 500 (or 1000), it has been tested on the other block of 5000 symbols to compute the SER; this is a *hold-out* type performance testing that avoids over-fitting. This procedure has been repeated 45 times, in order to obtain an average SER; the median of the 45 trials is plotted in figure 3. As a consequence, the minimal resolution is $(45 * 5000)^{-1} = 4.4 \times 10^{-6}$. After a SNR of 13dB, the SER falls below the latter resolution. These curves demonstrate the good behavior of the algorithm for short data blocks. As a basis for comparisons, the performances obtained with the exact inverse channel are also represented; it corresponds in the present case to the Zero-Forcing (ZF) equalizer, optimal in the absence of noise.

VI. Conclusions

The numerical algorithm described in section IV, performing a Partial Joint Diagonalization of cumulant matrices, was based on preliminary theoretical results reported in section III. This algorithm demonstrates that it is possible to equalize blindly FIR MIMO channels from data records as short as 500 symbols, contrary to what is often believed. In addition, the block approach we proposed is attractive in all TDMA transmission modes. Performances of the proposed algorithm remain quite attractive for random channels up to length 5, but could probably be improved by refining the paraunitary constraint. This is the subject of current research.

Acknowledgment

We wish to thank the anonymous reviewers for their careful reading. Their remarks have helped us to improve on the clarity of the paper.

References

[1] A. BELOUCHRANI, K. ABED-MERAIM, J. F. CARDOSO, and E. MOULINES. A blind source separation technique using second order statistics. *IEEE Trans. Signal Processing*, 45(2):434–444, February 1997.
[2] J. F. CARDOSO and A. SOULOUMIAC. Blind beamforming for non-Gaussian signals. *IEE Proceedings Part F*, 140(6):362–370, December 1993.
[3] J. F. CARDOSO and A. SOULOUMIAC. Jacobi angles for simultaneous diagonalization. *SIAM Jour. matrix Analysis*, 17(1):161–164, Jan. 1996.

[4] P. COMON. Analyse en Composantes Indépendantes et identification aveugle. *Traitement du Signal*, 7(3):435–450, December 1990. Numero special non lineaire et non gaussien.

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

[6] P. COMON. From source separation to blind equalization, contrast-based approaches. In *Int. Conf. on Image and Signal Processing (ICISP'01)*, Agadir, Morocco, May 3-5, 2001. invited plenary.

[7] P. COMON. Blind separation of convolutive mixtures, a partial joint diagonalization PAJOD approach. Research report 2002.04, I3S-CNRS, Feb. 2002.

[8] P. COMON, E.MOREAU, and L. ROTA. Blind separation of convolutive mixtures: a contrast-based joint diagonalization approach. In *3rd Int. Conf. Independant Component Analysis*, San Diego, Dec. 9-13 2001.

[9] P. COMON and E. MOREAU. Blind MIMO equalization and jointdiagonalization criteria. In *ICASSP*, Salt Lake City, May 711 2001.

[10] L. DeLATHAUWER. *Signal Processing on Multilinear Algebra*. Doctorate, Katholieke Universiteit Leuven, Sept. 1997.

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

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

[13] S. HAYKIN. *Unsupervised Adaptive Filtering*, volume 1. Wiley, 2000. series in Adaptive and Learning Systems for Communications, Signal Processing, and Control.

[14] Y. INOUYE and T. HABE. Blind equalization of multichannel linear time-invariant systems. *The Institute of Electronics Information and Communication Engineers*, 24(5):9–16, May 1995.

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

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

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

[18] A. SWAMI, G. GIANNAKIS, and S. SHAMSUNDER. Multichannel arma processes. *IEEE Trans. Sig. Proc.*, 42(4):898–913, Apr. 1994.

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

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

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

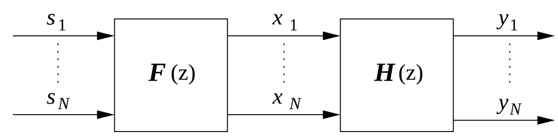

Fig. 1. Observation $\boldsymbol{x}$ is equalized by $\boldsymbol{H}$; the global system is denoted $\boldsymbol{G}$.

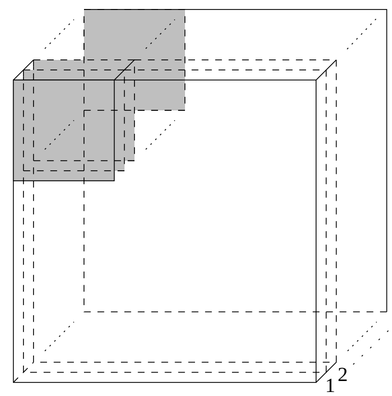

Fig. 2. The semi-unitary matrix $\mathbb{H}$ aims at diagonalizing jointly the $N \times N$ leading sub-matrices (shaded area) of the N^2L^2 matrices. In the above picture, they are stacked one above the other, as slices of a cube.

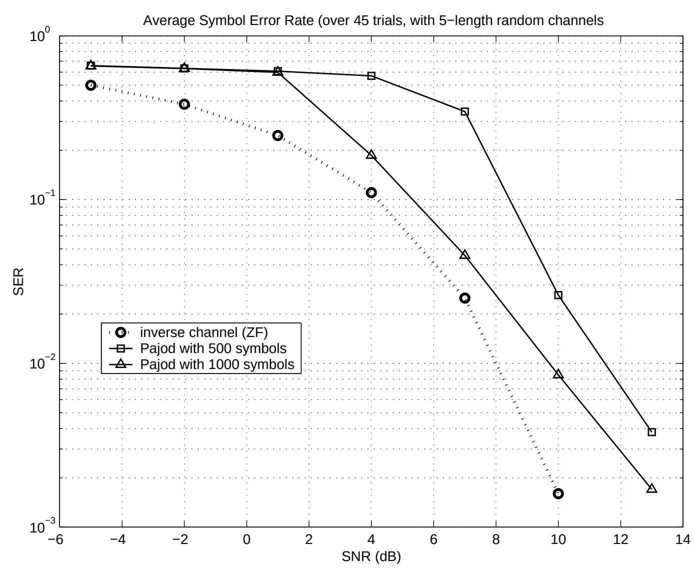

Fig. 3. Performances obtained for data blocks of length 500 and 1000 symbols: Symbol Error Rates (SER) are obtained for random paraunitary channels of length 5, and with a blind equalizer of length 5.