

Parafac-based blind channel identification using 4th-order cumulants

Carlos Estêvão R. Fernandes⁽¹⁾, Gérard Favier⁽¹⁾, João Cesar M. Mota⁽²⁾

Abstract—Strong relationships between joint-diagonalization and tensor decompositions have been established recently. In this paper we propose a finite impulse response (FIR) channel identification method based on the Parafac decomposition of a 3rd-order tensor composed of the 4-th order output cumulants. By avoiding the prewhitening operation required by joint-diagonalization based methods, our method is shown to improve the estimation performance provided by the classic joint-diagonalization algorithm.

Index Terms—Channel identification, Higher-order statistics, Joint-diagonalization, Tensor decomposition, Parafac.

I. INTRODUCTION

In data communications systems, the knowledge of the transmission channel plays a very important role. Blind channel equalization and identification concern the retrieval of unknown information about the transmission channel and source signals from the knowledge of the received signal only. For several years, higher-order statistics (HOS) have been an important interesting topic in diverse fields including data communication, speech and image processing and geophysical data processing. The higher-order spectra have the ability to preserve both magnitude and (nonminimum-) phase information. Moreover, it is well-known that for Gaussian signals, all cumulant spectra of order greater than 2 vanish. As a result, higher-order cumulants automatically eliminate additive Gaussian noise that corrupts non-Gaussian signals.

All over the last three decades, several results have been reported proposing many relationships connecting cumulant slices of different

orders and the parameters of FIR models [1]. A number of other methods were proposed to improve identification performance making use of a larger set of output cumulants and solving an overdetermined system [2]. In this context, the optimal combination of the statistical data arises as a challenging problem. Existing approaches to exploit the sample output cumulants include the joint-diagonalization of cumulant matrices [3]. However, to apply such joint-decomposition techniques, a *prewhitening* transformation over the cumulant matrices is necessary [4]. This operation is required in most of HOS-based blind channel identification methods and it is often a source of increased complexity and estimation errors [5].

Nevertheless, strong relationships between joint-diagonalization and tensor decompositions have been established recently [6]. In blind channel identification, factorization of *multi-way* arrays allows us to avoid the *prewhitening* operation over the cumulant matrices. Moreover, we fully exploit the three-dimensional nature of the 4th-order cumulant tensor, thus improving parametric estimation quality. Recently, some blind identification methods have been introduced based on the multi-way decomposition of a tensor containing output bispectra and making use of Parallel Factor Analysis (Parafac) [7], [8], [9]. However, there are some limitations in that frequency-domain approach: besides being computationally very complex, they use 3rd-order statistics, which are not appropriate to treat digital communication systems where input signals are often symmetrically distributed random variables.

In this paper, we propose a time-domain tensor-based approach making use of Parafac to decompose a three-dimensional tensor of 4th-order cumulants aiming to recover the parameters of an FIR channel. We consider the baseband rep-

⁽¹⁾I3S Lab., CNRS, University of Nice Sophia Antipolis, France. E-mails: cfernand@i3s.unice.fr, favier@i3s.unice.fr.

⁽²⁾Teleinformatics Engineering Dept., Federal University of Ceará, Fortaleza-CE, Brazil. E-mail: mota@deti.ufc.br.

representation of a digital *single-input single-output* (SISO) communication channel where the output signal $y(n)$ after sampling at the symbol rate is written as follows

$$\begin{aligned} y(n) &= x(n) + v(n), \\ x(n) &= \sum_{l=0}^L h_l s(n-l), \quad h_0 = 1. \end{aligned} \quad (1)$$

The following assumptions hold¹:

- A1 : The non-measurable, complex-valued, discrete input sequence $s(n)$ is non-Gaussian, stationary, independent and identically distributed (i.i.d.) with symmetric distribution of zero-mean and variance $\sigma_s^2 = 1$.
- A2 : The additive Gaussian noise sequence $v(n)$ has zero-mean and is independent from $s(n)$. Its autocorrelation function is unknown.
- A3 : The channel frequency-response is $H(\omega) = \sum_l h_l e^{-j\omega l}$ with complex coefficients h_l representing the equivalent discrete impulse response, including the pulse shaping filter, the channel itself and the receiving filters.
- A4 : The system is causal with memory L , i.e. $h_l = 0 \forall l \notin [0, L]$. In addition, $h_L \neq 0$ and $L \neq 0$.

The paper is organized as follows: in section II, we briefly introduce some fundamentals of Parafac decomposition; in section III, we build a 3rd-order tensor using the 4th-order output cumulants and then we decompose it using Parafac; Identifiability and uniqueness issues are briefly addressed and a relationship relating the tensor decomposition with the joint-diagonalization approach is established. in section IV, we propose our novel Parafac-based blind channel identification (PBCI) algorithm. In section V, performance results are illustrated through computer simulations and we finally draw our conclusions in section VI.

II. PARAFAC REVISITED

Let us consider a 3rd-order tensor \mathcal{X} of dimensions $I \times J \times K$ having the following F -component decomposition:

$$x_{i,j,k} = \sum_{f=1}^F a_{if} b_{jf} c_{kf} \quad (2)$$

¹ Imposing $h_0 = 1$ in (1) is not a restrictive assumption since it is equivalent to a simple unit-norm constraint.

where $i \in [1, I]$, $j \in [1, J]$ and $k \in [1, K]$. The sum of F factors expressed in (2) is called the Parafac decomposition of tensor \mathcal{X} and it can be represented as the outer product of three vectors, i.e. three rank-one elements.

Three matrix representations of (2) can be obtained by slicing tensor \mathcal{X} along each of the three possible directions. Taking the horizontal direction, the tensor slices define a set of I matrices $\mathbf{X}_{i..}$, $i \in [1, I]$, with dimensions $J \times K$, which are written as follows

$$\mathbf{X}_{i..} = \mathbf{B} \mathbf{D}_i(\mathbf{A}) \mathbf{C}^T, \quad i \in [1, I], \quad (3)$$

where $\mathbf{A} \in \mathbb{C}^{I \times F}$, $\mathbf{B} \in \mathbb{C}^{J \times F}$ and $\mathbf{C} \in \mathbb{C}^{K \times F}$ are composed entrywise as $[\mathbf{A}]_{i,f} = a_{if}$, $[\mathbf{B}]_{j,f} = b_{jf}$ and $[\mathbf{C}]_{k,f} = c_{kf}$, respectively, and $\mathbf{D}_i(\cdot)$ is a diagonal matrix whose main diagonal contains the elements in the i th row of the matrix argument. Then, stacking the slices $\mathbf{X}_{i..}$, $i \in [1, I]$, we get an unfolded matrix representation of tensor \mathcal{X} , as follows

$$\mathbf{X}_{[1]} = (\mathbf{A} \diamond \mathbf{B}) \mathbf{C}^T \in \mathbb{C}^{IJ \times K}, \quad (4)$$

where \diamond stands for the Khatri-Rao product [10]. Similarly, slicing \mathcal{X} along vertical direction, yields J matrices $\mathbf{X}_{.j.}$, $j \in [1, J]$, with dimensions $K \times I$, which are written as

$$\mathbf{X}_{.j.} = \mathbf{C} \mathbf{D}_j(\mathbf{B}) \mathbf{A}^T, \quad j \in [1, J], \quad (5)$$

and this gives rise to the following unfolded representation

$$\mathbf{X}_{[2]} = (\mathbf{B} \diamond \mathbf{C}) \mathbf{A}^T \in \mathbb{C}^{KJ \times I}. \quad (6)$$

Finally, by frontally slicing \mathcal{X} we get K matrices $\mathbf{X}_{..k}$, $k \in [1, K]$, with dimensions $I \times J$, which are written as

$$\mathbf{X}_{..k} = \mathbf{A} \mathbf{D}_k(\mathbf{C}) \mathbf{B}^T, \quad k \in [1, K], \quad (7)$$

yielding an unfolded representation as follows

$$\mathbf{X}_{[3]} = (\mathbf{C} \diamond \mathbf{A}) \mathbf{B}^T \in \mathbb{C}^{KI \times J}. \quad (8)$$

Any of the unfolded representations $\mathbf{X}_{[1]}$, $\mathbf{X}_{[2]}$ or $\mathbf{X}_{[3]}$ completely characterizes the Parafac decomposition of tensor \mathcal{X} .

The rank of a three-dimensional tensor is defined as the minimum number F of (3-way) factors needed to decompose the tensor in the form of (2). For $F > 1$ and under certain conditions,

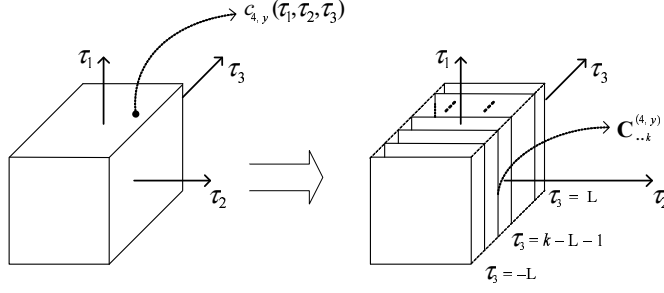


Fig. 1. Slicing tensor $\mathcal{C}^{(4,y)}$ along the frontal direction (mode-3).

the Parafac decomposition is shown to be essentially unique up to scaling factors. The uniqueness condition of Parafac can be stated as follows

$$k_A + k_B + k_C \geq 2(F + 1), \quad (9)$$

where k_A , k_B and k_C are respectively the k -ranks of the matrices \mathbf{A} , \mathbf{B} and \mathbf{C} [11]. The k -rank of a matrix is defined as follows: *An $n \times m$ matrix \mathbf{A} is said to have k -rank equal to k_A if every set containing $k_A \leq m$ columns of \mathbf{A} is linearly independent and, in addition, there is at least one set composed of $k_A + 1$ columns of \mathbf{A} that is linearly dependent.* From this definition, note that $k_A \leq \text{rank}(\mathbf{A}) \leq \min(n, m)$.

Satisfying (9) means that the Parafac decomposition is unique up to scaling and column permutation. This fact represents a great advantage of Parafac over bilinear decompositions, where some rotation is always possible without changing the fit of the model.

III. A TENSOR OF 4TH-ORDER CUMULANTS

Let us denote the output 4th-order cumulant as $c_{4,y}(\tau_1, \tau_2, \tau_3) \triangleq \text{cum}[y^*(n), y(n + \tau_1), y^*(n + \tau_2), y(n + \tau_3)]$, where $*$ stands for complex-conjugate. Using the channel model (1), taking into account assumption A1 and A2, and making use of the multilinearity property of cumulants, we get [12]:

$$c_{4,y}(\tau_1, \tau_2, \tau_3) = \gamma_{4,s} \sum_{l=0}^L h_l^* h_{l+\tau_1} h_{l+\tau_2}^* h_{l+\tau_3}, \quad (10)$$

where $\gamma_{4,s} = C_{4,s}(0, 0, 0)$ is the *kurtosis* of the input signal $s(n)$. Based on (10) and on assumption

A4, it is easy to verify that

$$c_{4,y}(\tau_1, \tau_2, \tau_3) = 0, \quad \forall |\tau_1|, |\tau_2|, |\tau_3| > L. \quad (11)$$

Hence, making the time-lags τ_1 , τ_2 and τ_3 vary in the interval $[-L, L]$ we can be sure to include all the nonzero 4th-order cumulant information. Such a choice allows us to construct a maximal redundant information model, in which the 4th-order cumulants are taken for time-lags τ_1 , τ_2 and τ_3 within the interval $[-L, L]$, defining the element in position (i, j, k) of the 3-way array $\mathcal{C}^{(4,y)}$, where $i = \tau_1 + L + 1$, $j = \tau_2 + L + 1$ and $k = \tau_3 + L + 1$ (fig. 1).

Therefore, $\mathcal{C}^{(4,y)}$ is clearly 3rd-order tensor represented as a cube of dimensions $(2L + 1) \times (2L + 1) \times (2L + 1)$. Slicing this cube along the frontal direction, as shown in fig. 1, yields $2L + 1$ matrices $\mathbf{C}_{..k}^{(4,y)} \in \mathbb{C}^{(2L+1) \times (2L+1)}$, $k \in [1, 2L + 1]$. These frontal slices can be easily shown to be written as

$$\mathbf{C}_{..k}^{(4,y)} = \gamma_{4,s} \mathbf{H} \mathbf{D}_k(\boldsymbol{\Sigma}) \mathbf{H}^H, \quad (12)$$

for all $k \in [1, 2L + 1]$, where $\mathbf{D}_k(\boldsymbol{\Sigma}) \in \mathbb{C}^{(L+1) \times (L+1)}$, with

$$\boldsymbol{\Sigma} = \mathbf{H} \text{Diag}(\mathbf{h}^H) \quad (13)$$

where the operator $\text{Diag}(\cdot)$ builds a diagonal matrix in which the main diagonal contains the elements of the vector argument and we have defined the channel coefficients matrix $\mathbf{H} \in \mathbb{C}^{(2L+1) \times (L+1)}$ as follows

$$\mathbf{H} \triangleq \mathcal{H}(\mathbf{h}) = [\mathbf{h}_0 \ \mathbf{h}_1 \ \dots \ \mathbf{h}_L] \quad (14)$$

with

$$\mathbf{h}_p \triangleq [h_{p-L} \ \dots \ h_p \ \dots \ h_{p+L}]^T, \quad p \in [0, L],$$

where $\mathcal{H}(\cdot)$ is an operator that builds a special Hankel matrix from the vector argument as shown above and the channel coefficients vector is $\mathbf{h} = [h_0 \dots h_L]^\top \in \mathbb{C}^{(L+1)}$.

Equation (12) establishes a direct link between the tensor decomposition presented here and the joint-diagonalization approach [3]. Details of the demonstration that yields (12) are omitted here due to the lack of space. However, since \mathbf{H} is not unitary, diagonalization-based approaches need a *prewhitening* step in order to make its columns orthonormal. This additional non-optimal processing is common in most HOS-based methods and is often responsible for increased estimation errors.

Note moreover that the full-rank (Hankel) structure of the channel matrix \mathbf{H} ensures that $k_H = \text{rank}(\mathbf{H}) = L+1$. The rank of Σ , however, equals the number of non-zero channel parameters in \mathbf{h} . Due to assumption A4, we have $\text{rank}(\Sigma) \geq 2$. Thus, according to (9), the identifiability condition of our Parafac decomposition with $F = L + 1$ components is satisfied. Stacking up matrices $\mathbf{C}_{\cdot,k}^{(4,y)}$ for $k \in [1, 2L+1]$, it is not difficult to get from (12) an unfolded representation of $\mathcal{C}^{(4,y)}$ as follows

$$\mathbf{C}_{[3]} = \gamma_{4,s} (\Sigma \diamond \mathbf{H}) \mathbf{H}^H. \quad (15)$$

From analogy with (8), we can easily deduce that $\mathbf{A} = \mathbf{H}$, $\mathbf{B} = \mathbf{H}^*$ and $\mathbf{C} = \Sigma$. Then, from (4) we get

$$\begin{aligned} \mathbf{C}_{[1]} &= \gamma_{4,s} (\mathbf{H} \diamond \mathbf{H}^*) \Sigma^\top \\ &= \gamma_{4,s} (\mathbf{H} \diamond \mathbf{H}^*) \text{Diag}(\mathbf{h}^*) \mathbf{H}^\top \end{aligned} \quad (16)$$

IV. PARAFAC-BASED BLIND CHANNEL IDENTIFICATION

Based on the knowledge of $\mathbf{C}_{[1]}$, we propose an iterative Least Squares (LS) procedure to estimate the channel coefficients vector \mathbf{h} . Recall that if $\mathbf{X} = \mathbf{A} \text{Diag}(\mathbf{d}) \mathbf{B}$, then the following property of the Khatri-Rao product holds [10]

$$\text{vec}(\mathbf{X}) = (\mathbf{B}^\top \diamond \mathbf{A}) \mathbf{d} \quad (17)$$

Therefore, it is straightforward to rewrite (16) as

$$\text{vec}(\mathbf{C}_{[1]}) = \gamma_{4,s} \mathbf{H} \diamond (\mathbf{H} \diamond \mathbf{H}^*) \mathbf{h}^* \quad (18)$$

where $\text{vec}(\cdot)$ stands for the *vectorizing* operator. At this point, it is crucial to take into account the

Hankel structure of \mathbf{H} to initialize the algorithm. This is the key to ensure the full-rank property that makes our Parafac decomposition essentially unique and free from permutation ambiguities. Also, the constraint $h_0 = 1$ is essential to avoid any kind of scaling ambiguities. For that reason, we initialize the algorithm with a Hankel matrix $\hat{\mathbf{H}}^{(r)}$, $r = 0$, in which the first column is $[\mathbf{0}_{(L)}^\top \hat{\mathbf{h}}^{(0)\top}]^\top$ and the last row is $[\hat{h}_L^{(0)} \mathbf{0}_{(L)}^\top]$, where $\hat{\mathbf{h}}^{(0)} = [1 \ \mathbf{v}^\top]^\top$ and $\mathbf{v} \in \mathbb{C}^{(L)}$ is a Gaussian random vector.

Thus, after that initialization step, we compute the vector $\hat{\mathbf{h}}^*$ by using the least squares (LS) algorithm to minimize the cost function

$$\mathbf{J}(\hat{\mathbf{h}}^*) \triangleq \left\| \text{vec}(\mathbf{C}_{[1]}) - \hat{\mathbf{G}} \hat{\mathbf{h}}^* \right\|_F^2 \quad (19)$$

where

$$\hat{\mathbf{G}} = \gamma_{4,s} \hat{\mathbf{H}} \diamond (\hat{\mathbf{H}} \diamond \hat{\mathbf{H}}^*). \quad (20)$$

The algorithm is iterated until convergence of the parametric estimator, i.e. until $\|\hat{\mathbf{h}}^{(r)} - \hat{\mathbf{h}}^{(r-1)}\| / \|\hat{\mathbf{h}}^{(r)}\| \leq \varepsilon$, where r is the iteration number and ε is an arbitrary small positive constant. For each iteration r , we use the new estimate $\hat{\mathbf{h}}^{(r)}$ to update $\hat{\mathbf{H}}^{(r)}$. The algorithm to estimate \mathbf{h} is as follows:

- 1) For iteration $r = 0$, initialize $\hat{\mathbf{h}}^{(0)}$ as a Gaussian random vector;
- 2) Build the channel matrix $\hat{\mathbf{H}}^{(0)}$ as in (14);
- 3) For $r \geq 1$, estimate $\hat{\mathbf{h}}^{(r)*}$ by minimizing the cost function (19). This yields

$$\hat{\mathbf{h}}^{(r)*} = \hat{\mathbf{G}}_{r-1}^\# \text{vec}(\mathbf{C}_{[1]}), \quad (21)$$

where $(\cdot)^\#$ stands for the pseudoinverse.

- 4) Update $\hat{\mathbf{H}}^{(r)}$ from $\hat{\mathbf{h}}^{(r)}$ using (14);
- 5) Reiterate until convergence of the parametric error, i.e. $\|\hat{\mathbf{h}}^{(r)} - \hat{\mathbf{h}}^{(r-1)}\| / \|\hat{\mathbf{h}}^{(r)}\| < \varepsilon$.

This strategy ensures an improved solution at each iteration [13]. Another advantage is that none permutation or scaling ambiguities remain since we explicitly exploit the Hankel structure of \mathbf{H} as well as the constraint $h_0 = 1$.

V. SIMULATION RESULTS

In this section, we present some computer simulations illustrating the applicability of the proposed *Parafac-based Blind Channel Identification* (PBCI) method. We compare our results with

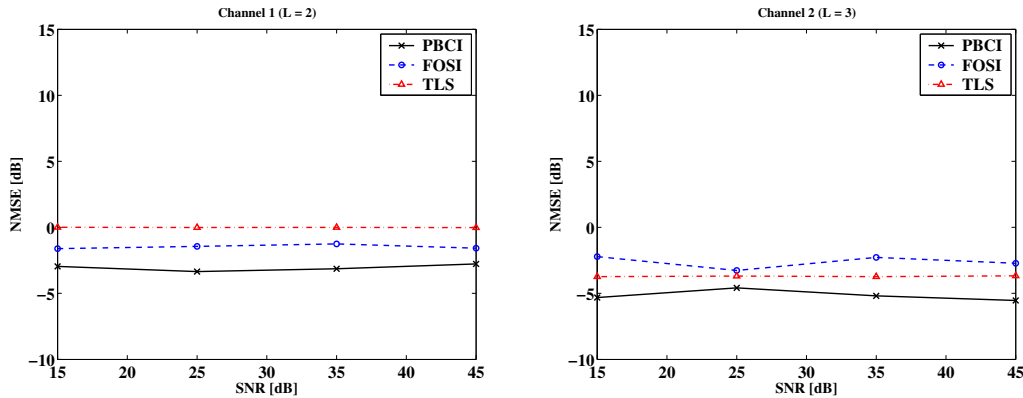


Fig. 2. Identification performances of Parafac (PBCI), joint-diagonalization (FOSI) and TLS methods with BPSK modulation.

those obtained from a joint diagonalization based approach, which will be referred to as the *Fourth Order System Identification* (FOSI) algorithm [3]. As suggested by the authors, the results obtained with the FOSI algorithm are the average of the two solutions proposed in [3] because we do not make any further assumptions allowing us to choose between one of them. In addition, we also compare our results with the total least squares (TLS) solution proposed in [2].

Parametric channel estimation performance is evaluated by means of the normalized mean squared error (NMSE) of the estimator, which is computed through the following formula:

$$\text{NMSE} = \frac{1}{M} \sum_{m=1}^M \frac{\|\hat{\mathbf{h}}^{(m)} - \mathbf{h}\|^2}{\|\mathbf{h}\|^2}, \quad (22)$$

where M is the number of Monte Carlo runs and $\hat{\mathbf{h}}^{(m)}$ is the channel estimate obtained at simulation $m \in [1, M]$. To obtain the following results we have used two particular nonminimum-phase channels, the coefficients of which being respectively given by $\mathbf{h} = [1, -1.4 - 1.21i, 0.49 + 0.85i]^T$ ($L = 2$) and $\mathbf{h} = [1, 0.7 - 1.44i, -0.35 - 0.44i, -0.87 + 0.2i]^T$ ($L = 3$). We estimate 4th-order cumulants from $N = 10000$ output data samples assuming perfect knowledge of the channel memory L . The following curves were averaged from $M = 50$ Monte Carlo runs where input and noise signal samples have been varied for each SNR value.

Figure 2 illustrates the estimation performance

of PBCI against FOSI and TLS solutions for the case of a BPSK modulated input signal. These plots clearly show that our approach over-perform both the joint-diagonalization (FOSI) algorithm and the TLS solution. In fig. 3 we observe the results concerning the case of a 8-PSK modulated input signal. Here again, PBCI presents the best results and, as expected, all techniques have superior performance than in the previous case, due to a simple information diversity principle. Finally, observe that the robustness of the HOS-based algorithms with respect to the additive noise is confirmed by the flatness of the curves.

VI. CONCLUSIONS AND PERSPECTIVES

We have introduced a new blind FIR channel identification method based on the Parafac decomposition of a 3rd-order tensor composed from the output 4th-order cumulants. Our method fully exploits the three-dimensional nature of the cumulant tensor and has the advantage of avoiding any kind of pre-processing. Computer simulations show that the Parafac-based approach provides better estimation performance than the classic joint-diagonalization algorithm. Convergence issues should be further investigated soon. Moreover, as tensor decomposition techniques are suitable for treating multidimensional problems, an extension to the multiple-input multiple-output (MIMO) case is straightforward. Such an extension of PBCI is presently under work.

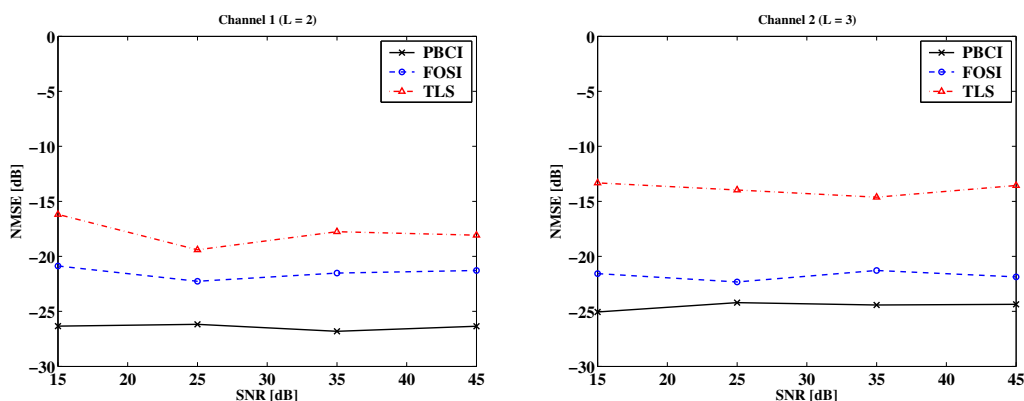


Fig. 3. Identification performances of Parafac (PBCI), joint-diagonalization (FOSI) and TLS methods with 8-PSK modulation.

REFERENCES

- [1] C. L. Nikias and J. M. Mendel, "Signal processing with higher-order spectra," *IEEE Signal Processing Magazine*, pp. 10–37, jul 1993.
- [2] P. Comon, "MA identification using fourth order cumulants," *Signal Processing*, vol. 26, no. 3, pp. 381–388, mar 1992.
- [3] A. Belouchrani and B. Derras, "An efficient fourth-order system identification FOSI algorithm utilizing the joint diagonalization procedure," in *Proc. of the 10-th IEEE Workshop on Statistical Signal and Array Processing*, Pennsylvania, USA, aug 2000, pp. 621–625.
- [4] J.-F. Cardoso and A. Souloumiac, "Blind beamforming for non gaussian signals," *IEE Proceedings-F*, vol. 140, no. 6, pp. 362–370, dec 1993.
- [5] 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, jun 1997.
- [6] J. Castaing and L. De Lathauwer, "Séparation aveugle de signaux DS-CDMA à l'aide de techniques algébriques," in *20ème colloque GRETSI sur le traitement du signal et des images*, Louvain-La-Neuve, Belgium, Sept. 2005, pp. 965–968.
- [7] T. D. Acar and A. P. Petropulu, "Blind MIMO system identification using PARAFAC decomposition of an output HOS-based tensor," in *Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, USA, nov 2003, pp. 1080–1084.
- [8] Y. Yu and A. Petropulu, "Blind MIMO system estimation based on PARAFAC decomposition of HOS tensors," in *IEEE Statistical Signal Processing Workshop*, Bordeaux, France, jul 2005.
- [9] Y. Yu and A. P. Petropulu, "Improved PARAFAC based blind MIMO system estimation," in *Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, USA, nov 2005, pp. 1010–1013.
- [10] G. Favier, *Calcul Matriciel et Tensoriel avec Applications à l'Automatique et au Traitement du Signal*, France, 2006, under preparation (in french).
- [11] J. B. Kruskal, "Three way arrays: rank and uniqueness of trilinear decompositions with applications to arithmetic complexity and statistics," *Linear Algebra and Its Applications*, vol. 18, pp. 95–138, 1977.
- [12] D. R. Brillinger and M. Rosenblatt, *Spectral Analysis of Time Series*. New York, USA: Wiley, 1967, ch. Computation and interpretation of k th-order spectra, pp. 189–232.
- [13] R. Bro, "PARAFAC. tutorial and applications." *Elsevier Chemometrics and Intelligent Laboratory Systems*, vol. 38, pp. 149–171, 1997.