ALTERNATING LEAST SQUARES IDENTIFICATION OF UNDER-DETERMINED MIXTURES BASED ON THE CHARACTERISTIC FUNCTION

Myriam Rajih, Pierre Comon

I3S, Algorithmes-Euclide-B, BP 121, F-06903 Sophia-Antipolis, France
phone: +33 4 9294 2793/2717, fax: +33 4 9294 2896, {rajih,comon}@i3s.unice.fr

ABSTRACT

Algorithm ALESCAF (Alternating LEast Squares identification based on the ChAracteristic Function) uses the derivatives of the second characteristic function (c.f.) of observations, without any need of sparsity assumption on sources, but assuming their statistical independence. ALESCAF was already proposed by the authors in [1], where only one derivative order was considered. In this paper, new versions of ALESCAF are proposed, that jointly use derivatives of different orders. We also propose ALESCAS, a new algorithm that uses the knowledge of source c.f.'s. Computer simulations demonstrate that both algorithms accelerate the convergence.

1. INTRODUCTION

Under-Determined Mixtures (UDM) refer to linear mixtures of independent random sources, where the number of sources, N, exceeds the number of sensors, P. Blind Channel Identification (BCI) of UDM's has been dealt with in [2] [3] [4] [5], where only data Fourth Order statistics were used.

The c.f. has been already utilized in [6] to blindly separate sources under the assumption that there are at most as many sources as sensors (i.e. over-determined mixtures). In this paper, we use the c.f. to identify linear mixtures where the number of sources exceeds the number of sensors (i.e. UDM).

The second c.f. was used in [7] for UDMs, but computer experiments were disappointing. In [1], we proposed an algebraic solution ALGECAF (ALGEbraic identification based on the ChAracteristic Function), proven to be very attractive for BCI of $2 \times N$ mixtures. However, ALGECAF turns out to be complicated to implement for $P > 2$. Hence, the importance of ALESCAF (Alternating LEast Squares identification based on the ChAracteristic Function), which reduced the BCI problem to a tensor decomposition, and accelerated the convergence [8].

Faster convergence can be obtained by exploiting the possible knowledge of source c.f.'s through algorithm ALESCAS [9] (Alternating LEast Squares identification based on the ChAracteristic function with known Source pdf's), and/or by using jointly derivatives of different orders, as subsequently shown. Some recent works dealt with BCI of UDM using tensor decomposition, see [10] and references therein. The number of sources is limited in [10], whereas the algorithms proposed here do not impose any bound on the number of sources, at least in theory. Practically, the complexity of the solution increases as the number of sources increases.

2. ASSUMPTIONS AND NOTATION

We assume the observation model below:

$$\boldsymbol{x} = \boldsymbol{A}\,\boldsymbol{s} + \boldsymbol{w} \tag{1}$$

where array variables are distinguished from scalars by bold faces, $\boldsymbol{x}$ and $\boldsymbol{s}$ are random vectors of size P and N respectively with $1 < P < N$, $\boldsymbol{A}$ is a $P \times N$ full rank matrix, and $\boldsymbol{w}$ accounts for modeling errors and additive noise of unknown distribution. From now on, its presence is just ignored in the remaining, except when running computer experiments. The entries s_n of vector $\boldsymbol{s}$ are assumed to be non Gaussian and statistically independent.

For simplicity, we shall restrict our attention in this paper to real variables and mixture, but most of the reasoning applies to the complex case, up to some complication in the notation. We also assume the following hypotheses:

- **H1** the columns of $\boldsymbol{A}$ are pairwise linearly independent
- **H2** source distributions are unknown and non Gaussian
- **H3** the number N of sources is known
- **H4** the moments of the sources are unknown, but finite up to some order larger than N

Under **H1**, **H2**, and **H3**, $\boldsymbol{A}$ can be shown to be essentially unique [11]. In algorithm ALESCAS, **H2** is replaced by:

H2s source distributions are known and non Gaussian.

The algorithms proposed are based on the core functional equation below, which is a direct consequence of source independence

$$\Psi_x(\boldsymbol{u}) = \sum_{n=1}^{N} \psi_n(\sum_{p=1}^{P} A_{pn} u_p) \tag{2}$$

where $\Psi_x(\boldsymbol{u})$ denotes the second c.f. of $\boldsymbol{x}$ defined as $\Psi_x(\boldsymbol{u}) = \log \mathrm{E}\{\exp(\jmath \boldsymbol{u}^\intercal \boldsymbol{x})\}$, and where $\psi_n(v)$ denotes the second c.f. of source s_n: $\psi_n(v) = \log \mathrm{E}\{\exp(\jmath v s_n)\}$, the dotless $\jmath$ being the square root of -1. This core equation can be used in an open neighborhood Ω of the origin, where Ψ_x does not vanish.

3. UNKNOWN SOURCE DISTRIBUTIONS: ALESCAF

In this section, we assume the hypothesis:

- **H4a** the source second characteristic functions ψ_n admit unknown finite derivatives up to order 3, at every point of some grid $\mathcal{G}$ of K values $\{\boldsymbol{u}[1], ..., \boldsymbol{u}[k]\} \in \Omega$

3.1. One Order Derivatives

For notation simplicity, suppose that $P \geq 3$, the case with 2 sensors will be discussed in 3.3. And take the second order derivatives of (2)

$$\frac{\partial^2 \Psi_x(\boldsymbol{u})}{\partial u_i \partial u_j} = \sum_{n=1}^{N} A_{in} A_{jn} \, \psi_n^{(2)}(\sum_q A_{qn} u_q) \qquad (3)$$

with $1 \leq i, j \leq P$. Take this equation on K points $\boldsymbol{u}[k] \in \mathcal{G} \subset \Omega$. Then, storing the left hand side of (3) in a family of symmetric matrices $T_{ij}[k]$, and denoting $D_{kn} = \psi_n^{(2)}(\sum_q A_{qn} u_q[k])$, (3) can be arranged in compact form as:

$$\boldsymbol{T}[k] = \boldsymbol{A} \, \mathrm{Diag}\{\boldsymbol{D}(k,:)\} \, \boldsymbol{A}^{\mathsf{T}}, \qquad (4)$$

with $1 \leq k \leq K$, where $\mathrm{Diag}\{\boldsymbol{v}\}$ denotes the diagonal matrix whose entries are those of vector $\boldsymbol{v}$, and where $\boldsymbol{D}(k,:)$ denotes the kth row of $\boldsymbol{D}$. Expression (4) can be alternatively written as:

$$\mathcal{T}_{ijk} = \sum_{n=1}^{N} A_{in} A_{jn} D_{kn} \qquad (5)$$

where $\boldsymbol{\mathcal{T}}$ is a $P \times P \times K$ tensor. (4) and (5) equivalently represent a 3-way PARAFAC model, and can be efficiently solved by using the ELS algorithm described in [8].

This procedure constitutes algorithm ALESCAF(2), and is able to compute $\boldsymbol{A}$ and $\boldsymbol{D}$ from tensor $\boldsymbol{\mathcal{T}}$, where the implicit dependence of $\boldsymbol{D}$ on $\boldsymbol{A}$ is ignored.

3.2. Multiple Order Derivatives

The previous algorithm only uses one (the second) order derivatives of the second c.f. of the observations to build tensor $\boldsymbol{\mathcal{T}}$, and proved to have some limitations in terms of identifiability and convergence speed. It can be made faster by adding extraneous terms to tensor $\boldsymbol{\mathcal{T}}$, that correspond to other order derivatives of (2). In fact, take the P further derivatives of (3):

$$\frac{\partial^3 \Psi_x(\boldsymbol{u})}{\partial u_i \partial u_j \partial u_p} = \sum_{n=1}^{N} A_{in} A_{jn} A_{pn} \, \psi_n^{(3)}(\sum_q A_{qn} u_q) \qquad (6)$$

with $1 \leq p \leq P$. Denote by $\boldsymbol{\mathcal{T}}[2]$ the $P \times P \times K$ tensor previously defined in (5), where the number 2 stands for the order of the derivatives used. Again, take equation (6) on K points $\boldsymbol{u}[k] \in \mathcal{G} \subset \Omega$. Note that the latter K points can be different from those of section (3.1). For simplicity, we keep the same K points $\boldsymbol{u}[k]$.

Then, (6) leads to P new $P \times P \times K$ tensors $\boldsymbol{\mathcal{T}}[3,p]$:

$$\mathcal{T}[3,p]_{::k} = \boldsymbol{A} \, \mathrm{Diag}\{\tilde{\boldsymbol{D}}_p(k,:)\} \, \boldsymbol{A}^{\mathsf{T}}$$

where $(\tilde{\boldsymbol{D}}_p)_{kn} = A_{pn} \psi_n^{(3)}(\sum_q A_{qn} u_q[k])$. By putting $\boldsymbol{\mathcal{T}}[2]$ and the P tensors $\boldsymbol{\mathcal{T}}[3,p]$ next to each other in the third mode, we obtain a bigger tensor of size $P \times P \times (P+1)K$. This rearrangement is shown in figure 1. This constitutes algorithm ALESCAF(2,3), that simultaneously uses the second and third order derivatives of the second c.f. of the observations, and accelerates the convergence as will be shown in the simulation section.

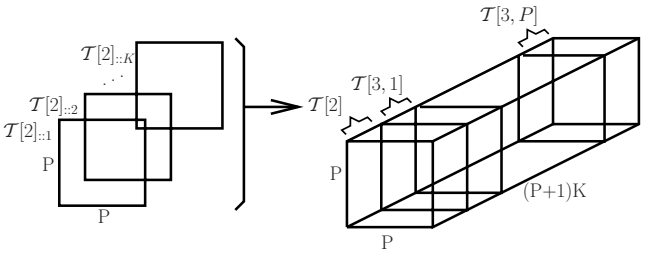

Fig. 1: Building new $P \times P \times (P+1)K$ tensor from tensors $\boldsymbol{\mathcal{T}}[2]$ and $\boldsymbol{\mathcal{T}}[3,p]$, $1 \leq p \leq P$.

3.3. Example 1

As mentioned in 3.1, let's take a more concrete example and suppose that $P = 2$, $N = 3$. In this case, we need to take the third order derivatives of the second c.f. of the observations, as the second order derivatives yield a 3-way PARAFAC model that does not achieve Kruskal uniqueness condition [12]:

$$2\, r_k(\boldsymbol{A}) + r_k(\tilde{\boldsymbol{D}}) \geq 2\, rank\{\boldsymbol{\mathcal{T}}[2]\} + 2 \qquad (7)$$

Then, we obtain a 4-way PARAFAC model defined by:

$$\mathcal{T}_{ijpk} = \sum_{n=1}^{N} A_{in} A_{jn} A_{pn} D_{kn}$$

where $\boldsymbol{\mathcal{T}}$ is a $P \times P \times P \times K$ tensor and $D_{kn} = \psi_n^{(3)}(\sum_q A_{qn} u_q[k])$. Now, uniqueness is achieved as (7) is verified: $3P + N = 9 \geq 2N + 3 = 9$.

We cannot use ALESCAF(2,3) because $r_k(\tilde{\boldsymbol{D}})$ cannot be at least as large as 5. But ALESCAF(3) generically works. But we can use ALESCAF(3,4). Then, entries of the new $P \times P \times P \times (P+1)K$ tensor $\boldsymbol{\mathcal{T}}$ are defined as, with $1 \leq p \leq P$:

$$\mathcal{T}_{::pk} = \begin{pmatrix} \frac{\partial^3 \Psi_x(\boldsymbol{u}[k])}{\partial u_1^2 \partial u_p} & \frac{\partial^3 \Psi_x(\boldsymbol{u}[k])}{\partial u_1 \partial u_2 \partial u_p} \\ \\ \frac{\partial^3 \Psi_x(\boldsymbol{u}[k])}{\partial u_1 \partial u_2 \partial u_p} & \frac{\partial^3 \Psi_x(\boldsymbol{u}[k])}{\partial u_2^2 \partial u_p} \end{pmatrix}, \; 1 \leq k \leq K$$

$$\mathcal{T}_{::pk} = \begin{pmatrix} \frac{\partial^4 \Psi_x(\boldsymbol{u}[k])}{\partial u_1^3 \partial u_p} & \frac{\partial^4 \Psi_x(\boldsymbol{u}[k])}{\partial u_1^2 \partial u_2 \partial u_p} \\ \\ \frac{\partial^4 \Psi_x(\boldsymbol{u}[k])}{\partial u_1^2 \partial u_2 \partial u_p} & \frac{\partial^4 \Psi_x(\boldsymbol{u}[k])}{\partial u_1 \partial u_2^2 \partial u_p} \end{pmatrix}, \; K+1 \leq k \leq 2K$$

$$\mathcal{T}_{::pk} = \begin{pmatrix} \frac{\partial^4 \Psi_x(\boldsymbol{u}[k])}{\partial u_1^2 \partial u_2 \partial u_p} & \frac{\partial^4 \Psi_x(\boldsymbol{u}[k])}{\partial u_1 \partial u_2^2 \partial u_p} \\ \\ \frac{\partial^4 \Psi_x(\boldsymbol{u}[k])}{\partial u_1 \partial u_2^2 \partial u_p} & \frac{\partial^4 \Psi_x(\boldsymbol{u}[k])}{\partial u_2^3 \partial u_p} \end{pmatrix}, \; 2K+1 \leq k \leq 3K$$

4. KNOWN SOURCE DISTRIBUTIONS: ALESCAS

In this section, we exploit the knowledge of the source distributions in order to accelerate the convergence. We assume the hypothesis:

H4b the source second characteristic functions admit finite derivatives up to order 8 at the origin

When source distributions are known, or are unknown but known to be equal, it is convenient to assure that $\mathcal{G} = \{0\}$, that is, to use a single point in the grid ($K = 1$). In fact, the argument of ψ_n in (2) then becomes independent of $\boldsymbol{A}$.

4.1. Equal source distributions: ALESCASE

We first consider that the N sources have the same (possibly unknown) distribution. The fourth order derivatives of (2) taken at the origin give:

$$\frac{\partial^4 \Psi_x(0)}{\partial u_i \partial u_j \partial u_p \partial u_q} = \psi^{(4)}(0) \sum_{n=1}^{N} A_{in} A_{jn} [A_{pn} A_{qn}] \quad (8)$$

with $1 \leq i, j, p, q \leq P$. (8) can be written as a 3-way tensor:

$$\mathcal{T}_{ijk} = \sum_{n=1}^{N} A_{in} A_{jn} \tilde{D}_{kn} \quad (9)$$

where

$$\tilde{\boldsymbol{D}} = \psi^{(4)}(0) \boldsymbol{A}^{\odot 2} \quad (10)$$

and $\boldsymbol{A}^{\odot q}$ is the column-wise non redundant symmetric Kronecker product of $\boldsymbol{A}$ with itself q times. If $\boldsymbol{A}$ is of size $2 \times N$ for instance, then $\boldsymbol{A}^{\odot q}$ is of size $q+1 \times N$.

Now, in the ELS algorithm, the dependence of $\tilde{\boldsymbol{D}}$ on $\boldsymbol{A}$ is exploited. In fact, at each iteration, $\tilde{\boldsymbol{D}}$ is not updated, as it is the case for ALESCAF, but computed from expression (10). This makes the convergence faster and the solution more stable, as will be seen in the simulation section.

Note that we do not need to know the source distributions, but we just need to know that those distributions are equal. In fact, $\psi^{(4)}(0)$ of expression (10) is a scale factor, and can be absorbed in $\boldsymbol{A}^{\odot 2}$. Then, $\tilde{\boldsymbol{D}}$ is updated by: $\tilde{\boldsymbol{D}} = \boldsymbol{A}^{\odot 2}$.

4.2. Unequal source distributions: ALESCASU

When the source distributions are different, $\tilde{\boldsymbol{D}}$ cannot be computed anymore from expression (10). It is even worse, due to the permutation ambiguity:

$$\tilde{\boldsymbol{D}} = \boldsymbol{A}^{\odot 2} \boldsymbol{P} \operatorname{Diag}\{[\psi_1^{(4)}(0), ..., \psi_N^{(4)}(0)]\} \boldsymbol{P}^{\intercal}$$

where $\boldsymbol{P}$ is a $N \times N$ permutation matrix. This ambiguity needs to be fixed, which leads to algorithm ALESCASU. As the symmetry constraint is relaxed in the ELS algorithm, minimizing the gap between both sides of (9) consists of minimizing:

$$\Upsilon = \sum_{ijk} \|\mathcal{T}_{ijk} - \sum_{n=1}^{N} A_{in} B_{jn} \tilde{D}_{kn}\|^2$$

After some (5 for example) ELS iterations on tensor $\boldsymbol{\mathcal{T}}$, ALESCASU consists of performing the following steps, at each iteration:

1. Estimate $\hat{\boldsymbol{A}}$ in the Least Squares (LS) sense from previous values of $\hat{\boldsymbol{B}}$ and $\tilde{\boldsymbol{D}}$, then normalize its columns

$$\hat{\boldsymbol{A}}_1 = \hat{\boldsymbol{A}} \operatorname{Diag}\{[\frac{1}{\|\hat{\boldsymbol{a}}_1\|}, ..., \frac{1}{\|\hat{\boldsymbol{a}}_N\|}]\}$$

   where $\|\hat{\boldsymbol{a}}_i\|$ is the norm of the i^{th} column of $\hat{\boldsymbol{A}}$

2. Estimate $\hat{\boldsymbol{D}}$, in the LS sense, from $\hat{\boldsymbol{A}}_1$, and previous values of $\hat{\boldsymbol{B}}$

3. Fix the permutation ambiguity by minimizing:

$$\epsilon = \|\hat{\boldsymbol{D}}_1 - \tilde{\boldsymbol{D}}_1\|_F^2$$

   where: $\hat{\boldsymbol{D}}_1 = \hat{\boldsymbol{D}} \operatorname{Diag}\{[\frac{1}{\|\hat{\boldsymbol{d}}_1\|}, ..., \frac{1}{\|\hat{\boldsymbol{d}}_N\|}]\}$

$$\tilde{\boldsymbol{D}} = \hat{\boldsymbol{A}}_1^{\odot 2} \boldsymbol{P}(\boldsymbol{D}\boldsymbol{\Lambda}) \boldsymbol{P}^{\intercal}$$

$$\boldsymbol{D} = \operatorname{Diag}\{[\psi_1^{(4)}(0), ..., \psi_N^{(4)}(0)]\}$$

$$\boldsymbol{\Lambda} = \operatorname{Diag}\{[\|\hat{\boldsymbol{a}}_1\|, ..., \|\hat{\boldsymbol{a}}_N\|]\}$$

   Practically, in order to fix $\boldsymbol{P}$, we try all possible permutations and keep the one that minimizes ϵ

4. Estimate $\hat{\boldsymbol{B}}$ in the LS sense
5. Use $\hat{\boldsymbol{A}}_1$, $\hat{\boldsymbol{B}}$, and $\tilde{\boldsymbol{D}}$ as initial values for the next iteration

ALESCASU reiterates and performs the same previous steps until the convergence criterion is met.

4.3. Example 2

Under conditions of example 1, and in the case of BPSK sources, we need to take the sixth order derivatives for ALESCAS. We then obtain a $2 \times 2 \times 2 \times 4$ tensor $\boldsymbol{\mathcal{T}}$:

$$\mathcal{T}_{::pk} = \begin{pmatrix} \frac{\partial^6 \Psi_x(0)}{\partial u_1^{6-k} \partial u_2^{k-1} \partial u_p} & \frac{\partial^6 \Psi_x(0)}{\partial u_1^{5-k} \partial u_2^{k} \partial u_p} \\ \frac{\partial^6 \Psi_x(0)}{\partial u_1^{5-k} \partial u_2^{k} \partial u_p} & \frac{\partial^6 \Psi_x(0)}{\partial u_1^{4-k} \partial u_2^{k+1} \partial u_p} \end{pmatrix}$$

with $1 \leq p \leq P$, and $1 \leq k \leq 4$.

As mentioned in the introduction, the number of sources N is theoretically unbounded. However, as N increases, the complexity of both ALESCAF and ALESCAS increases because the order of the derivatives and the tensor order need to be increased. In fact, if we keep $P = 2$ and increase N by one (N=4), uniqueness is no longer achieved (7). One solution is to use the 8^{th} order derivatives of (2) and build the corresponding 5-way tensor of size $2 \times 2 \times 2 \times 2 \times 5$:

$$\mathcal{T}_{::lpk} = \begin{pmatrix} \frac{\partial^8 \Psi_x(0)}{\partial u_1^{7-k} \partial u_2^{k-1} \partial u_l \partial u_p} & \frac{\partial^8 \Psi_x(0)}{\partial u_1^{6-k} \partial u_2^{k} \partial u_l \partial u_p} \\ \frac{\partial^8 \Psi_x(0)}{\partial u_1^{6-k} \partial u_2^{k} \partial u_l \partial u_p} & \frac{\partial^8 \Psi_x(0)}{\partial u_1^{5-k} \partial u_2^{k+1} \partial u_l \partial u_p} \end{pmatrix}$$

5. COMPUTER RESULTS

We consider the linear model of expression (1), with $P = 2$ and $N = 3$. Sources are BPSK, that is, they take their values in $\{-1, 1\}$ with equal probabilities. The channel matrix $\boldsymbol{A}$ is

$$\boldsymbol{A} = \begin{pmatrix} 1 & cos(\theta) & 0 \\ 0 & sin(\theta) & 1 \end{pmatrix}$$

Firstly, the influence of the joint use of several order derivatives of the second c.f. of the observations is analyzed. For this purpose, we compare the performances of ALESCAF(3),

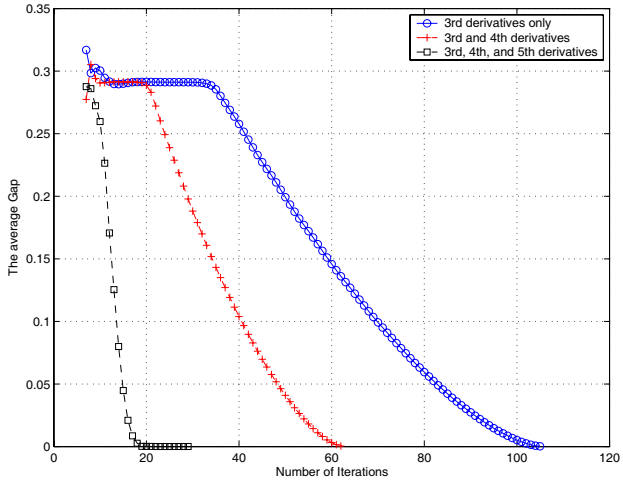

Fig. 2: Gap between original and estimated channel matrix using ALESCAF(3), ALESCAF(3,4), and ALESCAF(3,4,5) for $(P, N) = (2, 3)$, BPSK sources, infinite block, $\theta = \pi/6$.

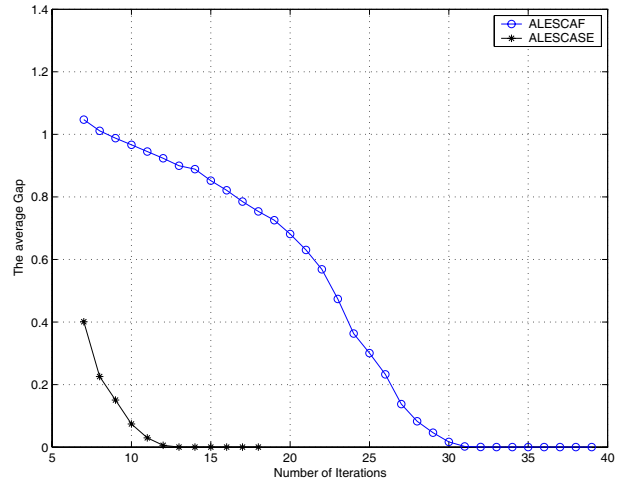

Fig. 3: Gap between original and estimated channel matrix using ALESCAF(6) and ALESCASE for $(P, N) = (2, 3)$, BPSK sources, infinite block, $\theta = \pi/6$.

ALESCAF(3,4), and ALESCAF(3,4,5), in figure 2. An "infinite block" of data is implicitly assumed by taking all the 2^3 possible combinations of $\{-1, 1\}$.

Secondly, we analyze the influence of the knowledge of the source distributions. In figure 3, we use ALESCASE where the 3 sources are BPSK. And in figure 4, we use ALESCASU where 2 sources are BPSK, and one source is 4PAM, that is, it takes its values in $\{-3, -1, 1, 3\}$ with equal probabilities.

6. CONCLUSION

We proposed new versions of algorithm ALESCAF, that simultaneously use derivatives of different orders of the second c.f. of the observations, and demonstrated that they accelerated the convergence. We also proposed two variants of ALESCAS, that exploit the knowledge of the source distributions: (i) ALESCASE supposes that the sources have equal distributions, and does not need the explicit value of the distribution ; (ii) ALESCASU allows to have sources with unequal distributions. We demonstrated that the convergence is accelerated in both cases.

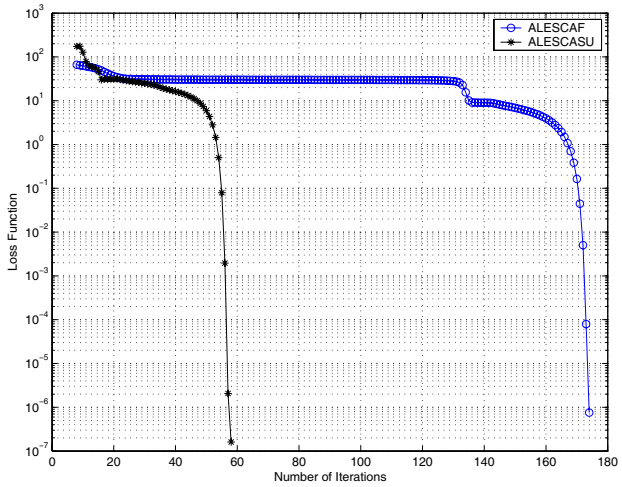

Fig. 4: The error Υ using ALESCAF(6) and ALESCASU for $(P, N) = (2, 3)$, 2 BPSK and one 4PAM sources, infinite block, $\theta = \pi/6$.

7. REFERENCES

[1] P. COMON and M. RAJIH, "Blind identification of under-determined mixtures based on the characteristic function," *ICASSP Conference*, Mar 18–23 2005.

[2] J. F. CARDOSO, "Super-symmetric decomposition of the fourth-order cumulant tensor. Blind identification of more sources than sensors," in *ICASSP'91*, 1991, pp. 3109–3112.

[3] P. COMON, "Blind identification and source separation in 2x3 under-determined mixtures," *IEEE Trans. Signal Processing*, pp. 11–22, Jan. 2004.

[4] L. de LATHAUWER, P. COMON, and B. de MOOR, "ICA algorithms for 3 sources and 2 sensors," in *6th Sig. Proc. Workshop on HOS*, Caesarea, Israel, June 14–16 1999, pp. 116–120.

[5] A. FERREOL, L. ALBERA, and P. CHEVALIER, "Fourth order blind identification of underdetermined mixtures of sources (FOBIUM)," in *ICASSP'03*, Hong Kong, China, April 6-10 2003, pp. 41–44.

[6] A. YEREDOR, "Blind source separation via the second characteristic function," *Sig. Proc.*, vol. 80, pp. 897–902, 2000.

[7] A. TALEB, "An algorithm for the blind identification of n independent signals with 2 sensors," in *6th ISSPA*, Kuala Lumpur, Aug. 13-16 2001, vol. 1, pp. 5–8.

[8] M. RAJIH and P. COMON, "Enhanced line search: A novel method to accelerate Parafac," in *Eusipco'05*, Antalya, Turkey, Sept. 4-8 2005.

[9] M. RAJIH and P. COMON, "Blind identification of under-determined mixtures based on the characteristic function : Influence of source pdf's," *CAMSAP*, Dec 13–15 2005.

[10] L. De Lathauwer, "Simultaneous matrix diagonalization: the overcomplete case," *ICA'03*, pp. 821–825, apr 1–4 2003.

[11] A. M. KAGAN, Y. V. LINNIK, and C. R. RAO, *Characterization Problems in Mathematical Statistics*, Wiley, 1973.

[12] J. B. KRUSKAL, "Three-way arrays: Rank and uniqueness of trilinear decomp.," *Lin. Alg. Appl.*, vol. 18, pp. 95–138, 1977.