

THIN QR AND SVD FACTORIZATIONS FOR SIMULTANEOUS BLIND SIGNAL EXTRACTION

Sergio Cruces¹, Andrzej Cichocki², Lieven De Lathauwer³

¹ Teoría de la Señal y Comunicaciones,
Ingenieros, 41092-Sevilla, Spain.
e-mail: sergio@us.es

² Lab. for Advanced Brain Signal Proc.,
Riken, Brain Science Institute, Japan.
e-mail: cia@bsp.brain.riken.go.jp

³ ETIS (ENSEA, UCP, CNRS)
Cergy-Pontoise, France.
e-mail: delathau@ensea.fr

ABSTRACT

This paper studies the problem of the simultaneous blind signal extraction of a subset of independent components from a linear mixture. In order to solve it in a robust manner, we consider the optimization of contrast functions that jointly exploit the information provided by several cumulant tensors of the observations. We develop hierarchical and simultaneous ICA extraction algorithms that are able to optimize the proposed contrast functions. These algorithms are based on the thin-QR and thin-SVD factorizations of a matrix of weighted cross-statistics between the observations and outputs. Simulations illustrate the good performance of the proposed methods.

1. INTRODUCTION

Blind signal extraction (BSE) consists in the estimation of a subset of the independent components that appear linearly combined in the observations. BSE includes Blind Signal Separation (BSS) as the particular case where one is interested in all the independent components. In the last decade powerful criteria and algorithms have been developed to solve this problem [1]-[7].

Popular techniques like JADE [2] and SOBI [3] are robust in the sense that they use joint criteria to obtain accurate estimates from the available data. Other approaches, like the higher-order power method (HOPM) [4] try to find the best least-squares rank-1 approximation to a high-order cumulant tensor. In this paper, we propose a robust reformulation of the later criterion which consists in finding the best weighted least-squares low-rank approximation to a set of cross-cumulant tensors of the observations. For the optimization of the resulting criterion we develop the Thin-ICA algorithm, whose hierarchical and simultaneous implementations are based, respectively, on the thin-QR and thin-SVD factorizations. This algorithm combines the advantages in flexibility of the simultaneous extraction methods with the good performance of the robust methods.

2. SIGNAL MODEL AND ASSUMPTIONS

Figure 1 shows the signal model. The complex vector of observations $\mathbf{x}(t) = [x_1(t), \dots, x_M(t)]^T$ obeys the following equation

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) \quad (1)$$

where $\mathbf{s}(t) = [s_1(t), \dots, s_N(t)]^T$ is the complex signal vector process of N independent components, $\mathbf{n}(t)$ the noise vector process, and $\mathbf{A} \in \mathbb{C}^{M \times N}$ is the mixing matrix ($M \geq N$).

We consider the following assumptions:

- A1** The components of $\mathbf{s}(t)$ are mutually independent, locally stationary and normalized to zero mean and unit variance.
- A2** The noise vector process $\mathbf{n}(t)$ is independent from $\mathbf{s}(t)$, locally stationary, Gaussian, white ($\mathbf{R}_{\mathbf{n}}(t_2, t_1) = \delta(t_2 - t_1)E[\mathbf{n}(t_1)(\mathbf{n}(t_2))^H]$) and with a known correlation matrix $\mathbf{R}_{\mathbf{n}}(t, t) = E[\mathbf{n}(t)(\mathbf{n}(t))^H]$ or one that can be accurately estimated from the observations.
- A3** The mixing matrix \mathbf{A} is full-column rank.

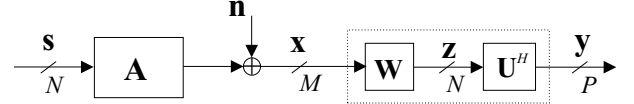


Figure 1: Signal model for the blind extraction of P sources.

- A4** There exists an order relation among the sources that is maximized by those we want to extract. For a given subset of P desired independent components $\{s_1(t), \dots, s_P(t)\}$, there exist time tuples $\theta = (t_1, \dots, t_q)$ contained in the set $\Theta = \{\theta_m \in \mathbb{R}^q, m = 1, \dots, r : \text{if } q > 2, \theta_m \in \mathbb{R}^2 \setminus \{(t, t), \forall t \in \mathbb{R}\} \text{ if } q = 2\}$ and some chosen positive weighting scalars w_θ (normalized so as to verify $\sum_{\theta \in \Theta} w_\theta = 1$) that sort the following statistic of the sources

$$\psi_\Theta(s_j) = \sum_{\theta \in \Theta} w_\theta |Cum(s_j(t_1), \dots, s_j(t_q))|^2,$$

in such a way that these inequalities hold true

$$\psi_\Theta(s_i) \geq \psi_\Theta(s_j), \quad 1 \leq i \leq P < j \leq N. \quad (2)$$

From (1) one obtains that $\mathbf{A}\mathbf{A}^H = \mathbf{R}_{\mathbf{x}}(t, t) - \mathbf{R}_{\mathbf{n}}(t, t)$. Let $\mathbf{Q}_1 \text{diag}(\sigma_1, \dots, \sigma_N) \mathbf{Q}_1^H$ denote the trimmed down version of the Schur decomposition of $\mathbf{R}_{\mathbf{x}}(t, t) - \mathbf{R}_{\mathbf{n}}(t, t)$. The $N \times M$ prewhitening system $\mathbf{W} = \text{diag}(\sigma_1^{-1/2}, \dots, \sigma_N^{-1/2}) \mathbf{Q}_1^H$ projects the observations onto the signal subspace and also spheres the resulting vector of preprocessed observations

$$\mathbf{z}(t) = \mathbf{W}\mathbf{x}(t) \quad (3)$$

In order to extract P desired independent components ($1 \leq P \leq N$) we multiply them by the $P \times N$ semi-unitary matrix \mathbf{U}^H , being $\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_P]$ formed by orthonormal columns. This way, we obtain the vector of P outputs or estimated sources

$$\mathbf{y}(t) = \mathbf{U}^H \mathbf{z}(t) \quad (4)$$

3. A ROBUST CRITERION

One of the most popular contrast functions for the blind signal extraction of a single source is based on the maximization of a higher order cumulant of the output subject to a normalizing constraint. A result in [5] shows that, subject to the semi-unitarity of \mathbf{U} (i.e. $\mathbf{U}^H \mathbf{U} = \mathbf{I}_P$), the following function

$$\Phi_\Theta(\mathbf{U}) = \sum_{i=1}^P \sum_{\theta \in \Theta} w_\theta |Cum(y_i(t_1), \dots, y_i(t_q))|^2 \quad (5)$$

is a contrast for the extraction of P independent components: $s_1(t), \dots, s_P(t)$. By considering cross-cumulants of the observations at different time tuples $\theta = (t_1, \dots, t_q)$, contained the set Θ ,

one exploits the temporal correlation of the observation process and its possibly long term non-stationarity.

Let us introduce, for later convenience, q semi-unitary matrices $\mathbf{U}^{[k]}, k = 1, \dots, q$, that try to estimate the mixing system $\mathbf{W}\mathbf{A}$, and their respective linear estimates $\mathbf{y}^{[k]}(t) = \mathbf{U}^{[k]H} \mathbf{z}(t)$, $k = 1, \dots, q$ of the vector with the desired independent components. A less constrained contrast than (5) results from allowing the arguments of the cumulants to be different [6], i.e.,

$$\Phi_{\Theta}(\mathbf{U}) = \sum_{i=1}^P \sum_{\theta \in \Theta} w_{\theta} \left| \text{Cum} \left(y_i^{[1]}(t_1), \dots, y_i^{[q]}(t_q) \right) \right|^2, \quad (6)$$

where $\mathbf{U} \equiv \{\mathbf{U}^{[1]}, \dots, \mathbf{U}^{[q]}\}$.

Let us denote by $\mathbf{D}_{\mathbf{y}}(\theta) = \text{diag}(d_{y_1}(\theta), \dots, d_{y_P}(\theta))$ a diagonal matrix of P complex scalars, each one depending on one of the outputs and on the vector θ . The robust nature of the proposed contrast function is a consequence of its least squares interpretation given in [4]. The contrast function is a performance index of how well the set of P -component q -way arrays $\{\mathcal{C}_q^{\mathbf{z}}(\mathbf{D}_{\mathbf{y}}(\theta), \mathbf{U}), \forall \theta \in \Theta\}$, with $1 \leq j_1, \dots, j_q \leq N$ elements

$$[\mathcal{C}_q^{\mathbf{z}}(\mathbf{D}_{\mathbf{y}}(\theta), \mathbf{U})]_{j_1, \dots, j_q} = \sum_{i=1}^P d_{y_i}(\theta) (u_{j_1 i}^{[1]} \dots u_{j_q i}^{[q]}), \quad (7)$$

approximate, in weighted least squares sense, the set of super-symmetric arrays of q -th-order cross-cumulants of the observations $\{\mathcal{C}_q^{\mathbf{z}}(\theta) \in \mathbb{C}^{N \times \dots \times N}, \forall \theta \in \Theta\}$, with $1 \leq j_1, \dots, j_q \leq N$ elements

$$[\mathcal{C}_q^{\mathbf{z}}(\theta)]_{j_1, \dots, j_q} = \text{Cum}(z_{j_1}(t_1), \dots, z_{j_q}(t_q)). \quad (8)$$

as indicated by the following lemma.

Lemma 1 *The constrained maximization of the contrast function $\Phi_{\Theta}(\mathbf{U})$ is equivalent to the constrained minimization of the weighted error function of the approximation*

$$\varepsilon_{\Theta}(\mathbf{U}) = \sum_{\theta \in \Theta} w_{\theta} \min_{\mathbf{D}_{\mathbf{y}}(\theta)} \|\mathcal{C}_q^{\mathbf{z}}(\theta) - \mathcal{C}_q^{\mathbf{z}}(\mathbf{D}_{\mathbf{y}}(\theta), \mathbf{U})\|_F^2$$

with respect to the set of semi-unitary matrices in \mathbf{U} .

Proof: The proof of the lemma follows from the fact that

$$\varepsilon_{\Theta}(\mathbf{U}) = \sum_{\theta \in \Theta} w_{\theta} \|\mathcal{C}_q^{\mathbf{z}}(\theta)\|_F^2 - \Phi_{\Theta}(\mathbf{U}),$$

where the operator $\|\cdot\|_F^2$ returns the accumulated energy of the elements in the array. \square

4. THE THIN-ICA EXTRACTION ALGORITHM

A good method to maximize the proposed contrast is to optimize it cyclically with respect to each of the matrix arguments while keeping fixed the others. For this purpose, we will find it useful to distinguish between sequential notation (k) which indicates that the variable takes its value at the k -th iteration and the cyclic or modulo- q notation specified by the corresponding superindex $[k]$ where $[k] = (k \bmod q) + 1$. Then, at the (k) -th iteration we maximize $\Phi_{\Theta}(\mathbf{U}^{[1]}, \dots, \mathbf{U}^{[q]})$ with respect to the extraction matrix $\mathbf{U}^{[k]} \equiv \mathbf{U}^{(k)}$ while $\mathbf{U}^{[k-1]}, \dots, \mathbf{U}^{[k-q+1]}$ are kept fixed. It is convenient to define a new function

$$\begin{aligned} \Phi_{\Theta}^{(k-1)}(\mathbf{U}^{(k)}) &= \Phi_{\Theta}(\dots, \mathbf{U}^{[k-q+1]}, \mathbf{U}^{(k)}, \mathbf{U}^{[k-1]}, \dots) \\ &= \sum_{i=1}^P \mathbf{u}_i^{(k)H} \mathbf{M}_i^{(k-1)} \mathbf{u}_i^{(k)} \end{aligned} \quad (9)$$

which separates the dependence of the variable to optimize $\mathbf{U}^{(k)}$ (the argument of this new function) from the previous ones

whose influence is collected into the constant matrices $\mathbf{M}_i^{(k-1)}, i = 1, \dots, P$, given by

$$\mathbf{M}_i^{(k-1)} = \sum_{\theta \in \Theta} w_{\theta} \mathbf{c}_{\mathbf{zy}_i}^{(k-1)}(\theta) \left(\mathbf{c}_{\mathbf{zy}_i}^{(k-1)}(\theta) \right)^H \quad (10)$$

$$\begin{aligned} \mathbf{c}_{\mathbf{zy}_i}^{(k-1)}(\theta) &= \text{Cum}(\mathbf{z}(t_{[k]}), y_i^{(k-1)}(t_{[k-1]}), \dots \\ &\quad \dots, y_i^{(k-q+1)}(t_{[k-q+1]})) \end{aligned} \quad (11)$$

The cyclic maximization of $\Phi_{\Theta}(\mathbf{U}^{[1]}, \dots, \mathbf{U}^{[q]})$ is implemented by the sequential maximization, through iterations, of the functions $\Phi_{\Theta}^{(k-1)}(\mathbf{U}^{(k)})$, updating the cyclic variable $\mathbf{U}^{[k]} = \mathbf{U}^{(k)}$ after each optimization.

The key role in the algorithm will be played by the following matrix weighted statistic

$$\tilde{\mathbf{C}}_{\mathbf{zy}}^{(k-1)} = [\mathbf{M}_1^{(k-1)} \mathbf{u}_1^{(k-q)}, \dots, \mathbf{M}_P^{(k-1)} \mathbf{u}_P^{(k-q)}] \quad (12)$$

which is proportional to the gradient of $\Phi_{\Theta}^{(k-1)}(\cdot)$ evaluated at $\mathbf{U}^{(k-q)}$. Let us define the thin-SVD factorization of the statistic

$$\tilde{\mathbf{C}}_{\mathbf{zy}}^{(k-1)} = \mathbf{V}_L \mathbf{\Lambda}_{P \times P} \mathbf{V}_R^H, \quad (13)$$

where \mathbf{V}_L is a $N \times P$ matrix formed by the P left singular vectors associated to the singular values of $\tilde{\mathbf{C}}_{\mathbf{zy}}^{(k-1)}$, $\mathbf{\Lambda}_{P \times P}$ is the diagonal matrix of singular values and \mathbf{V}_R is the $P \times P$ matrix of right singular vectors.

Theorem 1 (Simultaneous optimization) *In each iteration, the choice resulting from the thin-SVD factorization*

$$\mathbf{U}^{(k)} = \mathbf{V}_L \mathbf{V}_R^H \quad (14)$$

guarantees a monotonous ascent in the contrast function

$$\Phi_{\Theta}^{(k-2)}(\mathbf{U}^{(k-1)}) \leq \Phi_{\Theta}^{(k-1)}(\mathbf{U}^{(k)}) \quad (15)$$

Proof: We will prove that the new candidate $\mathbf{U}^{(k)}$ satisfies the following chain of inequalities

$$\Phi_{\Theta}^{(k-2)}(\mathbf{U}^{(k-1)}) = \sum_{i=1}^P \mathbf{u}_i^{(k-1)H} \mathbf{M}_i^{(k-2)} \mathbf{u}_i^{(k-1)} \quad (16)$$

$$= \sum_{i=1}^P \mathbf{u}_i^{(k-q)H} \mathbf{M}_i^{(k-1)} \mathbf{u}_i^{(k-q)} \quad (17)$$

$$\stackrel{(a)}{\leq} \sum_{i=1}^P \mathbf{u}_i^{(k)H} \mathbf{M}_i^{(k-1)} \mathbf{u}_i^{(k)} \quad (18)$$

$$\stackrel{(b)}{\leq} \sum_{i=1}^P \mathbf{u}_i^{(k)H} \mathbf{M}_i^{(k-1)} \mathbf{u}_i^{(k)} \quad (19)$$

$$= \Phi_{\Theta}^{(k-1)}(\mathbf{U}^{(k)}) \quad (20)$$

which guarantees the monotonous ascent. The first inequality can be rewritten in the form

$$\text{Tr}\{\mathbf{U}^{(k-q)H} \tilde{\mathbf{C}}_{\mathbf{zy}}^{(k-1)}\} \stackrel{(a)}{\leq} \text{Tr}\{\mathbf{U}^{(k)H} \tilde{\mathbf{C}}_{\mathbf{zy}}^{(k-1)}\},$$

and it is straightforward to show that it is true for $\mathbf{U}^{(k)} = \mathbf{V}_L \mathbf{V}_R^H$ since this proposed choice is the one that maximizes the right-hand-part of the equation, simultaneously, with respect to all the columns $\mathbf{u}_1^{(k)}, \dots, \mathbf{u}_P^{(k)}$.

Each of the hermitian and positive semidefinite matrices $\mathbf{M}_i^{(k-1)}$ have a unique hermitian and positive semidefinite square root $(\mathbf{M}_i^{(k-1)})^{1/2}$. Then, we can define the vectors

$$\begin{aligned}\alpha^{(k)} &= [(\mathbf{M}_1^{(k-1)})^{1/2} \mathbf{u}_1^{(k)}; \dots; (\mathbf{M}_P^{(k-1)})^{1/2} \mathbf{u}_P^{(k)}] \\ \alpha^{(k-1)} &= [(\mathbf{M}_1^{(k-2)})^{1/2} \mathbf{u}_1^{(k-1)}; \dots; (\mathbf{M}_P^{(k-2)})^{1/2} \mathbf{u}_P^{(k-1)}]\end{aligned}$$

and rewrite the chain of inequalities as

$$\Phi_{\Theta}^{(k-2)}(\mathbf{U}^{(k-1)}) = \|\alpha^{(k-1)}\|^2 \stackrel{(a)}{\leq} |\alpha^{(k)H} \alpha^{(k-1)}| \stackrel{(b)}{\leq} \Phi_{\Theta}^{(k-1)}(\mathbf{U}^{(k)})$$

By applying Cauchy-Schwarz inequality

$$|\alpha^{(k)H} \alpha^{(k-1)}| \leq \|\alpha^{(k)}\| \|\alpha^{(k-1)}\| \quad (21)$$

one can see from (a) that $\|\alpha^{(k-1)}\| \leq \|\alpha^{(k)}\|$. This also proves (b) and, thus, the monotonous ascent $\Phi_{\Theta}^{(k-2)}(\mathbf{U}^{(k-1)}) \leq \Phi_{\Theta}^{(k-1)}(\mathbf{U}^{(k)})$. \square

Due to the bounded nature of the contrast function, the monotonous ascent guarantees that the only strictly stable points of the algorithm are the local maxima of the contrast. To the contrary to other algorithms, here the monotonous ascent property does not depend on the signs of the cumulants of the sources.

A different implementation of the Thin-ICA algorithm can be obtained from the hierarchical optimization of the contrast function.

Theorem 2 (Hierarchical optimization) *In each iteration, the following choice*

$$\mathbf{U}^{(k)} = \mathbf{Q}, \quad (22)$$

resulting from the thin-QR factorization $\tilde{\mathbf{C}}_{\mathbf{zy}}^{(k-1)} = \mathbf{Q}\mathbf{R}_{P \times P}$, performs the hierarchical optimization of the contrast function with respect to the columns of $\mathbf{U}^{(k)}$. This choice does not guarantee an overall monotonous ascent in the contrast, but it does guarantee a hierarchical monotonous ascent in the first non-convergent mode, i.e., when $\mathbf{u}_1, \dots, \mathbf{u}_{i-1}$ have already converged to some fixed value then

$$\mathbf{u}_i^{(k-1)H} \mathbf{M}_i^{(k-2)} \mathbf{u}_i^{(k-1)} \leq \mathbf{u}_i^{(k)H} \mathbf{M}_i^{(k-1)} \mathbf{u}_i^{(k)}. \quad (23)$$

The idea for the proof of theorem consists in that the choice $\mathbf{U}^{(k)} = \mathbf{Q}$ hierarchically maximizes $\text{Tr}\{\mathbf{U}^{(k)H} \tilde{\mathbf{C}}_{\mathbf{zy}}^{(k-1)}\}$ with respect to the columns $\mathbf{u}_1, \dots, \mathbf{u}_P$, in such a way that each i -th column satisfies the constraint $\mathbf{u}_i^{(k)H} \mathbf{u}_j^{(k)} = \delta_{ij} \forall j \leq i$. The optimized vector $\mathbf{u}_i^{(k)}$ in this implementation is identified with the first column of a Householder reflection.

Taking into account that the first columns are less constrained than the last ones, together with assumption **A4**, it results convenient after each iteration to sort the columns of $\mathbf{U}^{(k)}$, and its predecessors, according to $\gamma_i^{(k)} = \mathbf{u}_i^{(k)H} \mathbf{M}_i^{(k-1)} \mathbf{u}_i^{(k)}$, so as to obtain

$$\gamma_1^{(k)} \geq \dots \geq \gamma_P^{(k)}. \quad (24)$$

4.1 Supersymmetry of the solutions

The mutual independence of the sources guarantees the theoretical supersymmetry of the arrays of cross-cumulants of the observations $\mathcal{C}_q^{\mathbf{z}}(\theta), \forall \theta \in \Theta$. However, in practice, small stochastic deviations occur at the estimation of the cross-cumulants from the observations, which result in a loss of this property. When this happens, the best approximation is not supersymmetric and the candidate matrices, in general, have different values $\mathbf{U}^{[1]} \neq \dots \neq \mathbf{U}^{[q]}$ at their convergence. Fortunately, the previous situation can be easily prevented by restoring the supersymmetry of the arrays after their estimation. This allows the algorithm to converge to the solutions $\mathbf{U}^{[1]} = \dots = \mathbf{U}^{[q]}$ which extract the independent components.

4.2 Projection

The knowledge of the supersymmetry of the solutions can be also exploited to accelerate the convergence of algorithm by adding, at the end of each k -th iteration, a projection step onto the supersymmetric manifold

$$\{\mathbf{U}^{(k)}, \mathbf{U}^{(k-1)}, \dots, \mathbf{U}^{(k-q+1)}\} \longrightarrow \{\mathbf{U}^{(k)}, \dots, \mathbf{U}^{(k)}\}$$

However, this projection should be applied with care. It is only justified in special cases, such as when for each $\theta \in \Theta$ all the signs of $\text{Cum}(s_i(t_1), \dots, s_i(t_q))$, $i = 1, \dots, N$, coincide. This is a situation where a convexity result applies to the contrast function, preserving the monotonous ascent (see [7] for more details).

4.3 Extraction of a single source

When one considers the extraction of a single source the QR and SVD implementations of the Thin-ICA algorithm coincide. In this case, and as long as the cumulants are not affected by the noise, a result presented in [6] proves that the contrast function has no deceptive maxima, i.e., all the local maxima correspond to solutions that extract independent components.

Additionally, in the case of the approximation to a single cumulant tensor, one can observe that the Thin-ICA algorithm without projection reduces to the high order power method [4], while with the projection step it reduces to symmetric version of the high order power method [7] or fixed point implementation of the Fast-ICA algorithm [8].

4.4 Combining statistics of different orders

It is also possible to combine cross-cumulants of different orders in the contrast function but this requires a slightly more cumbersome notation and, for this reason, it has not been initially considered in the paper. The basic idea is to use as much extraction candidates as the maximum order q_m of the involved cross-cumulants, and to define an invariant contrast function $\Phi_{\Theta}(\mathbf{U}^{[1]}, \dots, \mathbf{U}^{[q_m]})$ with respect to permutations in the matrix arguments. The permutation invariance can be obtained by including in the contrast function, for any given order $q \leq q_m$, cross-cumulants with all the possible combinations of q arguments that can be obtained from $\{y_i^{[1]}(t_1), \dots, y_i^{[q_m]}(t_{q_m})\}$.

In [9] one can find an example of application of a Thin-ICA algorithm to blind DS-CDMA detection where the estimation was improved by choosing a contrast function that combined fourth and sixth order cumulants.

5. SIMULATIONS

When we tested the proposed algorithm, we observed similar results to those of the SOBI and JADE when using the same cumulant tensors that these later algorithms exploit, and better results when some additional relevant tensors are taken into account. However, the Thin-ICA algorithm also allows the extraction of subsets of sources. In this section we will illustrate this last property.

In our example an array of 20 sensors registers 500 snapshots of the observations. These are a random instantaneous mixture of 10 independent signals, in presence of white additive Gaussian noise, and with a maximum signal to noise ratio of 15dB. The desired independent components are four correlated signals obtained, after normalization, from the filtering of four binary processes by the corresponding systems: $F_1(z^{-1}) = (1 + 0.9z^{-1})^{-1}$, $F_2(z^{-1}) = (1 - 0.9z^{-1})^{-1}$, $F_3(z^{-1}) = (1 + 0.9z^{-2})^{-1}$ and $F_4(z^{-1}) = (1 - 0.9z^{-2})^{-1}$. The other six independent components are samples of temporally i.i.d. uniform processes. We chose second order statistics $q = 2$ because for short data records like this, they are usually the most reliable, and we set $\Theta = \{(t, t-1), (t, t-2), \dots, (t, t-4)\}$ because these four pairs guarantee that the considered independent components can be ordered according to (2).

Table 1: Thin-ICA algorithm for extraction.

1. Set $P \leq N$ the number of independent components to extract from $\mathbf{x}(t)$.
2. Prewhitening $\mathbf{z}(t) = \mathbf{W}\mathbf{x}(t)$;
3. Estimate $\forall \theta \in \Theta$ the q -way arrays $\mathcal{C}_q^{\mathbf{z}}(\theta)$.
4. Initialization: $\mathbf{U}^{(-q+1)} = \dots = \mathbf{U}^{(0)} = \mathbf{I}_{N \times P}$;
 $\mathbf{y}^{(0)}(t) = \mathbf{U}^{(0)H} \mathbf{z}(t)$; $k = 1$;
5. Compute $\forall \theta \in \Theta$ the cross-cumulant vectors $\mathbf{c}_{\mathbf{zy}_i}^{(k-1)}(\theta)$, $i = 1, \dots, P$ of elements $(1 \leq j \leq N)$

$$[\mathbf{c}_{\mathbf{zy}_i}^{(k-1)}(\theta)]_j = \sum_{j_2=1}^N \dots \sum_{j_q=1}^N u_{j_2 i}^{(k-1)*} \dots u_{j_q i}^{(k-q+1)*} [\mathcal{C}_q^{\mathbf{z}}(\theta)]_{j, j_2, \dots, j_q}$$
 - (a) Obtain the statistics $(1 \leq i \leq P)$

$$\mathbf{M}_i^{(k-1)} = \sum_{\theta \in \Theta} w_{\theta} \mathbf{c}_{\mathbf{zy}_i}^{(k-1)}(\theta) (\mathbf{c}_{\mathbf{zy}_i}^{(k-1)}(\theta))^H,$$

$$\bar{\mathbf{C}}_{\mathbf{zy}}^{(k-1)} = [\mathbf{M}_1^{(k-1)} \mathbf{u}_1^{(k-q)}, \dots, \mathbf{M}_P^{(k-1)} \mathbf{u}_P^{(k-q)}]$$
 - (b) Ascend in the contrast using 1. or 2.
$$\left\{ \begin{array}{l} 1. \text{ Simultaneous approach:} \\ \quad [\mathbf{V}_L, \Sigma_{P \times P}, \mathbf{V}_R] = \text{svd}(\bar{\mathbf{C}}_{\mathbf{zy}}^{(k-1)}, 0); \\ \quad \mathbf{U}^{(k)} = \mathbf{V}_L \mathbf{V}_R^H; \\ 2. \text{ Hierarchical approach:} \\ \quad [\mathbf{U}^{(k)}, \mathbf{R}_{P \times P}] = \text{qr}(\bar{\mathbf{C}}_{\mathbf{zy}}^{(k-1)}, 0); \\ \quad \text{Sort the columns of } \mathbf{U}^{(k)}, \dots, \mathbf{U}^{(k-q+1)}, \\ \quad \text{according to the ordering in (24).} \end{array} \right.$$
6. IF projection, $\mathbf{U}^{(k-1)}, \dots, \mathbf{U}^{(k-q+1)} = \mathbf{U}^{(k)}$; END;
7. Estimate P independent components:
 $\mathbf{y}^{(k)}(t) = \mathbf{U}^{(k)H} \mathbf{z}(t)$;
8. IF Convergence STOP
ELSE $k = k + 1$; RETURN TO 5

The implementation¹ of the Thin-ICA algorithm is summarized in table 1. We set $P = 4$ and run the simultaneous version of the algorithm in one hundred random experiments. As can be observed in figure 2, in all the cases we extracted in a few iterations the desired subset of sources.

6. CONCLUSIONS

By extending a previous result in [4] we have suggested a robust contrast function for the extraction of a subset of desired independent components, which consists in the maximization of the weighted least-squares low-rank fit to a set of cross-cumulant tensors. In order to optimize this contrast function, we have proposed the Thin-ICA algorithm, and two different implementations based on the thin-SVD and thin-QR factorizations. This algorithm allows the simultaneous extraction of subsets of independent components from the mixture, and provides one possible robust extension of the high-order power method and Fast-ICA algorithms.

¹The thin Singular Value Decomposition and the thin QR decomposition both have, for $P \ll N$, a computational complexity of $O(NP^2)$ flops. An efficient implementation of them can be found in the MatLab commands $\text{svd}(\cdot, 0)$ and $\text{qr}(\cdot, 0)$.

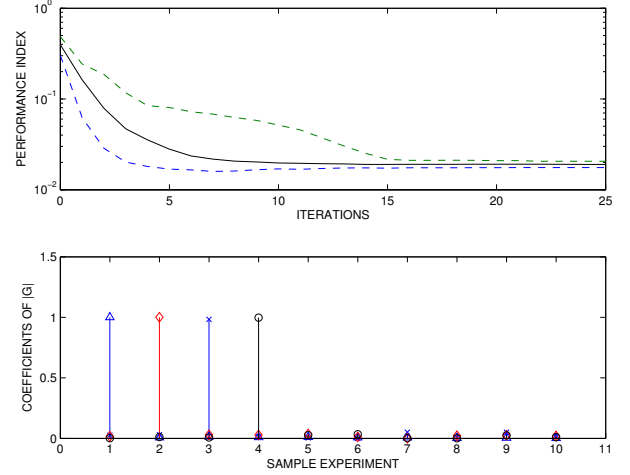


Figure 2: Lower figure presents the rows \mathbf{g}_i^H of the resulting 4×10 mixing matrix $\mathbf{G}^H = \mathbf{U}^H \mathbf{W} \mathbf{A}$ after the convergence, in one sample experiment. Upper figure shows the performance index $P_{index}(\mathbf{G}) = (P(N-1))^{-1} \sum_{i=1}^P (\|\mathbf{g}_i\|_1 / \|\mathbf{g}_i\|_\infty - 1)$ versus iterations. Continuous line is the median curve of convergence for 100 experiments, the dashed lines denote the 5th and 95th percentiles.

Acknowledgements

Part of this research was supported by the CICYT project (TIC2001-0751-C04-04) of the Government of Spain, the Brain Science Institute of Riken (Japan), the Research Council K.U.Leuven, Belgium (GOA-MEFISTO-666), the Flemish Government (F.W.O. project G.0240.99, F.W.O. Research Communities ICCoS and ANMMM) and the Belgian Federal Government (IUAP V-22).

REFERENCES

- [1] P. Comon, "Independent component analysis, a new concept?," *Signal Proc.*, vol. 3(36), pp. 287–314, 1994.
- [2] J.-F. Cardoso, A. Solumiac, "Blind beamforming for non-Gaussian signals," *IEE Proceedings-F*, vol. 140(6), pp. 362–370, 1993.
- [3] A. Belouchrani, K. Abel-Meraim, J.-F. Cardoso, and E. Moulines, "A blind source separation technique using second-order statistics," *IEEE Trans. on Signal Processing*, vol. 45(2), pp. 434–444, 1997.
- [4] L. De Lathauwer, B. De-Moor, and J. Vandewalle, "On the Best Rank-1 and Rank- (R_1, R_2, \dots, R_N) Approximation of Higher-order Tensors," *SIAM J. Matrix Anal. Appl.*, vol. 21(4), pp. 1324–1342, 2000.
- [5] S. Cruces, A. Cichocki, S.-i. Amari, "On a new blind signal extraction algorithm: different criteria and stability analysis," *IEEE Signal Proc. Letters*, vol. 9(8), pp. 233–236, 2002.
- [6] S. Cruces, A. Cichocki, "Combining Blind Source Extraction with Joint Approximate Diagonalization," *Proc. of the Int. Symp. on ICA and BSS*, pp. 463–468, Nara, Japan, 2003.
- [7] E. Kofidis and P.A. Regalia, "On the best rank-1 approximation of higher-order supersymmetric tensors," *SIAM J. Matrix Anal. Appl.*, vol. 23(3), pp. 863–884, 2002.
- [8] A. Hyvärinen and E. Oja, "A fast fixed-point algorithm for independent component analysis," *Neural Computation*, vol. 9, pp. 1483–1492, 1997.
- [9] I. Durán, S. Cruces, "An Application of ICA to Blind DS-CDMA Detection: A Joint Optimization Criterion," *Proc. of the Int. Workshop on Artificial Neural Networks (IWANN)*, vol. II, pp. 305–312, Maó, Spain, June 2003.