Independent Component Analysis of Largely Underdetermined Mixtures [1]

L. De Lathauwer[2], B. De Moor, J. Vandewalle, J.-F. Cardoso

April 2003

Published in:
Proc. Fourth International Symposium on Independent Component Analysis and Blind Signal Separation (ICA 2003), Nara, Japan, April 1–4, 2003

[1]This report is available by anonymous ftp from *ftp.esat.kuleuven.ac.be* in the directory *pub/SISTA/delathauwer/reports/ldl-02-196.ps.gz*

[2]K.U.Leuven, Dept. of Electrical Engineering (ESAT), Research group SCD-SISTA, Kasteelpark Arenberg 10, 3001 Leuven (Heverlee), Belgium, Tel. 32/16/32 11 43, Fax 32/16/32 19 70, WWW: *http://www.esat.kuleuven.ac.be/sista-cosic-docarch*. E-mail: *Lieven.DeLathauwer@esat.kuleuven.ac.be*, *Bart.DeMoor@esat.kuleuven.ac.be*, *Joos.Vandewalle@esat.kuleuven.ac.be*. Lieven De Lathauwer and Jean-François Cardoso hold a permanent research position with the French Centre National de la Recherche Scientifique (C.N.R.S.); Lieven De Lathauwer also holds a honorary research position with the K.U.Leuven. Bart De Moor and Joos Vandewalle are Full Professor with the K.U.Leuven. This work is supported by several institutions: (1) Research Council K.U.Leuven: Concerted Research Action GOA-MEFISTO-666 (Mathematical Engineering for Information and Communication Systems Technology), (2) Flemish Government: the F.W.O. projects G.0240.99 (Multilinear Generalisations of the Singular Value Decomposition and Applications in Signal Processing and System Identification) and G.0256.97 (Numerical Algorithms for Subspace System Identification, Extension to Special Cases), the F.W.O. Research Communities ICCoS (Identification and Control of Complex Systems) and ANMMM (Advanced Numerical Methods for Mathematical Modelling), (3) Belgian Federal Government: Interuniversity Poles of Attraction Programme IUAP IV-02 (1997-2001): Modeling, Identification, Simulation and Control of Complex Systems, and IUAP V-22 (2002-2006): Dynamical Systems and Control: Computation, Identification & Modelling. The scientific responsibility is assumed by the authors.

Abstract

In this paper we derive algebraic means for Independent Component Analysis (ICA) with more sources than sensors. The results are based on the structure of the fourth-order cumulant tensor. We derive bounds on the number of sources that generically guarantee uniqueness of the decomposition. The mixing matrix is computed by means of simultaneous diagonalization or off-diagonalization techniques.

INDEPENDENT COMPONENT ANALYSIS OF LARGELY UNDERDETERMINED MIXTURES

L. De Lathauwer

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

B. De Moor *J. Vandewalle*

KULeuven - SCD (SISTA)
Leuven
Belgium
{demoor,vdwalle}@esat.kuleuven.ac.be

J.-F. Cardoso

ENST - TSI
Paris
France
cardoso@tsi.enst.fr

ABSTRACT

In this paper we derive algebraic means for Independent Component Analysis (ICA) with more sources than sensors. The results are based on the structure of the fourth-order cumulant tensor. We derive bounds on the number of sources that generically guarantee uniqueness of the decomposition. The mixing matrix is computed by means of simultaneous diagonalization or off-diagonalization techniques.

1. INTRODUCTION

The basic ICA model is in this paper denoted as

$$Y = \mathbf{M}X + N, \tag{1}$$

in which the observed vector $Y \in \mathbb{C}^J(\mathbb{R}^J)$, the source vector $X \in \mathbb{C}^R(\mathbb{R}^R)$ and the noise vector $N \in \mathbb{C}^J(\mathbb{R}^J)$ are zero-mean random vectors. The components of X are mutually statistically independent and are assumed to be kurtic. The noise is Gaussian. In this paper, we consider the so-called *underdetermined* or *overcomplete* case, where $J \leqslant R$. We will focus on the estimation of the mixing matrix $\mathbf{M} \in \mathbb{C}^{J\times R}(\mathbb{R}^{J\times R})$; the estimation of the corresponding source values will not be considered.

Section 2 is an obvious but important modification of the technique proposed in [2]. In the latter paper, the solution was found via an Eigenvalue Decomposition (EVD) of a real symmetric matrix; instead, we will now resort to a simultaneous diagonalization, which is more robust. [2] contained a number of statements of which the proof was not explicited; these proofs will also be given in Section 2.

Section 3 is a variant of Section 2 in which the number of sources that can be allowed, is somewhat higher.

The algorithms of Section 2 and 3 work under the condition that the number of sources does not exceed a certain bound. The constraint of Section 3 is further weakened in Section 4. Here, the assessment whether the underlying cumulant tensor decomposition is unique, is based on heuristics. Whilst the algorithms of Section 2 and 3 take the form of a simultaneous diagonalization, Section 4 resorts to a simultaneous off-diagonalization.

2. ALGORITHM 1

Consider $\mathcal{C}^Y = \mathrm{Cum}\{Y, Y^*, Y^*, Y\}$. Due to the multilinearity property of tensors, we have:

$$c^Y_{ijkl} = \sum_{r=1}^{R} \kappa_r m_{ir} m^*_{jr} m^*_{kr} m_{lr}, \tag{2}$$

in which κ_r is the autocumulant of the rth source. This is a decomposition of a super-symmetric fourth-order tensor in a sum of rank-1 terms. If all the kurtosis values are positive, then we can absorb them in the corresponding outer products:

$$c^Y_{ijkl} = \sum_{r=1}^{R} a_{ir} a^*_{jr} a^*_{kr} a_{lr}. \tag{3}$$

If all the kurtosis values are negative, then we process $-\mathcal{C}^Y$, instead of $\mathcal{C}^Y$. In the case where there are terms with a different sign, our derivation can be reformulated in terms of an unknown J-orthogonal, instead of orthogonal, matrix $\mathbf{Q}$.

Associate with $\mathcal{C}^Y$ a matrix-to-matrix mapping as follows:

$$(\mathcal{C}^Y(\mathbf{M}))_{ij} = \sum_{kl} c^Y_{ijkl} m_{kl}. \tag{4}$$

Let this mapping be represented by a matrix $\mathbf{C}^Y \in \mathbb{C}^{J^2\times J^2}$. In terms of $\mathbf{C}^Y$, Eq. (3) can be rewritten as:

$$\mathbf{C}^Y = (\mathbf{A} \odot \mathbf{A}^*) \cdot (\mathbf{A} \odot \mathbf{A}^*)^H, \tag{5}$$

in which $\odot$ is the Katri-Rao or column-wise Kronecker product.

We now have:

Theorem 1. A complex $(J \times J \times J \times J)$ tensor $\mathcal{T}$, exhibiting $t_{klij} = t^*_{ijkl}$ and $t_{jilk} = t^*_{ijkl}$, can be eigen-decomposed as

$$t_{ijkl} = \sum_p^P \lambda_p (\mathbf{E}_p)_{ij} (\mathbf{E}_p)^*_{kl}, \qquad (6)$$

in which the matrices $\mathbf{E}_p$ are mutually orthonormal and Hermitean, λ_p are real and P is the rank of the matrix-to-matrix mapping defined in analogy with Eq. (4) (P can be as big as I^2).

Proof: Due to the first symmetry, the EVD of $\mathcal{T}$, considered as a matrix-to-matrix mapping, takes the form of Eq. (6), in which $\mathbf{E}_p$ are mutually orthonormal and λ_p are real. Hence the tensor $\mathcal{S}$, defined by $s_{ijkl} = t_{jilk}$, is given by

$$s_{ijkl} = \sum_p^P \lambda_p (\mathbf{E}_p)_{ji} (\mathbf{E}_p)^*_{lk}.$$

Also, the EVD of $\mathcal{T}^*$ is given by

$$t^*_{ijkl} = \sum_p^P \lambda_p (\mathbf{E}_p)^*_{ij} (\mathbf{E}_p)_{kl}.$$

Because of the second symmetry, $\mathcal{S} = \mathcal{T}^*$; assuming that all the eigenvalues are mutually different, we obtain that the projectors corresponding to the same eigenvalue should be the same:

$$e_{ji} e^*_{lk} = e^*_{ij} e_{kl}, \qquad (7)$$

in which we have dropped the index p for notational convenience. If λ is a multiple eigenvalue, then its rank-1 projectors can be chosen the same. Hence, the projectors satisfy the same symmetries as $\mathcal{T}$ itself.

Now we have to verify what this means in terms of $\mathbf{E}$. The projector does not change when $\mathbf{E}$ is multiplied by a unit-modulus factor. Call the result $\mathbf{E}'$. If some diagonal entry of $\mathbf{E}$, say e_{pp}, is non-vanishing, then we choose the unit-modulus factor such that ${e'}_{pp}$ is real. Since we have ${e'}_{ji} {e'}^*_{pp} = {e'}^*_{ij} {e'}_{pp}$ for all i, j, $\mathbf{E}'$ is Hermitean. If all the diagonal entries of $\mathbf{E}$ are zero, then we proceed as follows. First remark that Eq. (7) implies that all $|e_{ij}| = |e_{ji}|$. If, say, $e_{pq} \neq 0$, then we multiply $\mathbf{E}$ by a unit-modulus factor such that ${e'}_{pq} = {e'}^*_{qp}$. Since we have ${e'}_{ji} {e'}^*_{qp} = {e'}^*_{ij} {e'}_{pq}$, $\mathbf{E}'$ is Hermitean. ∎

We emphasize that the eigenmatrices are not Hermitean by default, as $\mathbf{E}_p$ may be multiplied by any unit-modulus factor; multiplication by j even yields anti-Hermitean eigenmatrices. The equivalent of Theor. 1 for real-valued tensors is trivial to prove; it involves real symmetric eigen-matrices and P is bounded by $J(J+1)/2$.

Recall that $\mathbf{C}^Y$ is, by assumption, positive (semi)definite. According to Theor. 1, and under the assumptions that $R \leqslant I^2$ and that the columns of $\mathbf{M}$ are as linearly independent as possible, it can be decomposed as

$$\mathbf{C}^Y = \mathbf{H} \cdot \mathbf{H}^H, \qquad (8)$$

in which $\mathbf{H} \in \mathbb{C}^{I^2 \times R}$ is full rank and has columns that correspond to Hermitean matrices $\mathbf{H_1}, \ldots, \mathbf{H_R}$. Comparison of Eqs. (5) and (8) shows that

$$\mathbf{A} \odot \mathbf{A}^* = \mathbf{H} \cdot \mathbf{Q},$$

in which $\mathbf{Q}$ is a real $(R \times R)$ orthogonal matrix (the fact that this matrix is real, and not complex unitary, maintains the Hermitean symmetry). The task is now to find $\mathbf{Q}$ such that the columns of $\mathbf{H} \cdot \mathbf{Q}$ correspond to rank-1 matrices.

To this end, the following theorem was proposed in [2]:

Theorem 2. Define a mapping $\Phi : \mathbb{C}^{J \times J} \times \mathbb{C}^{J \times J} \rightarrow \mathbb{C}^{J \times J \times J \times J}$ by

$$(\Phi(\mathbf{X}, \mathbf{Y}))_{ijkl} = x_{ij} y^*_{kl} + y_{ij} x^*_{kl} - x_{ik} y^*_{jl} - y_{ik} x^*_{jl}. \qquad (9)$$

Then we have that $\Phi(\mathbf{X}, \mathbf{X}) = 0$ if and only if $\mathbf{X}$ is at most rank-1.

Proof: We will give an alternative proof that is not explicitly based on the exploitation of symmetries. The "if" part is obvious. For the "only if" part, let the Singular Value Decomposition (SVD) of $\mathbf{X}$ be given by $\mathbf{U} \cdot \mathbf{\Sigma} \cdot \mathbf{V}^H$. We have:

$$\begin{aligned} x_{ij} x^*_{kl} &= \sum_{rs} \sigma_r \sigma_s u_{ir} v^*_{jr} u^*_{ks} v_{ls} \\ x_{ik} x^*_{jl} &= \sum_{rs} \sigma_r \sigma_s u_{ir} v^*_{js} u^*_{kr} v_{ls}. \end{aligned}$$

Rank-1 terms corresponding to the same $r = s$ cancel out in Eq. (9). However, due to the orthogonality of $\mathbf{U}$ and $\mathbf{V}$, the other terms are linearly independent. So we must have that $\sigma_r \sigma_s = 0$ whenever $r \neq s$; hence, $\mathbf{\Sigma}$ is at most rank-1. The proof still holds when $\mathbf{X}$ is Hermitean. ∎

Denote the $(J \times J)$ matrices represented by the columns of $\mathbf{H}$ as $\mathbf{H_1}, \ldots, \mathbf{H_R}$ and let $\Phi_{\mathbf{st}} = \Phi(\mathbf{H_s}, \mathbf{H_t})$. Let the columns of $\mathbf{A}$ be given by $\{A_p\}$. Due to the bilinearity of Φ, we have

$$\Phi_{\mathbf{st}} = \sum_{pq} (\mathbf{Q}^T)_{ps} (\mathbf{Q}^T)_{qt} \Phi(A_p A_p^H, A_q A_q^H). \qquad (10)$$

Now assume there exists a symmetric matrix $\mathbf{W}$ that satisfies

$$\sum_{st} w_{st} \Phi_{\mathbf{st}} = 0. \qquad (11)$$

If the tensors $\{\Phi(A_p A_p^H, A_q A_q^H)\}_{p<q}$ are linearly independent, then substitution of (10) in (11) shows that

$$\sum_{st} w_{st} (\mathbf{Q}^T)_{ps} (\mathbf{Q}^T)_{qt} = (\mathbf{\Lambda})_{pq} \delta_{pq} \qquad \forall p, q \qquad (12)$$

in which δ is the Kronecker delta. Hence, $\mathbf{Q}$ may be obtained from the EVD

$$\mathbf{W} = \mathbf{Q} \cdot \mathbf{\Lambda} \cdot \mathbf{Q}^T. \tag{13}$$

Actually, and this forms an improvement of [2], one can see that *any* diagonal matrix $\mathbf{\Lambda}$ generates a matrix $\mathbf{W}$ that satisfies Eq. (11). Hence, if the tensors $\{\Phi(A_p A_p^H, A_q A_q^H)\}_{p<q}$ are linearly independent, these matrices form an R-dimensional subspace of the symmetric $(R \times R)$ matrices. One can always find a real basis for this subspace. Let $\{\mathbf{W_r}\}$ represent such a basis. Then $\mathbf{Q}$ may be determined in a more robust way from a simultaneous real symmetric EVD:

$$\begin{aligned} \mathbf{W_1} &= \mathbf{Q} \cdot \mathbf{\Lambda_1} \cdot \mathbf{Q}^T \\ &\vdots \\ \mathbf{W_R} &= \mathbf{Q} \cdot \mathbf{\Lambda_R} \cdot \mathbf{Q}^T. \end{aligned} \tag{14}$$

The solution may be found by means of the technique developed in [3].

The question is now under which condition on R linear independence of $\{\Phi(A_p A_p^H, A_q A_q^H)\}_{p<q}$ can generically be guaranteed. We call a property "generic" when it holds everywhere, except for a set of Lebesgue measure 0. We have the following theorem:

Theorem 3. In the complex case, linear independence of $\{\Phi(A_p A_p^H, A_q A_q^H)\}_{p<q}$ is generically guaranteed if $R(R-1) \leqslant J^2(J-1)^2/2$. In the real case, R is bounded by R_1, given in Table 1.

Proof: $\Phi(A_p A_p^H, A_q A_q^H)$ can be represented in a vector format as follows:

$$\begin{aligned} & A_p \otimes A_q \otimes (A_p \otimes A_q - A_q \otimes A_p)^* \\ & + A_q \otimes A_p \otimes (A_q \otimes A_p - A_p \otimes A_q)^* \\ = \; & (A_p \otimes A_q - A_q \otimes A_p) \otimes (A_p \otimes A_q - A_q \otimes A_p)^* \\ = \; & [(\mathbf{I}-\mathbf{P})(A_p \otimes A_q)] \otimes [(\mathbf{I}-\mathbf{P})(A_p \otimes A_q)]^* \\ = \; & [(\mathbf{I}-\mathbf{P}) \otimes (\mathbf{I}-\mathbf{P})][A_p \otimes A_q \otimes A_p^* \otimes A_q^*], \end{aligned} \tag{15}$$

in which $\mathbf{P}$ is a specific permutation matrix. $\mathbf{P}$ is such that $\mathbf{I}-\mathbf{P}$ has rank $(J^2-J)/2$. In the complex case, the vectors $A_p \otimes A_q \otimes A_p^* \otimes A_q^*$ are generically as linearly independent as possible (second-order analogy: $A_p \otimes A_p^*$ are also as linearly independent as possible). Hence, the condition is that $R(R-1)/2 \leqslant (\text{rank}(\mathbf{I}-\mathbf{P}))^2 = J^2(J-1)^2/4$.

The real case is more difficult. Eq. (15) seems to suggest that $\Phi(A_p A_p^H, A_q A_q^H)$ can span a subspace of dimension $(\text{rank}(\mathbf{I}-\mathbf{P}))(\text{rank}(\mathbf{I}-\mathbf{P})+1)/2 = N(N-1)(N^2-N+2)/8$, but in fact there are a number of symmetries that decrease this dimensionality. For instance, we have that

$$\begin{aligned} & (a_{ip}a_{jq} - a_{jp}a_{iq})(a_{kp}a_{lq} - a_{lp}a_{kq}) \\ & -(a_{ip}a_{kq} - a_{kp}a_{iq})(a_{jp}a_{lq} - a_{lp}a_{jq}) \\ & +(a_{ip}a_{lq} - a_{lp}a_{iq})(a_{jp}a_{kq} - a_{kp}a_{jq}) = 0, \end{aligned}$$

J	2	3	4	5	6	7	8
R_1	2	4	6	10	15	20	26
R_2	2	5	8	12	18	24	31
R_{gen}	3	6	10	15	22	30	42
D	1	3	5	5	6	0	6

Table 1. Real case: the maximum number of sources that can be allowed in Alg. 1 (R_1) and in Alg. 2 (R_2); the generic rank of a super-symmetric fourth-order tensor (R_{gen}); the number of degrees of freedom (D) for a decomposition of a generic fourth-order tensor in rank-1 terms.

for arbitrary i, j, k, l. Hence, for $J \geqslant 4$, R is at least bounded by

$$\frac{R(R-1)}{2} \leqslant \frac{N(N-1)(N^2-N+2)}{8} - \frac{J!}{(J-4)!\,4!},$$

in which the last term corresponds to the number of ways in which one can choose 4 indices out of J. From $J = 7$ unwards, other symmetries additionally start to play a role. Table 1 lists the maximal value of R till $J = 8$. The values for $J = 7, 8$ have simply be obtained by checking the dimension of the subspace that is generically spanned by $\Phi(A_p A_p^T, A_q A_q^T)$. ■

3. ALGORITHM 2

Actually, it is not mandatory that the rank-1 detecting bilinear mapping is symmetric in its arguments, like the mapping Φ in the previous paragraph. We might as well resort to an unsymmetric mapping:

Theorem 4. Define a mapping $\Psi : \mathbb{C}^{J \times J} \times \mathbb{C}^{J \times J} \rightarrow \mathbb{C}^{J \times J \times J \times J}$ by

$$(\Psi(\mathbf{X}, \mathbf{Y}))_{ijkl} = x_{ij}y_{kl}^* - x_{ik}y_{jl}^*. \tag{16}$$

Then we have that $\Psi(\mathbf{X}, \mathbf{X}) = 0$ if and only if $\mathbf{X}$ is at most rank-1.

Proof: Since $\Phi(\mathbf{X}, \mathbf{X}) = 2\,\Psi(\mathbf{X}, \mathbf{X})$, Theor. 4 and Theor. 2 are equivalent. ■

One may then repeat the derivation of the previous section to find that the orthogonal matrix $\mathbf{Q}$ may be obtained from a simultaneous EVD of R real symmetric $(R \times R)$ matrices. The advantage is that working with Ψ imposes a weaker constraint on the number of sources that can be allowed:

Theorem 5. In the complex case, linear independence of $\{\Psi(A_p A_p^H, A_q A_q^H)\}_{p \neq q}$ is generically guaranteed if $R(R-1) \leqslant J^3(J-1)/2$. In the real case, R is bounded by R_2, given in Table 1.

Proof: $\Phi(A_p A_p^H, A_q A_q^H)$ can now be represented in a vector format as $[\mathbf{I} \otimes (\mathbf{I} - \mathbf{P})][A_p \otimes A_q \otimes A_p^* \otimes A_q^*]$, in which $\mathbf{P}$ is the same permutation matrix as in the proof of Theor. 3. Hence, in the complex case, the condition is $R(R-1) \leqslant [J^2][J(J-1)/2]$.

In the real case, we have symmetries like, e.g.,

$$\begin{gathered} a_{ip}a_{kq}(a_{ip}a_{jq} - a_{jp}a_{iq}) - a_{ip}a_{jq}(a_{ip}a_{kq} - a_{kp}a_{iq}) \\ -a_{ip}a_{iq}(a_{kp}a_{jq} - a_{jp}a_{kq}) = 0 \\ a_{ip}a_{iq}(a_{ip}a_{kq} - a_{kp}a_{iq}) - a_{ip}a_{iq}(a_{kp}a_{iq} - a_{ip}a_{kq}) = 0, \end{gathered}$$

which lead to the bounds on R given in Table 1. ■

4. ALGORITHM 3

In the two previous sections, the algorithms worked under the assumption of a certain bound on the number of sources. The fact that (14) is overdetermined, indicates that the decomposition of $\mathcal{C}^Y$ is unique under weaker constraints. Table 1 lists how many rank-1 terms are generically needed to build a real fourth-order tensor (the generic *rank*), and the corresponding dimension of the solution set [4]. This table shows, e.g., that generically a real $(7 \times 7 \times 7 \times 7)$-tensor can be decomposed in 30 rank-1 terms and that there are only a discrete number of ways to do so. Moreover, we may expect that, for $(7 \times 7 \times 7 \times 7)$-tensors that can be decomposed in 25, 26, ..., 29 terms, the decomposition is generically unique, and this supposition is indeed stronger than the results that had been derived in Sections 2 and 3.

Let us define a third rank-1 detecting mapping:

Theorem 6. Let $\mathcal{H}$ be the set of $(J \times J)$ Hermitean matrices. Define a mapping $\Gamma : \mathcal{H} \times \mathcal{H} \to \mathcal{H}$ by

$$\Gamma(\mathbf{X}, \mathbf{Y}) = \mathbf{XY} - \text{trace}(\mathbf{X})\mathbf{Y} + \mathbf{YX} - \text{trace}(\mathbf{Y})\mathbf{X}. \quad (17)$$

Then we have that $\Gamma(\mathbf{X}, \mathbf{X}) = 0$ if and only if $\mathbf{X}$ is at most rank-1.

Proof: It is easy to check that $\Gamma(\mathbf{X}, \mathbf{X}) = 0$ if $\mathbf{X}$ is rank-1. For the "only if" part, let the EVD of $\mathbf{X}$ be given by $\mathbf{U} \cdot \mathbf{\Lambda} \cdot \mathbf{U}^H$. $\Gamma(\mathbf{X}, \mathbf{X}) = 0$ iff

$$\begin{aligned} \mathbf{U} \cdot \mathbf{\Lambda}^2 \cdot \mathbf{U}^H &= \text{trace}(\mathbf{\Lambda})\, \mathbf{U} \cdot \mathbf{\Lambda} \cdot \mathbf{U}^H \\ \mathbf{\Lambda}^2 &= \text{trace}(\mathbf{\Lambda})\, \mathbf{\Lambda} \\ \mathbf{\Lambda} &= \text{trace}(\mathbf{\Lambda})\, \mathbf{I}, \end{aligned}$$

and hence at most one eigenvalue can be different from zero. ■

If $R \leqslant J^2$ (complex case) or $R \leqslant J(J+1)/2$ (real case), then R can still be estimated as the rank of $\mathbf{C}^Y$. Furthermore, Theor. 6 implies that, for any r, we have

$$\sum_{st} q_{sr} q_{tr} \Gamma(\mathbf{H_s}, \mathbf{H_t}) = 0. \quad (18)$$

If we define $\mathbf{B_{ij}} \in \mathbb{C}^{R \times R}$ by

$$(\mathbf{B_{ij}})_{st} = (\Gamma(\mathbf{H_s}, \mathbf{H_t}))_{ij},$$

then Eq. (18) can be rewritten as:

$$\text{diag}(\mathbf{Q}^T \cdot \mathbf{B_{ij}} \cdot \mathbf{Q}) = 0 \qquad \forall i, j \quad (19)$$

This simultaneous off-diagonalization problem can be solved by means of a Jacobi iteration using the expressions derived in [3]; one simply chooses in each step the Jacobi rotation that minimizes (instead of maximizes) the energy on the diagonals. A simultaneous off-diagonalization problem also appeared in [1].

Note that, in the real case, due to the symmetry of the mapping $\Gamma(\mathbf{H_s}, \mathbf{H_t})$, $\mathbf{B_{ij}} = \mathbf{B_{ji}}$, such that we only have to consider $J(J+1)/2$ real matrices in the off-diagonalization procedure. In the complex case, we have that $\mathbf{B_{ij}} = \mathbf{B_{ji}}^*$, and hence one can work with J^2 real matrices.

Using this algorithm, we have, e.g., been able to identify mixtures of 13 sources observed by 5 sensors (cf. Table 1).

Remark 1. One could also have started from the mapping Φ defined in Theor. 2. However, this would have been less economic in the number of matrices that have to be off-diagonalized.

Remark 2. In principle, R need not be bounded by J^2 (complex case) or $J(J+1)/2$ (real case). One then merely looks for a real row-wise orthonormal matrix $\mathbf{Q} \in \mathbb{R}^{J^2 \times R}$ (or $\mathbb{R}^{J(J+1)/2 \times R}$). However, such problems are by definition very ill-conditioned. In our simulations we observed that the convergence speed becomes too small to reach the optimum.

5. CONCLUSION

Underdetermined ICA involving J observations and $R > J$ sources can be achieved by resorting to the property that the rank of a fourth-order $(J \times J \times J \times J)$ tensor can be much higher than J. In this paper we showed that, in the complex case, the solution can be obtained from a simultaneous EVD of R real symmetric $(R \times R)$ matrices, if $R(R-1) \leqslant J^2(J-1)^2/2$ (Alg. 1) or $R(R-1) \leqslant J^3(J-1)/2$ (Alg. 2). The bounds on R in the real case are given in Table 1, for $J \leqslant 8$. Alg. 1 is a robust version of the technique derived in [2], in which a single EVD is replaced by a simultaneous EVD. The constraint on R is even further weakened in Alg. 3, which is based on a simultaneous off-diagonalization of $J(J+1)/2$ (real case) or J^2 (complex case) real symmetric $(R \times R)$ matrices.

Acknowledgements

The authors would like to thank Prof. Alle-Jan van der Veen (T.U.Delft, The Netherlands) for bringing the rank-1 revealing mapping Γ, used in Section 4, to their attention.

Lieven De Lathauwer and Jean-François Cardoso hold a permanent research position with the French Centre National de la Recherche Scientifique (C.N.R.S.); Lieven De Lathauwer also holds a honorary research position with the K.U.Leuven. Bart De Moor and Joos Vandewalle are Full Professor with the K.U.Leuven.

This work is supported by several institutions:

1. Research Council K.U.Leuven: Concerted Research Action GOA-MEFISTO-666 (Mathematical Engineering for Information and Communication Systems Technology),
2. Flemish Government: the F.W.O. projects G.0240.99 (Multilinear Generalisations of the Singular Value Decomposition and Applications in Signal Processing and System Identification) and G.0256.97 (Numerical Algorithms for Subspace System Identification, Extension to Special Cases), the F.W.O. Research Communities ICCoS (Identification and Control of Complex Systems) and ANMMM (Advanced Numerical Methods for Mathematical Modelling),
3. Belgian Federal Government: Interuniversity Poles of Attraction Programme IUAP IV-02 (1997-2001): Modeling, Identification, Simulation and Control of Complex Systems, and IUAP V-22 (2002-2006): Dynamical Systems and Control: Computation, Identification & Modelling.

The scientific responsibility is assumed by the authors.

6. REFERENCES

[1] K. Abed-Meraim, M. G. Amin, A. M. Zoubir, “Joint Anti-Diagonalization for Blind Source Separation”, *Proc. ICASSP-01*, May 2001.

[2] J.-F. Cardoso, “Super-Symmetric Decomposition of the Fourth-Order Cumulant Tensor. Blind Identification of More Sources than Sensors.” *Proc. ICASSP-91*, Vol. 5, pp. 3109–3112.

[3] J.-F. Cardoso, A. Souloumiac, “Blind Beamforming for Non-Gaussian Signals”, *IEE Proceedings-F*, Vol. 140, No. 6, 1994, pp. 362–370.

[4] P. Comon, B. Mourrain, “Decomposition of Quantics in Sums of Powers of Linear Forms”, *Signal Process.*, Special Issue on *Higher Order Statistics*, Vol. 53, 1996, pp. 93–108.