**Katholieke Universiteit Leuven**

Departement Elektrotechniek ESAT-SISTA/TR 2002-77

# Jacobi Algorithm for Joint Block Diagonalization in Blind Identification[1]

L. De Lathauwer, C. Févotte, B. De Moor and J. Vandewalle[2]

May 2002

Published in:
*Proc. 23rd Symp. on Information Theory in the Benelux*,
May 29-31, 2002, Louvain-la-Neuve, Belgium, pp. 155-162.

[1]This report is available by anonymous ftp from *ftp.esat.kuleuven.ac.be* in the directory *pub/SISTA/delathauwer/reports/ldl-02-77.ps.Z*

[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 04, 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.* C. Févotte holds a Ph.D. research position with the French Ministry of Research and is affiliated with IRCCyN, UMR 6597, 1 rue de la Noé, BP 92101, Nantes, France. Tel: +33-(0)2.40.37.69.24; fax: +33-(0)2.40.37.69.30; e-mail: *cedric.fevotte@irccyn.ec-nantes.fr*; URL: *http://www.irccyn.ec-nantes.fr.* Lieven De Lathauwer holds a permanent research position with the French Centre National de la Recherche Scientifique (C.N.R.S.); he also holds a honorary post-doctoral research mandate with the Fund for Scientific Research-Flanders (F.W.O.), affiliated 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, (2) Flemish Government: the F.W.O. projects G.0240.99 and G.0256.97, the F.W.O. Research Communities ICCoS and ANMMM, (3) Belgian Federal Government: Interuniversity Poles of Attraction Programme IUAP IV-02 and IUAP V-22. The scientific responsibility is assumed by its authors.

**Abstract**

In this paper we generalize the well-known SOBI technique for blind source separation to convolutive mixtures. Our algorithm is based on joint block diagonalization of a set of covariance matrices by means of a unitary similarity transformation. For the latter problem we derive a Jacobi procedure.

# Jacobi Algorithm for Joint Block Diagonalization in Blind Identification

L. De Lathauwer C. Févotte B. De Moor J. Vandewalle*

**Abstract**

In this paper we generalize the well-known SOBI technique for blind source separation to convolutive mixtures. Our algorithm is based on joint block diagonalization of a set of covariance matrices by means of a unitary similarity transformation. For the latter problem we derive a Jacobi procedure.

## 1 Introduction

In this paper we show that, under some constraints, blind separation of the inputs of a MIMO FIR system can be achieved by resorting only to the second-order statistics of the output, if we know that the inputs are spatially white but individually correlated in time. This is a convolutive equivalent of the well-known SOBI approach to the blind separation of instantaneous linear mixtures [2]. SOBI relies on the joint diagonalization of a set of covariance matrices by means of a unitary similarity transformation; for this problem one can resort to a Jacobi iteration. Convolutive mixtures could be identified by means of joint block diagonalization of a set of covariance matrices. We will derive a Jacobi algorithm for the latter problem.

In Section 2 we will define our data model. Section 3 outlines the computational procedure. In Section 4 we discuss in detail how the joint block diagonalization can be realized. The real case is covered in Section 4.1, the complex case in Section 4.2.

## 2 Model

We consider the following discrete-time MIMO model:

$$X(t) = \sum_{l=0}^{L} \mathbf{A}(l) S(t-l) + N(t),$$

where:

- $X(t) = [x_1(t), \ldots, x_m(t)]^T$ is the vector of observations.
- $S(t) = [s_1(t), \ldots, s_n(t)]^T$ contains unknown zero-mean stationary sources, mutually independent but individually correlated in time. We assume that $n < m$.
- $\mathbf{A}(l) = [a_{ij}(l)]$, $l = 0, \ldots, L$ are the unknown $(m \times n)$ Markov parameters.

*L. De Lathauwer is with the research group ETIS, UPRES-A 8051, 6 avenue du Ponceau, BP 44, F 95014 Cergy-Pontoise Cedex, France. Tel: +33-(0)1-30.73.66.10; fax: +33-(0)1-30.73.62.82; e-mail: delathau@ensea.fr; URL: http://www-rech.ensea.fr. C. Févotte is with IRCCyN, UMR 6597, 1 rue de la Noé, BP 92101, Nantes, France. Tel: +33-(0)2.40.37.69.24; fax: +33-(0)2.40.37.69.30; e-mail: cedric.fevotte@irccyn.ec-nantes.fr; URL: http://www.irccyn.ec-nantes.fr. B. De Moor and J. Vandewalle are with the group SCD-SISTA of the E.E. Dept. (ESAT) of the K.U.Leuven, Kard. Mercierlaan 94, B-3001 Leuven (Heverlee), Belgium. Tel: +32-(0)16-32.17.09; fax: +32-(0)16-32.19.70; e-mail: {demoor,vdwalle}@esat.kuleuven.ac.be; URL: http://www.esat.kuleuven.ac.be/sista-cosic-docarch.

- $N(t)$ is i.i.d. white noise, independent from the sources, with variance $\sigma_N^2$.

Let $L'$ be an integer such that $mL' \geqslant n(L+L')$. We define:

$$\begin{array}{rcll} S_j(t) & = & [s_j(t), \ldots, s_j(t-(L+L')+1)]^T, & j=1,\ldots,n, \\ X_i(t) & = & [x_i(t), \ldots, x_i(t-L'+1)]^T, & i=1,\ldots,m, \\ N_i(t) & = & [n_i(t), \ldots, n_i(t-L'+1)]^T, & i=1,\ldots,m, \\ \mathbf{S}(t) & = & [S_1(t)^T, \ldots, S_n(t)^T]^T, & \\ \mathbf{X}(t) & = & [X_1(t)^T, \ldots, X_m(t)^T]^T, & \\ \mathbf{N}(t) & = & [N_1(t)^T, \ldots, N_m(t)^T]^T. & \end{array}$$

Let, for $i=1,\ldots,m, j=1,\ldots,n$, $\mathbf{A}_{ij} \in \mathbb{C}^{L' \times (L+L')}$ be given by

$$\mathbf{A}_{ij} = \begin{bmatrix} a_{ij}(0) & \cdots & a_{ij}(L) & \cdots & 0 \\ & \ddots & & \ddots & \\ 0 & \cdots & a_{ij}(0) & \cdots & a_{ij}(L) \end{bmatrix},$$

and define the matrix $\mathbf{A} \in \mathbb{C}^{mL' \times n(L+L')}$ as

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \cdots & \mathbf{A}_{1n} \\ \vdots & \ddots & \vdots \\ \mathbf{A}_{m1} & \cdots & \mathbf{A}_{mn} \end{bmatrix}.$$

Then we have:

$$\mathbf{X}(t) = \mathbf{A}\mathbf{S}(t) + \mathbf{N}(t), \tag{1}$$

in which we assume that $\mathbf{A}$ is full rank. This is the typical equation for blind separation of instantaneous mixtures, with the individual sources correlated in time. However, a major difference with [2] is the fact that the sources in Eq. (1) are only partially spatially white, because time-shifted versions of $s_j(t)$ may be correlated ($j=1,\ldots,n$). Formally, the covariance matrices $\mathbf{R}_\mathbf{S}(\tau) = E\{\mathbf{S}(t)\mathbf{S}(t+\tau)^H\}$ are block-diagonal, instead of diagonal, with $((L+L') \times (L+L'))$ blocks given by $\mathbf{R}_{S_j}(\tau) = E\{S_j(t)S_j(t+\tau)^H\}$ ($j=1,\ldots,n$). As a consequence, we will only be able to separate filtered versions of the sources. If necessary, the remaining SIMO filters will have to be estimated afterwards.

# 3 Procedure

We have

$$\mathbf{R}_\mathbf{X}(\tau) = \mathbf{A} \cdot \mathbf{R}_\mathbf{S}(\tau) \cdot \mathbf{A}^H + \sigma_N^2 \tilde{\mathbf{I}}(\tau), \tag{2}$$

in which $\tilde{\mathbf{I}}(\tau)$ contains only zero entries, except for ones on the diagonal ($\tau = 0$), the $\tau$th subdiagonal ($\tau > 0$) or superdiagonal ($\tau < 0$). The noise variance $\sigma_N^2$ can be estimated as the mean of the $mL' - n(L+L')$ smallest eigenvalues of $\mathbf{R}_\mathbf{X}(0)$. In what follows, we will only work with the noise-compensated matrices

$$\tilde{\mathbf{R}}_\mathbf{X}(\tau) = \mathbf{R}_\mathbf{X}(\tau) - \sigma_N^2 \tilde{\mathbf{I}}(\tau).$$

If the Eigenvalue Decomposition (EVD) of $\tilde{\mathbf{R}}_\mathbf{X}(0)$ is given by

$$\tilde{\mathbf{R}}_\mathbf{X}(0) = \mathbf{U} \cdot \mathbf{\Lambda}^2 \cdot \mathbf{U}^H, \tag{3}$$

in which $\mathbf{U} \in \mathbb{C}^{mL' \times n(L+L')}$ has orthonormal columns and $\mathbf{\Lambda} \in \mathbb{R}^{n(L+L') \times n(L+L')}$ is diagonal, and if $\mathbf{R}_\mathbf{S}(0)^{\frac{1}{2}}$ represents an arbitrary square root of $\mathbf{R}_\mathbf{S}(0)$, then we have

$$\mathbf{A} \cdot \mathbf{R}_\mathbf{S}(0)^{\frac{1}{2}} = \mathbf{U} \cdot \mathbf{\Lambda} \cdot \mathbf{V}, \tag{4}$$

for some unitary matrix $\mathbf{V} \in \mathbb{C}^{n(L+L') \times n(L+L')}$. Combination with (2) yields, for different values of $\tau \neq 0$,

$$\mathbf{\Lambda}^{-1} \cdot \mathbf{U}^H \cdot \tilde{\mathbf{R}}_\mathbf{X}(\tau) \cdot \mathbf{U} \cdot \mathbf{\Lambda}^{-1} = \mathbf{V} \cdot \Omega(\tau) \cdot \mathbf{V}^H, \tag{5}$$

in which $\Omega(\tau) \in \mathbb{C}^{n(L+L')\times n(L+L')}$ is a block-diagonal matrix given by

$$\Omega(\tau) = \mathbf{R_S}(0)^{-\frac{1}{2}} \cdot \mathbf{R_S}(\tau) \cdot \mathbf{R_S}(0)^{-\frac{H}{2}}.$$

The simultaneous block-diagonalization in Eq. (5) leads to a unitary matrix $\mathbf{Q}$ that is equal to $\mathbf{V}$ up to some trivial indeterminacies:

$$\mathbf{V} = \mathbf{Q} \cdot \mathbf{P} \cdot \mathbf{D_V},$$

with $\mathbf{D_V}$ an unkown block-diagonal matrix containing $(L+L') \times (L+L')$ unitary blocks and $\mathbf{P}$ an unkown permutation of these blocks. Combination with Eq. (4) yields

$$\mathbf{A} = \mathbf{U} \cdot \mathbf{\Lambda} \cdot \mathbf{Q} \cdot \mathbf{P} \cdot \mathbf{D_V} \cdot \mathbf{R_S}(0)^{-\frac{1}{2}} = \mathbf{U} \cdot \mathbf{\Lambda} \cdot \mathbf{Q} \cdot \mathbf{P} \cdot \mathbf{D_A}, \tag{6}$$

in which $\mathbf{D_A} \in \mathbb{C}^{n(L+L')\times n(L+L')}$ is an unknown nonsingular block-diagonal matrix. In this way, the sources have been separated, but for each source we have introduced some filtering, represented by the corresponding block of $\mathbf{D_A}$. If required, the original sources may be recovered using a blind SIMO technique [1].

# 4 Simultaneous Block-diagonalization

The problem we want to solve is the following. Given matrices $\mathbf{M}_1, \ldots, \mathbf{M}_P \in \mathbb{C}^{N\times N}$, with $N = nR$, find a unitary matrix $\mathbf{Q}$ that maximizes

$$f(\mathbf{Q}) = \sum_{p=1}^{P} \|\texttt{blockdiag}(\mathbf{Q}^H \cdot \mathbf{M}_p \cdot \mathbf{Q})\|^2, \tag{7}$$

in which $\texttt{blockdiag}(\cdot)$ takes the block-diagonal, consisting of $n$ $(R\times R)$-blocks, of its argument. It is clear that $\mathbf{Q}$ can only be determined up to a block-diagonal unitary transformation and a block permutation, as illustrated in Eq. (3). We will propose a Jacobi-iteration scheme.

## 4.1 Real Case

Consider a Jacobi rotation matrix $\mathbf{J} \in \mathbb{R}^{N\times N}$, equal to the identity matrix, except for the entries at position $(i,i)$ and $(j,j)$, which are given by $\cos\theta$, and the entries at position $(i,j)$ and $(j,i)$, which are given by $\sin\theta$ and $-\sin\theta$, respectively $(i<j)$. Let the $(r,s)$ entry of the transformed matrix

$$\mathbf{M'}_p = \mathbf{J}^T \cdot \mathbf{M}_p \cdot \mathbf{J}$$

be represented by $m'_{rsp}$. Then we have to maximize

$$f = \sum_p \left( (m'_{iip})^2 + (m'_{jjp})^2 + \sum_{k\in B_1, k\neq i} ((m'_{kip})^2 + (m'_{ikp})^2) + \sum_{k\in B_2, k\neq j} ((m'_{kjp})^2 + (m'_{jkp})^2) \right)$$

in which $B_1$ is the set of indices in the diagonal block to which $(i,i)$ belongs and $B_2$ is the set of indices in the diagonal block to which $(j,j)$ belongs. We have:

$$\begin{aligned}
(m'_{iip})^2 &= (m_{iip}\cos^2\theta - (m_{ijp}+m_{jip})\cos\theta\sin\theta + m_{jjp}\sin^2\theta)^2 \\
(m'_{jjp})^2 &= (m_{iip}\sin^2\theta + (m_{ijp}+m_{jip})\cos\theta\sin\theta + m_{jjp}\cos^2\theta)^2 \\
(m'_{ikp})^2 &= (m_{ikp}\cos\theta - m_{jkp}\sin\theta)^2\,(\cos^2\theta+\sin^2\theta) \qquad (k\neq i) \\
(m'_{jkp})^2 &= (m_{ikp}\sin\theta + m_{jkp}\cos\theta)^2\,(\cos^2\theta+\sin^2\theta) \qquad (k\neq j) \\
(m'_{kip})^2 &= (m_{kip}\cos\theta - m_{kjp}\sin\theta)^2\,(\cos^2\theta+\sin^2\theta) \qquad (k\neq i) \\
(m'_{kjp})^2 &= (m_{kip}\sin\theta + m_{kjp}\cos\theta)^2\,(\cos^2\theta+\sin^2\theta) \qquad (k\neq j)
\end{aligned}$$

$f$ can be maximized by computing the critical points of its derivative $f'$, which is also homogeneous in $\cos\theta$ and $\sin\theta$, and retaining the global optimum. Computing the critical points of $f'$ amounts to finding the roots of a polynomial of degree 4 in $t = \tan\theta$. (Note that the SOBI equivalent is a quadratic problem.) This approach is more direct than the one in [3] (cf. infra).

## 4.2 Complex Case

### 4.2.1 Polynomial Rooting Approach

In [3] the following problem is tackled by means of a Jacobi algorithm:

$$\max_{\mathbf{E}} \sum_p \|\mathbf{E}^H \cdot \mathbf{M}_p \cdot \mathbf{E}\|^2,$$

with $\mathbf{E} \in \mathbb{C}^{N \times R}$ having orthonormal columns. We may use the subresults to solve our problem.

Consider a Jacobi rotation matrix $\mathbf{J} \in \mathbb{C}^{N \times N}$. The entries at position $(i,i)$ and $(j,j)$ are equal to $c$, and the entries at position $(i,j)$ and $(j,i)$ are given by $s$ and $-s^*$, respectively, where

$$c = \cos\theta \qquad s = \sin\theta\, e^{i\alpha}.$$

Let the $(r,s)$ entry of the transformed matrix

$$\mathbf{M'}_p = \mathbf{J}^H \cdot \mathbf{M}_p \cdot \mathbf{J}$$

be represented by $m'_{rsp}$. Then we have to maximize

$$f = \sum_p \left( |m'_{iip}|^2 + |m'_{jjp}|^2 + \sum_{k \in B_1, k \neq i} (|m'_{kip}|^2 + |m'_{ikp}|^2) + \sum_{k \in B_2, k \neq j} (|m'_{kjp}|^2 + |m'_{jkp}|^2) \right)$$

in which $B_1$ is the set of indices in the diagonal block to which $(i,i)$ belongs and $B_2$ is the set of indices in the diagonal block to which $(j,j)$ belongs. Define

$$V = \begin{pmatrix} \cos(2\theta) & \sin(2\theta)\cos\alpha & \sin(2\theta)\sin\alpha \end{pmatrix}^T,$$

$$\begin{aligned}
\mathbf{G}_p &= \mathtt{Re}(Z_p Z_p^H) \\
\mathbf{G'}_p &= \mathtt{Re}(Z'_p {Z'}_p^H) \\
G_p &= \mathtt{Re}((m_{iip}^* + m_{jjp}^*) Z_p) \\
G'_p &= \mathtt{Re}((m_{iip}^* + m_{jjp}^*) Z'_p) \\
Z_p &= \frac{1}{2} \begin{pmatrix} m_{iip} - m_{jjp} \\ -(m_{ijp} + m_{jip}) \\ i(m_{ijp} - m_{jip}) \end{pmatrix} \\
Z'_p &= \frac{1}{2} \begin{pmatrix} m_{iip} - m_{jjp} \\ m_{ijp} + m_{jip} \\ i(m_{jip} - m_{ijp}) \end{pmatrix} \\
X_{kp} &= \begin{pmatrix} \frac{1}{2}(|m_{kjp}|^2 - |m_{kip}|^2) \\ \mathtt{Re}(m_{kip}^* m_{kjp}) \\ \mathtt{Im}(m_{kip}^* m_{kjp}) \end{pmatrix} \\
Y_{kp} &= \begin{pmatrix} \frac{1}{2}(|m_{jkp}|^2 - |m_{ikp}|^2) \\ \mathtt{Re}(m_{ikp} m_{jkp}^*) \\ \mathtt{Im}(m_{ikp} m_{jkp}^*) \end{pmatrix}.
\end{aligned}$$

Then we have:

$$\begin{aligned}
|m'_{iip}|^2 &= V^T \mathbf{G}_p V + G_p^T V + \frac{1}{4} |m_{iip} + m_{jjp}|^2 \\
|m'_{jjp}|^2 &= V^T \mathbf{G'}_p V + {G'}_p^T V + \frac{1}{4} |m_{iip} + m_{jjp}|^2 \\
|m'_{ikp}|^2 &= -V^T Y_{kp} + \frac{1}{2} (|m_{ikp}|^2 + |m_{jkp}|^2) \quad (k \neq i) \\
|m'_{jkp}|^2 &= V^T Y_{kp} + \frac{1}{2} (|m_{ikp}|^2 + |m_{jkp}|^2) \quad (k \neq j) \\
|m'_{kip}|^2 &= -V^T X_{kp} + \frac{1}{2} (|m_{kip}|^2 + |m_{kjp}|^2) \quad (k \neq i) \\
|m'_{kjp}|^2 &= V^T X_{kp} + \frac{1}{2} (|m_{kip}|^2 + |m_{kjp}|^2) \quad (k \neq j).
\end{aligned}$$

Hence $f$ takes the form

$$f = V^T \mathbf{G} V + G^T V + \text{constant}, \tag{8}$$

with

$$\begin{aligned} \mathbf{G} &= \sum_p (\mathbf{G}_p + \mathbf{G}'_p) \\ G &= \sum_p \left( G_p + G'_p - \sum_{k \in B_1, k \neq i} (X_{kp} + Y_{kp}) + \sum_{k \in B_2, k \neq j} (X_{kp} + Y_{kp}) \right). \end{aligned}$$

Maximization of (8), under the constraint $\|V\| = 1$, using a Lagrange multiplier, leads to

$$2(\mathbf{G} + \lambda \mathbf{I})V + G = 0, \tag{9}$$

and, hence,

$$V = -\frac{1}{2}(\mathbf{G} + \lambda \mathbf{I})^{-1} G, \tag{10}$$

where $\lambda$ is a real scalar chosen in such a way that

$$\|V\|^2 = 1 \Leftrightarrow \frac{1}{4} \sum_{i=1}^{3} \frac{(U_i^T G)^2}{(\lambda_i + \lambda)^2} = 1, \tag{11}$$

$U_i$ and $\lambda_i$ being the eigenvectors and -values of $\mathbf{G}$. As we can see from (11), the problem amounts to rooting a polynomial of degree 6 and selecting the root of which the corresponding $V$ maximizes $f$. (SOBI involves a polynomial of degree 3.) The elements of the optimal Jacobi rotation follow from $\cos\theta = \sqrt{1 + \cos(2\theta)/2}$ and $\sin\theta \, e^{i\alpha} = (\sin(2\theta)\, \cos\alpha + i\, \sin(2\theta)\, \sin\alpha)/(2\cos\theta)$.

If $G$ is orthogonal to one of the eigenvectors of $\mathbf{G}$, the corresponding $-\lambda_i$ is a double root of (11). In this case, we have to resort to (9) instead of (10) for the computation of $V$:

$$V = -\frac{1}{2}(\mathbf{G} - \lambda_i \mathbf{I})^{\dagger} G + c_i U_i,$$

where $c_i$ is a real constant chosen such that $\|V\| = 1$ and $f$ is maximum. If it exists, it is given by

$$c_i = \operatorname{sign}(U_i^T G) \sqrt{1 - \|(\mathbf{G} - \lambda_i \mathbf{I})^{\dagger} G\|^2 / 4}.$$

[3] mentions that this problem is observed near convergence. Actually, the situation is worse: when there is little noise, $G$ becomes proportional to the canonical unit vector $E_1$ and $\mathbf{G}$ converges to a multiple of $E_1 E_1^T$. Hence, $\lambda = 0$ is a quadruple root and the corresponding $V$ follows from

$$V = -\frac{1}{2}\mathbf{G}^{\dagger} G + c_2 U_2 + c_3 U_3,$$

in which the first term is proportional to $E_1$ and $U_2$, $U_3$ are orthogonal to it; $c_1$ and $c_2$ are real constants chosen such that $\|V\| = 1$ and $f$ is maximum. There are two possibilities: (1) if $\|\frac{1}{2}\mathbf{G}^{\dagger} G\| > 1$, then this does not yield a solution, and (2) if $\|\frac{1}{2}\mathbf{G}^{\dagger} G\| < 1$, then the solution is not unique. However, the remaining possibilities for $\lambda$, given by

$$(U_1^T G)^2 = (\lambda_1 + \lambda)^2,$$

lead to solutions that are close to $\pm E_1$.

#### 4.2.2 Tensorial Rank-1 Approximation Approach

The optimal Jacobi rotation can also be computed via the best super-symmetric rank-1 approximation of a partially symmetric $(2 \times 2 \times 2 \times 2)$-tensor. This approach has the conceptual advantage that we aim directly at the global optimum, instead of determining all the critical points of $f$ and selecting the best one. However, current implementations of the technique may not be faster than the Matlab `roots` command.

Call $X = (c \;\; s)^T$. Let the outer product be symbolized by "$\circ$":

$$(\mathcal{A} \circ \mathcal{B})_{i_1 i_2 \ldots j_1 j_2 \ldots} = (\mathcal{A})_{i_1 i_2 \ldots} (\mathcal{B})_{j_1 j_2 \ldots}.$$

Define the following mapping from $\mathbb{C}^{2\times 2}$ to $\mathbb{C}^{2\times 2}$:

$$\mathcal{M}(\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}) = \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}.$$

Define the following matrices:

$$\begin{aligned}
\mathbf{F}_p &= \begin{pmatrix} m_{iip} & m_{ijp} \\ m_{jip} & m_{jjp} \end{pmatrix} \\
\mathbf{G}_{1p} &= \begin{pmatrix} \cdots & m_{ikp} & \cdots \\ \cdots & m_{jkp} & \cdots \end{pmatrix} \cdot \begin{pmatrix} \vdots & \vdots \\ m_{ikp} & m_{ijp} \\ \vdots & \vdots \end{pmatrix}^* \qquad k \in B_1, k \neq i \\
\mathbf{G}_{2p} &= \begin{pmatrix} \cdots & m_{ikp} & \cdots \\ \cdots & m_{jkp} & \cdots \end{pmatrix} \cdot \begin{pmatrix} \vdots & \vdots \\ m_{ikp} & m_{ijp} \\ \vdots & \vdots \end{pmatrix}^* \qquad k \in B_2, k \neq j \\
\mathbf{H}_{1p} &= \begin{pmatrix} \cdots & m_{kip} & \cdots \\ \cdots & m_{kjp} & \cdots \end{pmatrix}^* \cdot \begin{pmatrix} \vdots & \vdots \\ m_{kip} & m_{kjp} \\ \vdots & \vdots \end{pmatrix} \qquad k \in B_1, k \neq i \\
\mathbf{H}_{2p} &= \begin{pmatrix} \cdots & m_{kip} & \cdots \\ \cdots & m_{kjp} & \cdots \end{pmatrix}^* \cdot \begin{pmatrix} \vdots & \vdots \\ m_{kip} & m_{kjp} \\ \vdots & \vdots \end{pmatrix} \qquad k \in B_2, k \neq j.
\end{aligned}$$

Then we have:

$$\begin{aligned}
|m'_{iip}|^2 &= |X^T \mathbf{F}_p X^*|^2 = \sum_{rstu} (\mathbf{F}_p \circ \mathbf{F}_p^*)_{rstu} x_r x_s^* x_t^* x_u \\
|m'_{jjp}|^2 &= |X^T \mathcal{M}(\mathbf{F}_p) X^*|^2 = \sum_{rstu} (\mathcal{M}(\mathbf{F}_p) \circ \mathcal{M}(\mathbf{F}_p)^*)_{rstu} x_r x_s^* x_t^* x_u \\
\sum_{k\in B_1, k\neq i} |m'_{ikp}|^2 &= X^T \mathbf{G}_{1p} X^* = \sum_{rstu} \frac{1}{2} (\mathbf{G}_{1p} \circ \mathbf{I} + \mathbf{I} \circ \mathbf{G}_{1p})_{rstu} x_r x_s^* x_t^* x_u \\
\sum_{k\in B_2, k\neq j} |m'_{ikp}|^2 &= X^T \mathcal{M}(\mathbf{G}_{2p}) X^* = \sum_{rstu} \frac{1}{2} (\mathcal{M}(\mathbf{G}_{2p}) \circ \mathbf{I} + \mathbf{I} \circ \mathcal{M}(\mathbf{G}_{2p}))_{rstu} x_r x_s^* x_t^* x_u \\
\sum_{k\in B_1, k\neq i} |m'_{kip}|^2 &= X^T \mathbf{H}_{1p} X^* = \sum_{rstu} \frac{1}{2} (\mathbf{H}_{1p} \circ \mathbf{I} + \mathbf{I} \circ \mathbf{H}_{1p})_{rstu} x_r x_s^* x_t^* x_u \\
\sum_{k\in B_2, k\neq j} |m'_{ikp}|^2 &= X^T \mathcal{M}(\mathbf{H}_{2p}) X^* = \sum_{rstu} \frac{1}{2} (\mathcal{M}(\mathbf{H}_{2p}) \circ \mathbf{I} + \mathbf{I} \circ \mathcal{M}(\mathbf{H}_{2p}))_{rstu} x_r x_s^* x_t^* x_u.
\end{aligned}$$

Hence, $f$ takes the form

$$f = \sum_{rstu} a_{rstu} x_r x_s^* x_t^* x_u, \tag{12}$$

in which $\mathcal{A}$ is a $(2\times 2\times 2\times 2)$-tensor that has the symmetry $a_{rstu} = a_{turs}^*$. Maximization of $f$ is equivalent to the best least-squares approximation of $\mathcal{A}$ by a super-symmetric rank-1 tensor $\lambda X \circ X^* \circ X^* \circ X$ [5].

For the computation of the best rank-1 approximation of a given tensor, a higher-order Rayleigh Quotient Iteration was proposed in [8]. This algorithm converges quadratically in a neighbourhood of the solution.

We adapt the algorithm to impose super-symmetry on the rank-1 approximation ($\delta_{rs}$ is the Kronecker delta):

Given initial estimate $X^{(0)}$
**for** $l = 0, 1, \ldots$
 normalize $X^{(l)}$ so that $\|X^{(l)}\| = 1$
 Compute $\lambda^{(l)}$, $B$, $\mathbf{M}_1$, $\mathbf{M}_2$:

$$\begin{aligned}
\lambda^{(l)} &= \sum_{rstu} a_{rstu} x_r^{(l)} x_s^{(l)^*} x_t^{(l)^*} x_u^{(l)} \\
b_r &= \sum_{stu} a_{rstu} x_s^{(l)^*} x_t^{(l)^*} x_u^{(l)} \\
(\mathbf{M}_1)_{rs} &= \sum_{tu} a_{rtus} x_t^{(l)^*} x_u^{(l)^*} - \lambda^{(l)} \delta_{rs} \\
(\mathbf{M}_2)_{rs} &= \sum_{tu} a_{rstu} x_t^{(l)^*} x_u^{(l)} + a_{rtsu} x_t^{(l)^*} x_u^{(l)}
\end{aligned}$$

 Solve for $X^{(l+1)}$:

$$\begin{pmatrix} \mathtt{Re}(\mathbf{M}_1 + \mathbf{M}_2) & -\mathtt{Im}(\mathbf{M}_1 - \mathbf{M}_2) \\ \mathtt{Im}(\mathbf{M}_1 + \mathbf{M}_2) & \mathtt{Re}(\mathbf{M}_1 - \mathbf{M}_2) \end{pmatrix} \cdot \begin{pmatrix} \mathtt{Re}(X^{(l+1)}) \\ \mathtt{Im}(X^{(l+1)}) \end{pmatrix} = \begin{pmatrix} \mathtt{Re}(B) \\ \mathtt{Im}(B) \end{pmatrix} \tag{13}$$

**endfor**

Two interesting initial values can be obtained as follows. Associate to $\mathcal{A}$ the matrix-to-matrix mapping defined by a summation over its last two indices. Compute the dominant eigenmatrix $\mathbf{E}$ of this $(4 \times 4)$ Hermitean mapping. It makes sense to choose $X^{(0)}$ such that $X^H \mathbf{E} X$ is maximized. This the case for the two Schur vectors associated with the dominant eigenvalue of $\mathbf{E}$ [6]. An alternative is to associate to $\mathcal{A}$ the mapping from third-order tensors to vectors defined by a summation over its last three indices. A computationally cheaper but perhaps less accurate initialization is then to choose $X^{(0)}$ as the left singular vector associated to the dominant singular value of this $(2 \times 8)$ mapping, as proposed in [5].

# 5 Conclusion

In this paper we derived a convolutive equivalent of SOBI [2], relying on joint block-diagonalization, rather than on joint diagonalization. This problem can still be solved by means of a Jacobi iteration, but the computation of an optimal Jacobi rotation now involves the rooting of a polynomial of degree 4 (real data) or 6 (complex data). The problem can also be formulated in terms of the best super-symmetric rank-1 approximation of a $(2 \times 2 \times 2 \times 2)$-tensor. This approximation can be calculated, e.g., by means of a higher-order Rayleigh-Quotient Iteration. The principle of this paper can also be used to generalize [7, 8] to convolutive mixtures.


### Acknowledgements

Lieven De Lathauwer holds a permanent research position with the French Centre National de la Recherche Scientifique (C.N.R.S.); he also holds a honorary post-doctoral research mandate with the Fund for Scientific Research-Flanders (F.W.O.), affiliated with the K.U.Leuven. Cédric Févotte holds a Ph.D. research position with the French Ministry of Research. 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 its authors.

# References

[1] K. Abed-Meraim, W. Qiu, Y. Hua, "Blind System Identification", *Proc. of the IEEE*, Vol. 85, No. 8, Aug. 1997, pp. 1310–1322.

[2] A. Belouchrani, et al., "A Blind Source Separation Technique using Second Order Statistics", *IEEE Trans. Signal Processing*, Vol. 45, No. 2, Feb. 1997, pp. 434–444.

[3] A. Belouchrani, K. Abed-Meraim, Y. Hua, "Jacobi-like Algorithms for Joint Block Diagonalization: Application to Source Localization", *Proc. IEEE Int. Work. on Intelligent Sig. Proc. and Comm. Sys. (ISPACS)*, Melbourne, Australia, Nov. 1998.

[4] J. Cambré, L. De Lathauwer, B. De Moor, J. Vandewalle, "Best Rank-$(R, R, R)$ Super-Symmetric Tensor Approximation — A Continuous-Time Approach", *Proc. IEEE Signal Processing Workshop on Higher-Order Statistics (HOS'99)*, Caesarea, Israel, June 14–16, 1999, pp. 242–246.

[5] L. De Lathauwer, B. De Moor, J. Vandewalle, "On the Best Rank-1 and Rank-$(R_1, R_2, \ldots, R_N)$ Approximation of Higher-Order Tensors", *SIAM J. Matrix Anal. Appl.*, Vol. 21, No. 4, April 2000, pp. 1324–1342.

[6] G.H. Golub, C.F. Van Loan, *Matrix Computations*, 3rd ed., Johns Hopkins University Press, Baltimore, Maryland, 1996.

[7] A. Holobar, C. Févotte, C. Doncarli, D. Zazula, "Single Autoterms Selection for Blind Source Separation in Time-frequency Plane", *Proc. XIth European Signal Processing Conf. (EUSIPCO-2002)*, Toulouse, France, Sept. 3–6, 2002.

[8] D.-T. Pham, J.-F. Cardoso, "Blind Separation of Instantaneous Mixtures of Non-stationary Sources", *IEEE Trans. Signal Processing*, Vol. 49, No. 9, pp. 1837–1848.

[9] T. Zhang, G.H. Golub, "Rank-one approximation to High Order Tensors", *SIAM J. Matr. Anal. Appl.*, Vol. 23, No. 2, pp. 534–550.