ON THE BLIND SEPARATION OF NON-CIRCULAR SOURCES

Lieven De Lathauwer and Bart De Moor† [1]
ETIS, ENSEA/UCP, Cergy-Pontoise, France
†E.E. Dept. (ESAT), K.U.Leuven, Leuven, Belgium
e-mail: lieven.delathauwer@ensea.fr, bart.demoor@esat.kuleuven.ac.be

ABSTRACT

In this paper we address the blind separation of an instantaneous complex mixture of statistically independent non-circular signals. We show that, if the sources are non-circular at order 2, by exploiting all the second-order information, the mixing matrix can be estimated up to a real orthogonal factor. This is based on a link with the Takagi factorization, for the computation of which we derive a Jacobi-type algorithm. We prove that, if the sources are non-circular at order 4, after a classical prewhitening, the remaining unitary factor can be found via a simultaneous Takagi factorization / Hermitian Eigenvalue Decomposition (EVD). We also describe a variant in which no hard prewhitening is carried out. In addition, we pay some attention to the issue of dimensionality reduction, in the case where there are fewer sources than sensors.

1 INTRODUCTION

The basic statistical model for Independent Component Analysis (ICA), or Blind Source Separation (BSS), is in this paper denoted as

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

in which the observed vector $Y \in \mathbb{C}^J$, the source vector $X \in \mathbb{C}^I$ and the noise vector $N \in \mathbb{C}^J$ are zero-mean random vectors. The components of X are mutually statistically independent, as well as statistically independent from the noise components. We assume that $I \leqslant J$ and that the mixing matrix $\mathbf{M} \in \mathbb{C}^{J\times I}$ is non-singular. The goal of ICA consists of the estimation of $\mathbf{M}$ and the corresponding realizations of X, given only realizations of Y.

[1]This work is supported by (1) the Research Council K.U.Leuven: GOA-MEFISTO-666, (2) the Flemish Government: (a) the Fund for Scientific Research-Flanders (F.W.O.) projects G.0256.97, G.0240.99, (b) the F.W.O. Research Communities ICCoS and ANMMM, (3) the Belgian Federal Government: IUAP IV-02 and IUAP V-22. 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 F.W.O., affiliated with the K.U.Leuven. Bart De Moor is a Full Professor with the K.U.Leuven. The scientific responsibility is assumed by its authors.

In this paper we assume that the sources are non-circular. Non-circularity at order 2 is dealt with in Section 2; it applies to, e.g., BPSK constellations. Non-circularity at order 4 is dealt with in Section 4; it applies to, e.g., QAM4 and QAM16 constellations. On the other hand, it is natural to assume that the noise is circular (e.g. Gaussian).

In what follows, $\cdot^T$ denotes the transpose, $\cdot^*$ the complex conjugate and $\cdot^H$ the Hermitian adjoint.

2 SECOND-ORDER ANALYSIS

Let us define $\mathbf{C}_Y^{(1,1)} \stackrel{\text{def}}{=} \mathtt{E}\{YY^H\}$ and $\mathbf{C}_Y^{(2,0)} \stackrel{\text{def}}{=} \mathtt{E}\{YY^T\}$. We have

$$\mathbf{C}_Y^{(1,1)} = \mathbf{M}\cdot\mathbf{C}_X^{(1,1)}\cdot\mathbf{M}^H + \mathbf{C}_N^{(1,1)}, \tag{2}$$
$$\mathbf{C}_Y^{(2,0)} = \mathbf{M}\cdot\mathbf{C}_X^{(2,0)}\cdot\mathbf{M}^T, \tag{3}$$

in which, due to the mutual statistical independence of the sources, $\mathbf{C}_X^{(1,1)} = \mathtt{diag}\{\sigma_1^2,\ldots,\sigma_I^2\}$, with $\sigma_i^2 = \mathtt{E}\{|x_i|^2\}$, and $\mathbf{C}_X^{(2,0)} = \mathtt{diag}\{\tilde{\sigma}_1^2,\ldots,\tilde{\sigma}_I^2\}$, with $\tilde{\sigma}_i^2 = \mathtt{E}\{x_i^2\}$. We take $\sigma_i \in \mathbb{R}^+$. In this section we make the assumption that at most one of the entries of $\mathbf{C}_X^{(2,0)}$ vanishes. Note that, unlike (2), (3) does not contain a noise term.

The mixing matrix can only be determined modulo a permutation and scaling of its columns. In the equivalence class we may consider

$$\mathbf{M}' \stackrel{\text{def}}{=} \mathbf{M}\cdot\mathtt{diag}\{\sigma_1 e^{i\phi_1},\ldots,\sigma_I e^{i\phi_I}\}, \tag{4}$$

with ϕ_i defined by $\tilde{\sigma}_i^2 = |\tilde{\sigma}_i|^2 e^{i2\phi_i}$. Let a Singular Value Decomposition (SVD) of $\mathbf{M}'$ be given by

$$\mathbf{M}' = \mathbf{U}\cdot\mathbf{S}\cdot\mathbf{V}, \tag{5}$$

in which $\mathbf{U} \in \mathbb{C}^{J\times I}$ has orthonormal columns, $\mathbf{S} \in \mathbb{R}^{I\times I}$ is positive diagonal, and $\mathbf{V} \in \mathbb{C}^{I\times I}$ is unitary. Then $\mathbf{U}$ and $\mathbf{S}$ can be found from an EVD of $\mathbf{C}_Y^{(1,1)}$:

$$\mathbf{C}_Y^{(1,1)} = \mathbf{U}\cdot\mathbf{S}^2\cdot\mathbf{U}^H \tag{6}$$

(neglecting the noise term, for clarity; if $J > I$ and the noise is spatially white, then the noise variance may be estimated as the mean of the $J-I$ smallest eigenvalues of $\mathbf{C}_Y^{(1,1)}$, and the noise term may be compensated by subtracting this value from the I biggest eigenvalues, as is well-known). Substituting these results in (3) leads to

$$\mathbf{S}^{-1} \cdot \mathbf{U}^H \cdot \mathbf{C}_Y^{(2,0)} \cdot \mathbf{U}^* \cdot \mathbf{S}^{-1} = \mathbf{V} \cdot \mathtt{diag}\{\tfrac{|\tilde{\sigma}_1|^2}{\sigma_1^2}, \ldots, \tfrac{|\tilde{\sigma}_I|^2}{\sigma_I^2}\} \cdot \mathbf{V}^T, \quad (7)$$

which is a Takagi factorization [9]:

Theorem 1 *If $\mathbf{A} \in \mathbb{C}^{I\times I}$ is symmetric ($\mathbf{A} = \mathbf{A}^T$), then there exists a unitary $\mathbf{E}$ and a real nonnegative diagonal $\mathbf{\Lambda}$ such that $\mathbf{A} = \mathbf{E}\mathbf{\Lambda}\mathbf{E}^T$. The columns of $\mathbf{E}$ are an orthonormal set of eigenvectors for $\mathbf{A}\mathbf{A}^*$, and the corresponding diagonal entries of $\mathbf{\Lambda}$ are the nonnegative square roots of the corresponding eigenvalues of $\mathbf{A}\mathbf{A}^*$.*

Let us first consider the hypothetical situation in which all the values $|\tilde{\sigma}_i|/\sigma_i$ are distinct. In this case $\mathbf{V}$ can be uniquely determined, up to the sign of its rows, so that we don't need Higher-Order Statistics (HOS) to solve the problem. However, perturbations of a matrix may result in significant changes of eigenvectors, when the corresponding eigenvalues are close. In this respect, it may be wise to resort to HOS anyway, to obtain extra constraints on the mixing matrix. For instance, it is clear that, by embedding (7) in a joint diagonalization (JODI), like we will do in Section 4, the estimation of $\mathbf{V}$ becomes more robust, cf. [2].

Now let us turn to the more common situation in which all the sources have the same distribution. It is still possible to extract useful information from $\mathbf{C}_Y^{(2,0)}$.

The key is that, when $\mathbf{A}\mathbf{A}^*$ in the Takagi factorization theorem has an eigenvalue with multiplicity greater than 1, not every orthonormal set of corresponding eigenvectors can be used for the corresponding columns of $\mathbf{E}$. Assume that λ^2 is an eigenvalue of $\mathbf{A}\mathbf{A}^*$ with multiplicity $R \leqslant I$, that we have the EVD

$$\mathbf{A}\mathbf{A}^* = (\mathbf{E}_1\, \mathbf{E}_2) \cdot \begin{pmatrix} \lambda^2 \mathbf{I} & \\ & \mathbf{\Lambda}_2^2 \end{pmatrix} \cdot \begin{pmatrix} \mathbf{E}_1^H \\ \mathbf{E}_2^H \end{pmatrix} \quad (8)$$

with the obvious partitioning, and that a Takagi factorization of $\mathbf{A}$ is given by

$$\mathbf{A} = (\mathbf{E}_1\, \mathbf{E}_2) \cdot \begin{pmatrix} \lambda \mathbf{I} & \\ & \mathbf{\Lambda}_2 \end{pmatrix} \cdot \begin{pmatrix} \mathbf{E}_1^T \\ \mathbf{E}_2^T \end{pmatrix}. \quad (9)$$

In (8), $\mathbf{E}_1$ may be replaced by $\mathbf{E}_1\mathbf{Q}$, in which $\mathbf{Q} \in \mathbb{C}^{R\times R}$ is unitary, i.e., $\mathbf{Q}\mathbf{Q}^H = \mathbf{I}$. If we replace $\mathbf{E}_1$ in (9) by $\mathbf{E}_1\mathbf{Q}$, then this gives only a factorization of $\mathbf{A}$ when $\mathbf{Q}\mathbf{Q}^T = \mathbf{I}$. Hence, $\mathbf{Q} = \mathbf{Q}^*$, or $\mathbf{Q}$ can only be a real orthogonal matrix.

This means that, when the sources are identically distributed, $\mathbf{V}$ can be found from (7) up to a real orthogonal factor. This factor has to be estimated from the HOS of Y. The fact that the factor is real, drastically reduces the computational complexity — cf. [1, 3, 6]. Note that in this way, half of the independent parameters of $\mathbf{V}$ are obtained from $\mathbf{C}_Y^{(2,0)}$, and the other half from the HOS of Y.

Instead of calculating $\mathbf{C}_Y^{(1,1)}$ explicitly, it is numerically preferable to work via the SVD of the dataset, such that the singular values are not squared. For a square-root version of (7) we may resort to the following theorem:

Theorem 2 *A matrix $\mathbf{A} \in \mathbb{C}^{I\times T}$ may be decomposed as $\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T$, with $\mathbf{U} \in \mathbb{C}^{I\times I}$ unitary, $\mathbf{S} \in \mathbb{R}^{I\times T}$ diagonal containing R strictly positive entries and $\mathbf{V} \in \mathbb{C}^{T\times T}$ complex orthogonal (i.e., $\mathbf{V}^T\mathbf{V} = \mathbf{I}$) iff $\mathtt{rank}(\mathbf{A}\mathbf{A}^T) = \mathtt{rank}(\mathbf{A}) = R$.*

This theorem can be proved in analogy with the SVD theorem. Note that $\mathbf{V}^T\mathbf{V}$ does not prevent the entries of $\mathbf{V}$ from being big, which is a numerical disadvantage. Due to lack of space, we will not discuss procedures for the computation of the decomposition.

3 JACOBI ALGORITHM FOR TAKAGI'S FACTORIZATION

The components of the Takagi factorization of a matrix $\mathbf{A}$ can be derived from an EVD of $\mathbf{A}\mathbf{A}^*$; however, this approach has the numerical disadvantage that the condition number is squared. In this section we will work directly on $\mathbf{A}$. We propose a Jacobi-type algorithm for the calculation of the decomposition; as far as we know, this has not appeared in the literature yet. The derivation is analogous to that of its Hermitian EVD counterpart [8].

By left multiplication of $\mathbf{A}$ with a Jacobi-rotation affecting rows p and q, and right multiplication with its transpose, the Frobenius-norm of the (off-)diagonal part of $\mathbf{A}$ can only be changed through the transformation of the entries a_{pp}, a_{pq} and a_{qq}. Let us represent the Jacobi-rotation by

$$\mathbf{J} = \begin{pmatrix} \cos\alpha & -\sin\alpha\, e^{j\phi} \\ \sin\alpha\, e^{-j\phi} & \cos\alpha \end{pmatrix}. \quad (10)$$

The off-diagonal entry of

$$\begin{pmatrix} b_{pp} & b_{pq} \\ b_{qp} & b_{qq} \end{pmatrix} \stackrel{\text{def}}{=} \mathbf{J}^T \cdot \begin{pmatrix} a_{pp} & a_{pq} \\ a_{qp} & a_{qq} \end{pmatrix} \cdot \mathbf{J} \quad (11)$$

is then given by $b_{pq} = \frac{1}{2} G^T \cdot V$, in which

$$V^T \stackrel{\text{def}}{=} (\cos 2\alpha, \sin 2\alpha \cos\phi, \sin 2\alpha \sin\phi), \quad (12)$$
$$G^T \stackrel{\text{def}}{=} (2\, a_{pq}, a_{qq} - a_{pp}, -j\,(a_{qq} + a_{pp})). \quad (13)$$

Hence b_{pq} can be made zero by choosing V as a real unit-norm vector that satisfies

$$\begin{pmatrix} \mathtt{Re}(G^T) \\ \mathtt{Im}(G^T) \end{pmatrix} \cdot V = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (14)$$

The elements of the optimal Jacobi rotation follow from $\cos\alpha = \sqrt{1+\cos(2\alpha)/2}$ and $\sin\alpha\, e^{i\phi} = (\sin(2\alpha)\,\cos\phi + i\,\sin(2\alpha)\,\sin\phi)/(2\cos\alpha)$. By choosing the first entry of V to be positive, $\mathbf{J}$ can be restricted to the inner rotations ($\alpha \in (-\pi/4, \pi/4]$).

If we assume that a_{pq} was the off-diagonal entry with largest modulus, then the squared Frobenius-norm of the off-diagonal part of $\mathbf{A}$ was reduced with at least a factor $1 - \frac{2}{I(I-1)}$. Hence, by repeating this procedure we necessarily converge to a diagonal form. (In practice, we address the off-diagonal entries in a cyclic way.) A complex phase of the diagonal entries can be incorporated in the overall unitary factor. This is a constructive proof of Theorem 1.

4 PREWHITENING-BASED COMPUTATION

Let us assume that a classical prewhitening, i.e., diagonalization of $\mathbf{C}_Y^{(1,1)}$, has been carried out. We consider then the transformed observation vector $Z \stackrel{\text{def}}{=} \mathbf{S}^{-1}\mathbf{U}^H Y$, and the task becomes the estimation of the unitary matrix $\mathbf{V}$.

The classical approach is to exploit conditions on the structure of $\mathcal{C}_Z^{(2,2)} = \mathtt{Cum}\{Z, Z^*, Z, Z^*\}$. Due to the statistical independence of the components of X, we have

$$(\mathcal{C}_Z^{(2,2)})_{i_1 i_2 i_3 i_4} = \sum_i \kappa_i\, v_{i_1 i} v_{i_2 i}^* v_{i_3 i} v_{i_4 i}^*, \qquad (15)$$

in which $\kappa_i \stackrel{\text{def}}{=} \mathtt{Cum}\{x_i, x_i^*, x_i, x_i^*\}$, which we will write as

$$\mathcal{C}_Z^{(2,2)} = \mathcal{C}_X^{(2,2)} \times_1 \mathbf{V} \times_2 \mathbf{V}^* \times_3 \mathbf{V} \times_4 \mathbf{V}^*, \qquad (16)$$

in which $\mathcal{C}_X^{(2,2)} = \mathtt{diag}\{\kappa_1, \ldots, \kappa_I\}$. In [1] the condition of diagonality of $\mathcal{C}_X^{(2,2)}$ is exploited in a simultaneous Hermitian EVD:

$$\begin{aligned} \mathbf{A}_1^{(2,2)} &= \mathbf{V} \cdot \mathbf{D}_1^{(2,2)} \cdot \mathbf{V}^H \\ &\vdots \\ \mathbf{A}_K^{(2,2)} &= \mathbf{V} \cdot \mathbf{D}_K^{(2,2)} \cdot \mathbf{V}^H. \end{aligned} \qquad (17)$$

$\mathbf{A}_k^{(2,2)} \in \mathbb{C}^{I \times I}$ $(1 \leqslant k \leqslant K)$ are Hermitian matrices, derived from $\mathcal{C}_Z^{(2,2)}$, and the goal is to estimate $\mathbf{V}$ as the unitary matrix that makes $\mathbf{D}_k^{(2,2)}$ $(1 \leqslant k \leqslant K)$ simultaneously as diagonal as possible in the Frobenius sense. In [10] it is shown that the problem can also be rephrased as

$$\begin{aligned} \tilde{\mathbf{A}}_1^{(2,2)} &= \mathbf{V} \cdot \tilde{\mathbf{D}}_1^{(2,2)} \cdot \mathbf{V}^T \\ &\vdots \\ \tilde{\mathbf{A}}_{\tilde{K}}^{(2,2)} &= \mathbf{V} \cdot \tilde{\mathbf{D}}_{\tilde{K}}^{(2,2)} \cdot \mathbf{V}^T, \end{aligned} \qquad (18)$$

in which $\tilde{\mathbf{D}}_k^{(2,2)}$ $(1 \leqslant k \leqslant K)$ are (theoretically) diagonal complex matrices. We will call this a simultaneous Takagi factorization.

For sources that are non-circular at order 4, we may also consider $\mathcal{C}_Z^{(4,0)} = \mathtt{Cum}\{Z, Z, Z, Z\}$. We have

$$(\mathcal{C}_Z^{(4,0)})_{i_1 i_2 i_3 i_4} = \sum_i \tilde{\kappa}_i\, v_{i_1 i} v_{i_2 i} v_{i_3 i} v_{i_4 i}, \qquad (19)$$

in which $\tilde{\kappa}_i \stackrel{\text{def}}{=} \mathtt{Cum}\{x_i, x_i, x_i, x_i\}$, which we write as

$$\mathcal{C}_Z^{(4,0)} = \mathcal{C}_X^{(4,0)} \times_1 \mathbf{V} \times_2 \mathbf{V} \times_3 \mathbf{V} \times_4 \mathbf{V}, \qquad (20)$$

in which $\mathcal{C}_X^{(4,0)} = \mathtt{diag}\{\tilde{\kappa}_1, \ldots, \tilde{\kappa}_I\}$. This again leads to a simultaneous Takagi factorization:

$$\begin{aligned} \mathbf{A}_1^{(4,0)} &= \mathbf{V} \cdot \mathbf{D}_1^{(4,0)} \cdot \mathbf{V}^T \\ &\vdots \\ \mathbf{A}_L^{(4,0)} &= \mathbf{V} \cdot \mathbf{D}_L^{(4,0)} \cdot \mathbf{V}^T. \end{aligned} \qquad (21)$$

By taking this extra information into account, we may expect to enhance the accuracy. In addition, we may also have Eq. (7) and equations related to $\mathcal{C}_Z^{(3,1)}$.

In [4, Section 4] we proposed a Jacobi-type approach to solve a simultaneous Takagi factorization. This technique has independently been derived in [10]. [4] mentions that the technique can also be used for a simultaneous Takagi factorization combined with a simultaneous Hermitian EVD, which applies when one decides to resort to Eqs. (17) instead of Eqs. (18). It was shown that the computation of an elementary rotation amounts to the computation of the dominant eigenvector of a real symmetric (3×3)-matrix. One can associate weights to the different equations, depending on their supposed relative reliability and importance. In case one starts with a "full" prewhitening, in which also $\mathbf{C}_Y^{(2,0)}$ is diagonalized, under the conditions specified in Section 2, the remaining unknown factor is real orthogonal. The computation of an elementary rotation then amounts to the computation of the dominant eigenvector of a real symmetric (2×2)-matrix.

5 SOFT WHITENING

It seems strange to consider second-order constraints on the mixing matrix as infinitely more reliable than higher-order constraints (note that (2) is the only equation that is explicitly affected by Gaussian noise). In this section we will handle second- and higher-order constraints simultaneously, instead of sequentially. Without loss of generality, we assume that $I = J$. The problem of dimensionality reduction will be discussed in Section 6.

If we do not perform an explicit prewhitening, then we obtain a weighted system of equations of the type

$$\begin{aligned} \mathbf{B}_1 &= \mathbf{M} \cdot \mathbf{D}_1 \cdot \mathbf{M}^H \\ &\vdots \\ \mathbf{B}_P &= \mathbf{M} \cdot \mathbf{D}_P \cdot \mathbf{M}^H \end{aligned} \qquad (22)$$

$$\begin{aligned}\tilde{\mathbf{B}}_1 &= \mathbf{M}\cdot\tilde{\mathbf{D}}_1\cdot\mathbf{M}^T\\ &\vdots\\ \tilde{\mathbf{B}}_{\tilde{P}} &= \mathbf{M}\cdot\tilde{\mathbf{D}}_{\tilde{P}}\cdot\mathbf{M}^T. \end{aligned}\tag{23}$$

If possible, we assume that (an estimate of) the noise contribution to $\mathbf{C}_Y^{(1,1)}$ has been subtracted. If not (e.g., due to an unknown colour of the noise), then the equation related to $\mathbf{C}_Y^{(1,1)}$ can be dropped or given a little weight.

(The problem is reduced to the one in Section 4 by picking one of the subequations of (22) (and possibly (23)) and giving them an infinite weight.) For a set of equations like (22), an algorithm has been proposed in [11]. This technique can easily be adapted to take (23) into account as well. One only has to make sure that in the Z-steps of the extended QZ iteration, the contributions related to (23) are complex conjugated. An alternative scheme is proposed in [7].

6 DIMENSIONALITY REDUCTION

Let us assume that $J > I$, that $\mathbf{F}_p \in \mathbb{C}^{J\times J}$ is the equivalent of $\mathbf{B}_p$ ($1 \leqslant p \leqslant P$) and $\tilde{\mathbf{F}}_{\tilde{p}} \in \mathbb{C}^{J\times J}$ the equivalent of $\tilde{\mathbf{B}}_{\tilde{p}}$ ($1 \leqslant \tilde{p} \leqslant \tilde{P}$). From (22) and (23) it is clear that the column space of $\mathbf{F}_p$ and $\tilde{\mathbf{F}}_{\tilde{p}}$ is equal to the column space of $\mathbf{M}$. This vector space can be estimated as the space generated by the dominant left singular vectors of

$$\left(\mathbf{F}_1\ldots\mathbf{F}_P\,\tilde{\mathbf{F}}_1\ldots\tilde{\mathbf{F}}_{\tilde{P}}\right),\tag{24}$$

and I itself can be determined by looking for a gap in the singular value spectrum. If the dominant subspace is represented by $\mathbf{X} \in \mathbb{C}^{J\times I}$ with orthonormal columns, then a dimensionality reduction can be realized by taking $\mathbf{B}_p = \mathbf{X}^H\mathbf{F}_p\mathbf{X}$ and $\tilde{\mathbf{B}}_{\tilde{p}} = \mathbf{X}^H\tilde{\mathbf{F}}_{\tilde{p}}\mathbf{X}^*$. However, when resorting to estimates $\hat{\mathbf{C}}_Y^{(2,0)}$, $\hat{\mathbf{C}}'^{(1,1)}_Y$ (a noise-compensated version of $\hat{\mathbf{C}}_Y^{(1,1)}$), $\hat{\mathcal{C}}_Y^{(2,2)}$, $\hat{\mathcal{C}}_Y^{(3,1)}$ and $\hat{\mathcal{C}}_Y^{(4,0)}$, this is in principle not equivalent to maximization of

$$\begin{aligned} f(\mathbf{X}) \stackrel{\text{def}}{=}\; & w_1^2\,\|\mathbf{X}^H\hat{\mathbf{C}}'^{(1,1)}_Y\mathbf{X}\|^2 + w_2^2\,\|\mathbf{X}^H\hat{\mathbf{C}}_Y^{(2,0)}\mathbf{X}^*\|^2\\ &+w_3^2\,\|\hat{\mathcal{C}}_Y^{(2,2)}\times_1\mathbf{X}^H\times_2\mathbf{X}^T\times_3\mathbf{X}^H\times_4\mathbf{X}^T\|^2\\ &+w_4^2\,\|\hat{\mathcal{C}}_Y^{(3,1)}\times_1\mathbf{X}^H\times_2\mathbf{X}^H\times_3\mathbf{X}^H\times_4\mathbf{X}^T\|^2\\ &+w_5^2\,\|\hat{\mathcal{C}}_Y^{(4,0)}\times_1\mathbf{X}^H\times_2\mathbf{X}^H\times_3\mathbf{X}^H\times_4\mathbf{X}^H\|^2, \end{aligned}\tag{25}$$

although the two approaches are usually close. One could, e.g., start an Alternating Least Squares (ALS) iteration, in analogy with [5]. Typically, in iteration step k a column-wise orthonormal matrix $\mathbf{X}^{(k)}$ is calculated of which the column space is equal to the space generated by the dominant left singular vectors of the matrix containing all the columns of $w_1\,\hat{\mathbf{C}}'^{(1,1)}_Y\mathbf{X}^{(k-1)}$, $w_2\,\hat{\mathbf{C}}_Y^{(2,0)}\mathbf{X}^{(k-1)^*}$, $w_3\,\hat{\mathcal{C}}_Y^{(2,2)}\times_2\mathbf{X}^{(k-1)^T}\times_3\mathbf{X}^{(k-2)^H}\times_4\mathbf{X}^{(k-3)^T}$, $w_4\,\hat{\mathcal{C}}_Y^{(3,1)}\times_2\mathbf{X}^{(k-1)^H}\times_3\mathbf{X}^{(k-2)^H}\times_4\mathbf{X}^{(k-3)^T}$ and $w_5\,\hat{\mathcal{C}}_Y^{(4,0)}\times_2\mathbf{X}^{(k-1)^H}\times_3\mathbf{X}^{(k-2)^H}\times_4\mathbf{X}^{(k-3)^H}$.

7 CONCLUSION

For non-circular random variables more statistics are available than for circular random variables. In this paper we have exploited this extra knowledge in the context of BSS. Different approaches were derived, depending on the relative importance of the statistics.

References

[1] J.F. Cardoso, A. Souloumiac, "Blind Beamforming for Non-Gaussian Signals", *IEE Proc.-F*, Vol. 140, No. 6, 1994, pp. 362–370.

[2] J.F. Cardoso, "Perturbation of Joint Diagonalizers", Tech. Report No. 94d027, ENST Paris, 1994.

[3] P. Comon, "Independent Component Analysis, A New Concept?" *Signal Processing*, Vol. 36, No. 3, April 1994, pp. 287–314.

[4] L. De Lathauwer, B. De Moor, J. Vandewalle, "ICA Techniques for More Sources than Sensors", *Proc. IEEE Signal Processing Workshop on Higher-Order Statistics (HOS'99)*, Caesarea, Israel, June 14-16, 1999, pp. 121–124.

[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] L. De Lathauwer, B. De Moor, J. Vandewalle, "Independent Component Analysis and (Simultaneous) Third-Order Tensor Diagonalization", *IEEE Trans. Signal Processing*, Vol. 49, No. 10, Oct. 2001, pp. 2262–2271.

[7] L. De Lathauwer, B. De Moor, J. Vandewalle, "An Algorithm for Joint Diagonalization by a Congruence Transformation", *Proc. 23rd Symp. on Information Theory in the Benelux*, May 29–31, 2002, Louvain-la-Neuve, Belgium.

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

[9] R.A. Horn, C.R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.

[10] A. Souloumiac, E. Moreau, "Blind Source Separation via Joint Diagonalization of Symmetric Complex Matrices", *Proc. 2002 IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP 2002)*, Orlando, FL, USA, May 13-17, 2002.

[11] A.J. van der Veen, A. Paulraj, "An Analytical Constant Modulus Algorithm", *IEEE Trans. Signal Processing*, Vol. 44, No. 5, May 1996, pp. 1136–1155.