IEEE Sig. Proc. Workshop on Higher-Order Statistics, June 14-16, 1999, Caesarea, Israel, pp.116–120

ICA algorithms for 3 sources and 2 sensors

L. De Lathauwer[1] P. Comon[2] B. De Moor[1] J. Vandewalle[1]

(1) K.U.LEUVEN, E.E.Dept. (ESAT), SISTA/COSIC,
Kard. Mercierlaan 94, B-3001 Leuven (Heverlee), Belgium
Lieven.DeLathauwer@esat.kuleuven.ac.be

(2) Lab. I3S, 2000 route des Lucioles,
Sophia-Antipolis, F-06410 Biot, France
comon@i3s.unice.fr

Abstract

In this paper we develop efficient algorithms to identify the mixing matrix in the context of an Independent Component Analysis with 3 sources and 2 sensors. Our contribution is two-fold. First, by applying twice a theorem by Sylvester, it is shown that the mixture can be obtained analytically by solving a linear system involving cumulants of both orders 3 and 4. Secondly, the latter theorem is extended to the case of "polynomials" of complex variables in which each monomial counts the same number of complex conjugated unknowns; this leads to an algorithm allowing to identify the mixture by solving a linear system involving only 4th order cumulants.

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, the source vector X and the noise vector N are zero-mean random vectors with values in $\mathbb{C}$ (or $\mathbb{R}$). The components of X are mutually statistically independent, as well as statistically independent from the noise components. The goal of ICA consists of the estimation of the "mixing matrix" $\mathbf{M}$ and the corresponding realizations of X, given only realizations of Y.

In this paper we focus on the situation in which there are 3 sources but only 2 sensor elements. We will show that in this set-up the mixing matrix can be estimated accurately at a low computational cost, which makes it attractive for e.g. mobile communications. The rectangular size of the mixing matrix makes it of course impossible to estimate the source values by mere matrix inversion, but the source values may nevertheless be reconstructed by exploiting prior knowledge on the source distributions, like e.g. the finite alphabet property [3].

Our approach is based on multilinear algebra. In this framework the ICA problem can be formulated as the decomposition of super-symmetric higher-order cumulant tensors in a minimal number of rank-1 terms, in which each rank-1 term consists of the contribution of one particular source signal to the overall linear combination. Higher-order tensors have the advantage, compared to matrices, that the number of rank-1 terms is not bounded by the dimension of the column nor the row space, which makes it a priori possible to address the ICA problem in a context where there are more sources than sensors.

The algorithms of this paper will be derived by exploring an old theorem proposed by Sylvester. First we will introduce some basic material in Section 2. Section 3 states Sylvester's theorem. The actual application to ICA is discussed in Section 4.

2. Tensors and quantics

In this paper, a *higher-order tensor* can intuitively be imagined as a multi-indexed array of numerical values. First we touch on some basic definitions in multilinear algebra.

Definition 1 (Outer product) *The outer product $\mathcal{A} \circ \mathcal{B}$ of a tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \ldots \times I_P}$ and a tensor $\mathcal{B} \in \mathbb{C}^{J_1 \times J_2 \times \ldots \times J_Q}$, is defined by*

$$(\mathcal{A} \circ \mathcal{B})_{i_1 i_2 \ldots i_P j_1 j_2 \ldots j_Q} \stackrel{\mathrm{def}}{=} a_{i_1 i_2 \ldots i_P} b_{j_1 j_2 \ldots j_Q},$$

for all values of the indices.

This product generalizes the construction $\mathbf{C} = U\,V^T$, in which U and V are vectors, in matrix algebra. As an example, an Nth-order tensor $\mathcal{A}$, equal to the outer product of N vectors $U^{(1)}$, $U^{(2)}$, ..., $U^{(N)}$, is given by the element-wise equation $a_{i_1 i_2 \ldots i_N} = u^{(1)}_{i_1} u^{(2)}_{i_2} \ldots u^{(N)}_{i_N}$. Such a tensor will be called a rank-1 tensor:

Definition 2 (Rank-1 tensor) *An Nth-order tensor $\mathcal{A}$ has rank 1 when it equals the outer product of N vectors $U^{(1)}$, $U^{(2)}$, ..., $U^{(N)}$:*

$$\mathcal{A} = U^{(1)} \circ U^{(2)} \circ \ldots \circ U^{(N)}.$$

Definition 3 (Rank) *The rank of an arbitrary Nth-order tensor $\mathcal{A}$, represented by $R = \texttt{rank}(\mathcal{A})$, is the minimal number of rank-1 tensors that yield $\mathcal{A}$ in a linear combination.*

In multilinear algebra the concept of rank is not subject to the same limitations as in the matrix case. The basis of this paper is the fact that the rank of a $(K \times K \times \ldots \times K)$-tensor is not bounded by K in general.

Next, we provide straightforward generalizations of the definitions of scalar product and orthogonality.

Definition 4 (Scalar product) *The scalar product $\langle \mathcal{A}, \mathcal{B} \rangle$ of two tensors $\mathcal{A}, \mathcal{B} \in \mathbb{C}^{I_1 \times I_2 \times \ldots \times I_N}$ is defined as*

$$\langle \mathcal{A}, \mathcal{B} \rangle \stackrel{\text{def}}{=} \sum_{i_1} \sum_{i_2} \cdots \sum_{i_N} b^*_{i_1 i_2 \ldots i_N} a_{i_1 i_2 \ldots i_N}.$$

Definition 5 (Orthogonality) *Tensors of which the scalar product equals 0, are mutually orthogonal.*

A *quantic* is an homogeneous polynomial. One can associate a quantic $f(x_1, x_2, \ldots, x_K)$ of degree N with a given Nth-order $(K \times K \times \ldots \times K)$-tensor $\mathcal{F}$ as follows:

$$f(x_1, x_2, \ldots, x_K) = \sum_{k_1 k_2 \ldots k_N = 1}^{K} f_{k_1 k_2 \ldots k_N} x_{k_1} x_{k_2} \ldots x_{k_N}. \qquad (2)$$

Conversely, one can also associate a tensor with a given quantic. To make the representation unique we may assume that the tensor is invariant under arbitrary index permutations. For binary quantics, set $x_1 = x$ and $x_2 = y$ to simplify the notation. Then, the link amounts to

$$f(x, y) = \sum_{n=0}^{N} \binom{N}{n} \, f_{\mathbf{k}} \, x^{N-n} \, y^n, \qquad (3)$$

in which the multi-index $\mathbf{k}$ contains $N - n$ times the value 1 and n times the value 2, and $\binom{N}{n} = \frac{N!}{(N-n)!\,n!}$.

We now define the apolar scalar product of quantics; we restrict ourselves to polynomials in two unknowns, which is the case of interest for this paper.

Definition 6 (Apolar scalar product) *The apolar scalar product $\langle f, g \rangle$ of two binary quantics of degree N is given by*

$$\langle f, g \rangle \stackrel{\text{def}}{=} \sum_{n=0}^{N} \binom{N}{n} \, f_{\mathbf{k}} \, g^*_{\mathbf{k}}.$$

This product simply corresponds to the scalar product of the associated tensors; the variational coefficient takes into account the number of times a given entry appears due to the symmetry.

For complex tensors with a specific Hermitean-type symmetry it may be more appropriate to consider the complex conjugate of a number of unknowns in the right-hand side of Eq. (2). Eq. (3) and the expression for the apolar scalar product of binary quantics then take a more complicated form.

3. A theorem by Sylvester

The following theorem is an old result of multilinear algebra obtained by J.J. Sylvester [1]:

Theorem 1 *A binary quantic $f(x, y)$ of degree N,*

$$f(x, y) = \sum_{n=0}^{N} \phi_n \, \binom{N}{n} \, x^{N-n} \, y^n,$$

can be written as a sum of Nth powers of R distinct linear forms,

$$f(x, y) = \sum_{r=1}^{R} \lambda_r \, (\alpha_r \, x + \beta_r \, y)^N,$$

iff (i) there exists a vector $G \in \mathbb{C}^{R+1}$, with components g_r $(0 \leqslant r \leqslant R)$, such that

$$\begin{bmatrix} \phi_0 & \phi_1 & \cdots & \phi_R \\ \phi_1 & \phi_2 & \cdots & \phi_{R+1} \\ \vdots & \vdots & & \vdots \\ \phi_{N-R} & \phi_{N-R+1} & \cdots & \phi_N \end{bmatrix} G^* = 0, \qquad (4)$$

and (ii) the polynomial $g(x, y) \stackrel{\text{def}}{=} \sum_{r=0}^{R} g_r \, x^{R-r} \, y^r$ admits R distinct roots, $f(x, y)$ and $g(x, y)$ being related in the sense that $g(x, y) = \prod_{r=1}^{R} (\beta^*_r \, x - \alpha^*_r \, y)$.

Proof. Call $X_r = (\alpha_r \; \beta_r)^T$ and $Y_r = (\beta^*_r \; -\alpha^*_r)^T$. The quantics $f(x, y)$ and $g(x, y)$ can be associated with the Nth order tensor $\mathcal{F} \stackrel{\text{def}}{=} \sum_{r=1}^{R} \lambda_r X_r \circ X_r \circ \cdots \circ X_r$ and the Rth order tensor $\mathcal{G} \stackrel{\text{def}}{=} Y_1 \circ Y_2 \circ \cdots \circ Y_R$ respectively. Eq. (4) is an apolar scalar product formulation of the fact that $\mathcal{G}$ is orthogonal to the $N - R + 1$ different Rth order tensors that can be obtained from $\mathcal{F}$ by keeping $N - R$ arbitrary indices fixed. ∎

Sylvester's theorem not only proves the existence of the R forms, but also gives a means to compute them. In fact, given the set of coefficients $\{\phi_n\}$, it is always possible to find the vector G from (4), and then deduce the forms from the roots of the associated polynomial $g(x, y)$. More precisely, if R is unknown, one starts with an $(N - 1 \times 2)$ Hankel matrix, and one assumes $R = 1$. If this matrix is full column rank, one goes to $R = 2$ and tests the rank of the $(N - 2 \times 3)$ Hankel matrix, and so forth. At some point, the number of columns exceeds the number of rows, and the algorithm stops. The generic rank R is obtained precisely at this stage, when $2R > N$. For odd values of N, we have thus a generic rank of $R = \frac{N+1}{2}$, whereas for even values of N, $R = \frac{N}{2} + 1$, generically. It is then clear that when N is even, there are at least two vectors satisfying (4), because the Hankel matrix is of size $\frac{N}{2} \times (\frac{N}{2} + 2)$. To be more concrete, take as example $N = 4$. The first Hankel matrix having more columns than rows is of size 2×4, and obviously has generically 2 vectors in its null space [2]. As a conclusion, when N is odd, there is generically a unique vector G satisfying (4), but there are two of them when N is even.

4. Application in ICA

Due to the multilinearity of cumulants, $\mathcal{C}^{(4,0)} \stackrel{\text{def}}{=} \mathtt{Cum}\{Y, Y, Y, Y\}$ can be expanded as

$$\mathcal{C}^{(4,0)} = \sum_{r=1}^{R} \kappa_r^{(4,0)} M_r \circ M_r \circ M_r \circ M_r , \qquad (5)$$

in which $\kappa_r^{(4,0)} = \mathtt{Cum}\{X_r, X_r, X_r, X_r\}$ is the kurtosis of the rth source and M_r the rth column of the mixing matrix. Hence identifying the mixing matrix from $\mathcal{C}^{(4,0)}$ is equivalent to decomposing the quartic associated with it into a sum of fourth powers of R linear terms.

For $R = 2$, the problem consists of finding the kernel of the (3×3) Hankel matrix constructed as in Eq. (4). In the noiseless case, this matrix is rank deficient by definition. With noise, the set of equations is overdetermined and can be solved in a least-squares sense by computing the third right singular vector of the Hankel matrix.

For $R = 3$, the unknown G in Eq. (4) can only be restricted to a two-dimensional linear subspace. In order to fix this indeterminacy, one can resort to another cumulant, which should admit a related decomposition [2]. Up to complex conjugation, we have the following different cumulants for orders three and four:

$$\begin{aligned}
\mathcal{C}^{(3,0)} &\stackrel{\text{def}}{=} \mathtt{Cum}\{Y, Y, Y\} \\
&= \sum_{r=1}^{R} \gamma_r^{(3,0)} M_r \circ M_r \circ M_r , && (6) \\
\mathcal{C}^{(2,1)} &\stackrel{\text{def}}{=} \mathtt{Cum}\{Y, Y, Y^*\} \\
&= \sum_{r=1}^{R} \gamma_r^{(2,1)} M_r \circ M_r \circ M_r^* , && (7) \\
\mathcal{C}^{(4,0)} &\stackrel{\text{def}}{=} \mathtt{Cum}\{Y, Y, Y, Y\} \\
&= \sum_{r=1}^{R} \kappa_r^{(4,0)} M_r \circ M_r \circ M_r \circ M_r , && (8) \\
\mathcal{C}^{(3,1)} &\stackrel{\text{def}}{=} \mathtt{Cum}\{Y, Y, Y, Y^*\} \\
&= \sum_{r=1}^{R} \kappa_r^{(3,1)} M_r \circ M_r \circ M_r \circ M_r^* , && (9) \\
\mathcal{C}^{(2,2)} &\stackrel{\text{def}}{=} \mathtt{Cum}\{Y, Y, Y^*, Y^*\} \\
&= \sum_{r=1}^{R} \kappa_r^{(3,1)} M_r \circ M_r \circ M_r^* \circ M_r^* . && (10)
\end{aligned}$$

When the sources are symmetrically distributed about the origin, $\mathcal{C}^{(3,0)}$, $\mathcal{C}^{(2,1)}$ and $\mathcal{C}^{(3,1)}$ vanish. In [2] it was proposed to perform an exhaustive search for the free complex parameter in Eq. (4) that yields the factors that give the best match of $\mathcal{C}^{(2,2)}$ in Eq. (10). This was formulated as a search in $\mathbb{R}^2$ to obtain an associated 35-dimensional vector that is closest to an associated linear 3-dimensional subspace of the 35-dimensional space.

On the other hand, in Section 5 we will show that, for sources that are *not* symmetrically distributed about the origin, the solution can be obtained in a non-iterative way.

In Section 6 we will formulate a modification of Theor. 1 for tensors that have a symmetry like $\mathcal{C}^{(2,2)}$. This will be exploited in Section 7 to obtain an algorithm that works for symmetric source distributions at a lower cost than the scheme derived in [2].

5. Unsymmetric source distributions

Let us assume that the sources have a probability density that is not symmetric about the origin. Call $M_r = (\alpha_r \, \beta_r)^T$ and $Y_r = (\beta_r^* \;\; -\alpha_r^*)^T$. For $R = 3$, the tensor $\mathcal{G} \stackrel{\text{def}}{=} Y_1 \circ Y_2 \circ Y_3$ is not only orthogonal to the two third-order tensors obtained from $\mathcal{C}^{(4,0)}$ by keeping the fourth index fixed, as reflected by the two equations of (4), but also to $\mathcal{C}^{(3,0)}$ and to the third-order tensors obtained from $\mathcal{C}^{(3,1)}$ by keeping the fourth index fixed. Consequently, G may be obtained from the following overdetermined system of equations:

$$\begin{bmatrix}
c_{1111}^{(4,0)} & c_{2111}^{(4,0)} & c_{2211}^{(4,0)} & c_{2221}^{(4,0)} \\
c_{1112}^{(4,0)} & c_{1122}^{(4,0)} & c_{1222}^{(4,0)} & c_{2222}^{(4,0)} \\
c_{111}^{(3,0)} & c_{112}^{(3,0)} & c_{122}^{(3,0)} & c_{222}^{(3,0)} \\
c_{1111}^{(3,1)} & c_{2111}^{(3,1)} & c_{2211}^{(3,1)} & c_{2221}^{(3,1)} \\
c_{1112}^{(3,1)} & c_{1122}^{(3,1)} & c_{1222}^{(3,1)} & c_{2222}^{(3,1)}
\end{bmatrix} G^* = 0, \qquad (11)$$

in which the equations may be weighted inversely proportionally to the variance of the cumulant estimates. Hence G can be estimated as the complex conjugate of the right singular vector that corresponds to the smallest singular value of the matrix in (11), which essentially amounts to finding a root of a polynomial of degree 4; this can be done in a non-iterative way. Of course one can append additional equations to the system by considering cumulants of order higher than 4 as well; e.g. the sixth-order cumulant $\mathcal{C}^{(3,3)}$, defined in analogy to (6 - 10), may lead to 4 extra equations, corresponding to the number of independent third-order tensors that can be obtained from it by keeping the last 3 indices fixed.

6. Variations on a theme

Sylvester's theorem can be modified to take into account complex conjugates of the unknowns, as indicated in Section 2. As a matter of fact, this was already the principle on which the fourth and fifth equation in (11) were based. For convenience we restrict ourselves in this section to fourth-order tensors with the same symmetry as $\mathcal{C}^{(2,2)}$.

Theorem 2 *The binary quartic*[1]

$$f(x,y) = \sum_{m,n=0}^{2} \phi_{m,n} \binom{2}{m}\binom{2}{n} x^{2-m} (x^{2-n})^* y^m (y^n)^*$$

can be expanded as

$$f(x,y) = \sum_{r=1}^{R} \lambda_r |\alpha_r x + \beta_r y|^4,$$

- *with $R = 1$, iff there exists a vector $G = (g_0\, g_1)^T$, such that*

$$\begin{bmatrix} \phi_{00} & \phi_{10} \\ \phi_{10} & \phi_{20} \\ \phi_{01} & \phi_{11} \\ \phi_{11} & \phi_{21} \\ \phi_{02} & \phi_{12} \\ \phi_{12} & \phi_{22} \end{bmatrix} G^* = 0. \tag{12}$$

  The linear form $g(x,y) \stackrel{\mathrm{def}}{=} g_0 y + g_1 x$ equals $\beta_1^* x - \alpha_1^* y$.

- *with $R = 2$, iff (i) there exists a vector $G \in \mathbb{C}^3$, with components g_r $(0 \leqslant r \leqslant 2)$, such that*

$$\begin{bmatrix} \phi_{00} & \phi_{10} & \phi_{20} \\ \phi_{01} & \phi_{11} & \phi_{21} \\ \phi_{02} & \phi_{12} & \phi_{22} \end{bmatrix} G^* = 0, \tag{13}$$

  and (ii) the polynomial $g(x,y) \stackrel{\mathrm{def}}{=} \sum_{r=0}^{2} g_r x^{2-r} y^r$ admits 2 distinct roots. $f(x,y)$ and $g(x,y)$ are related in the sense that $g(x,y) = \prod_{r=1}^{2}(\beta_r^* x - \alpha_r^* y)$.

- *with $R = 3$, iff (i) there exists a vector $G \in \mathbb{C}^6$ such that*

$$\begin{bmatrix} \phi_{00} & \phi_{10} & \phi_{20} & \phi_{01} & \phi_{11} & \phi_{21} \\ \phi_{01} & \phi_{11} & \phi_{21} & \phi_{02} & \phi_{12} & \phi_{22} \end{bmatrix} G^* = 0, \tag{14}$$

  with $G = G'^* \otimes G''$, in which $G' \in \mathbb{C}^2$ and $G'' \in \mathbb{C}^3$. Call the components of G' and G'' g'_r $(0 \leqslant r \leqslant 1)$ and g''_r $(0 \leqslant r \leqslant 2)$ respectively. (ii) The roots of $g(x,y) \stackrel{\mathrm{def}}{=} g'(x,y)\, g''(x,y)$ are mutually different. $f(x,y)$ and $g(x,y)$ are related in the sense that $g(x,y) = \prod_{r=1}^{3}(\beta_r^* x - \alpha_r^* y)$.

- *with $R = 4$, iff (i) there exist vectors $G', G'' \in \mathbb{C}^3$, with components g'_r and g''_r respectively $(0 \leqslant r \leqslant 2)$, such that*

$$G'^H \begin{bmatrix} \phi_{00} & \phi_{01} & \phi_{02} \\ \phi_{10} & \phi_{11} & \phi_{12} \\ \phi_{20} & \phi_{21} & \phi_{22} \end{bmatrix} G'' = 0, \tag{15}$$

  and (ii) the polynomial $g(x,y) \stackrel{\mathrm{def}}{=} g'(x,y)\, g''(x,y)$ admits 4 distinct roots. $f(x,y)$ and $g(x,y)$ are related in the sense that $g(x,y) = \prod_{r=1}^{4}(\beta_r^* x - \alpha_r^* y)$.

Proof. Call $X_r = (\alpha_r\, \beta_r)^T$ and $Y_r = (\beta_r^* \; -\alpha_r^*)^T$. The quantic $f(x,y)$ can be associated with the 4th order tensor $\mathcal{F} \stackrel{\mathrm{def}}{=} \sum_{r=1}^{R} \lambda_r X_r \circ X_r \circ X_r^* \circ X_r^*$. Eq. (12) is an apolar scalar product formulation of the fact that $G = Y_1$ is orthogonal to the 6 different vectors that can be obtained from $\mathcal{F}$ by varying only the first index. Eq. (13) is equivalent to the fact that the matrix $\mathbf{G} \stackrel{\mathrm{def}}{=} Y_1 Y_2^T$ is orthogonal to the 3 different matrices that can be obtained from $\mathcal{F}$ by keeping the third and the fourth index fixed. Eq. (14) expresses that $Y_1^* \circ Y_2 \circ Y_3$, $Y_2^* \circ Y_3 \circ Y_1$ and $Y_3^* \circ Y_1 \circ Y_2$ are orthogonal to the 2 third-order tensors that can be obtained from $\mathcal{F}$ by keeping the fourth index fixed. Eq. (15) means that $Y_1 \circ Y_2 \circ Y_3^* \circ Y_4^*$ is orthogonal to $\mathcal{F}$ (also valid when the subscripts are permuted). ■

7. Symmetric source distributions

7.1. Method 1

Consider $\mathcal{C}^{(4,0)}$ and $\mathcal{C}^{(2,2)}$ as linear matrix-to-matrix mappings, according to $(\mathcal{C}(\mathbf{A}))_{ij} = \sum_{p,q=1}^{2} c_{ijpq} a_{pq}$, and let $\mathbf{C}^{(4,0)}, \mathbf{C}^{(2,2)} \in \mathbb{C}^{4\times 4}$ be the corresponding matrix representations. As explained in [5], the unknown mixing matrix can be estimated from the Eigenvalue Decomposition

$$\mathbf{C}^{(4,0)} \cdot (\mathbf{C}^{(2,2)})^{\dagger *} \cdot (\mathbf{C}^{(4,0)})^* \cdot (\mathbf{C}^{(2,2)})^{\dagger} = \tilde{\mathbf{M}} \cdot \mathbf{D} \cdot \tilde{\mathbf{M}}^{\dagger}, \tag{16}$$

iff there are no mutually collinear pairs among $(|\kappa_1^{(4,0)}|, |\kappa_1^{(2,2)}|)$, $(|\kappa_2^{(4,0)}|, |\kappa_2^{(2,2)}|)$ and $(|\kappa_3^{(4,0)}|, |\kappa_3^{(2,2)}|)$: under the zero noise assumption, the columns $\tilde{M}_r$ corresponding to non-vanishing eigenvalues, are equal to $M_r \otimes M_r$; rearranged in a matrix format we will represent the eigenvectors as $\tilde{\mathbf{M}}_r$. However these eigenvectors may be ill-conditioned, if eigenvalues are close. To increase the

[1] f is actually an homogeneous polynomial in four real variables. Even if it is not correct strictly speaking, it is sometimes referred to as a real "binary quartic" in complex variables.

accuracy, the conditions expressed by the first 2 equations of (11) can be incorporated as follows.

Construct a sixth-order tensor $\mathcal{A}$ according to

$$a_{ijklmn} \stackrel{\text{def}}{=} \sum_{\sigma} (\tilde{\mathbf{M}}_1)_{\sigma(i)j}\, (\tilde{\mathbf{M}}_2)_{\sigma(k)l}\, (\tilde{\mathbf{M}}_3)_{\sigma(m)n}\,,$$

in which the summation is taken over the different permutations of (i, j, k). Now consider $\mathcal{A}$ as a mapping from $\mathbb{C}^{2\times 2\times 2}$ to $\mathbb{C}^{2\times 2\times 2}$ by summation over the indices l, m, n. By construction, the tensor $\mathcal{G}$ in Section 5 is orthogonal to any tensor in the range of $\mathcal{A}$, and each of these tensors has the same symmetry as $\mathcal{C}^{(3,0)}$. Due to the symmetry, the rank of the mapping is bounded by 4. This means that 4 extra equations can be added to the first 2 equations of (11). Without noise the rank equals 1.

7.2. Method 2

In the previous subsection we only obtained one extra equation of the type (11) in the noiseless case. However it is possible to derive two independent conditions from $\mathcal{C}^{(2,2)}$. To understand this, let us have a closer look at Theor. 2. Let the kernel of Eq. (14) be represented by the full-rank matrix $\mathbf{K} \in \mathbb{C}^{6\times 4}$, and let this matrix be partitioned as $\mathbf{K}^T = (\mathbf{K}_1\ \mathbf{K}_2)^T$. A vector $G \stackrel{\text{def}}{=} \mathbf{K}\, V$, with $V \in \mathbb{C}^{4\times 1}$, is then of the form required in Theor. 2 iff there exist complex values λ_1 and λ_2 such that

$$(\lambda_1 \mathbf{K}_1 - \lambda_2 \mathbf{K}_2)\, V = 0. \tag{17}$$

It is clear that for any value of λ_1/λ_2 an appropriate vector V can be determined, such that $\mathcal{C}^{(2,2)}$, like $\mathcal{C}^{(4,0)}$, leaves one free complex parameter in the solution set. However this algebraic variety is not a linear vector space, which makes $\mathcal{C}^{(2,2)}$ harder to handle here. The solutions we are interested in, consist of $(\lambda_1\ \lambda_2)$ proportional to $(\alpha_{\tilde{r}}\ \beta_{\tilde{r}})$, while $g''(x, y) = \prod_{r\neq\tilde{r}} (\beta_r\, x - \alpha_r\, y)$ $(1 \leqslant \tilde{r} \leqslant 3)$; the vector G'' is proportional to $\mathbf{K}_1\, V$. The equations for $\mathcal{C}^{(4,0)}$ can then be expressed as

$$\begin{bmatrix} c^{(4,0)}_{1111} & c^{(4,0)}_{2111} & c^{(4,0)}_{2211} & c^{(4,0)}_{2221} \\ c^{(4,0)}_{1112} & c^{(4,0)}_{1122} & c^{(4,0)}_{1222} & c^{(4,0)}_{2222} \end{bmatrix} \cdot \begin{bmatrix} \lambda_2^* & 0 & 0 \\ \lambda_1^* & \lambda_2^* & 0 \\ 0 & \lambda_1^* & \lambda_2^* \\ 0 & 0 & \lambda_1^* \end{bmatrix} \cdot \mathbf{K}_1 \cdot V = 0. \tag{18}$$

Combination of Eqs. (17-18) shows that λ_1/λ_2 should be calculated as the value that minimizes the smallest singular value of the (5×4)-matrix obtained by stacking the equations; this can be implemented by using a general-purpose optimization routine. The search can be initialized by the solution found in the previous subsection. It is clear that $\mathcal{C}^{(3,0)}$, $\mathcal{C}^{(2,1)}$ and $\mathcal{C}^{(3,1)}$, or cumulants of order higher than 4, can be taken into account by simply adding more equations to the set.

8. Conclusion

For unsymmetric source distributions a (2×3) mixing matrix can be identified by solving an homogeneous linear system of equations, followed by the rooting of a polynomial of degree 3. For symmetric source distributions the mixing matrix can be estimated in a non-iterative way as well. For the latter case we have also developed a fine-tuning iteration that takes into account all the available information and is computationally cheaper than the scheme proposed earlier in [2].

Acknowledgements

This research was partially supported by the Flemish Government (the Concerted Research Action GOA-MIPS, the FWO (Fund for Scientific Research - Flanders) project G.0240.99, the FWO Research Community Advanced Numerical Methods for Mathematical Modelling, the IWT Action Programme on Information Technology ITA/GBO/T23), the Belgian State, Prime Minister's Office - Federal Office for Scientific, Technical and Cultural Affairs (the Interuniversity Poles of Attraction Programmes IUAP P4-02 and IUAP P4-24) and the French CNRS Telecommunications Program TL 97104. The scientific responsibility is assumed by the authors.

References

[1] P. Comon and B. Mourrain. Decomposition of quantics in sums of powers of linear forms. *Signal Processing*, 53(2-3): 93-108, 1996.

[2] P. Comon. Blind channel identification and extraction of more sources than sensors. *Proc. SPIE Conf.*, 2-13, San Diego, CA, U.S.A., July 19-24, 1998.

[3] P. Comon and O. Grellier. Non-linear inversion of underdetermined mixtures. *Proc. ICA'99*, 461-465, Aussois, France, Jan. 11-15, 1999.

[4] L. De Lathauwer. *Signal Processing based on Multilinear Algebra*. Ph.D. thesis, K.U.Leuven, E.E.Dept. (ESAT), Belgium, Sept. 1997.

[5] L. De Lathauwer, B. De Moor and J. Vandewalle. ICA techniques for more sources than sensors. *Proc. HOS'99*.