Katholieke Universiteit Leuven

Departement Elektrotechniek ESAT-SISTA/TR 2001-66

An algebraic algorithm for ICA with more sources than sensors[1]

Lieven De Lathauwer, Bart De Moor and Joos Vandewalle[2]

July 2001

Published in:
J. McWhirter, I. Proudler (Eds.),
Mathematics in Signal Processing V,
The Institute of Mathematics & Its Applications
Conference Series, No. 71,
Clarendon Press, Oxford, 2002, pp. 37–46.

[1] This report is available by anonymous ftp from *ftp.esat.kuleuven.ac.be* in the directory *pub/SISTA/delathauwer/reports/ldl-01-66.ps.gZ*

[2] K.U.Leuven, Dept. of Electrical Engineering (ESAT), Research group SISTA/COSIC, Kardinaal Mercierlaan 94, 3001 Leuven, Belgium, Tel. 32/16/32 11 04, Fax 32/16/32 19 70, WWW: *http://www.esat.kuleuven.ac.be/sista.* E-mail: *Lieven.DeLathauwer@esat.kuleuven.ac.be*, *Bart.DeMoor@esat.kuleuven.ac.be*, *Joos.Vandewalle@esat.kuleuven.ac.be*. This research was partially supported by the Flemish Government: (1) Research Council K.U.Leuven: Concerted Research Action GOA-MEFISTO-666, IDO, (2) the Fund for Scientific Research-Flanders (F.W.O.) projects G.0240.99 and G.0256.97, (3) the F.W.O. Research Communities ICCoS and ANMMM, (4) BIL 98 with South Africa and by 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. 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. The scientific responsibility is assumed by the authors.

Abstract

In this paper we derive an algorithm to identify the mixing matrix in the context of an independent component analysis with more sources than sensors. The technique can cope with $I(I+1)/2$ sources for only I sensors. It exploits the fact that for non-circular complex-valued data, depending on the type of complex symmetry, 2 different fourth-order cumulants are available.

1

An algebraic algorithm for independent component analysis with more sources than sensors

L. De Lathauwer*, **B. De Moor and J. Vandewalle****

Equipe Traitement des Images et du Signal, Ecole Nationale Supérieure de l'Electronique et de ses Applications and Université de Cergy-Pontoise, France,
** *Department of Electrical Engineering, Katholieke Universiteit Leuven, Belgium*

Abstract

In this paper we derive an algorithm to identify the mixing matrix in the context of an independent component analysis with more sources than sensors. The technique can cope with $I(I+1)/2$ sources for only I sensors. It exploits the fact that for non-circular complex-valued data, depending on the type of complex symmetry, 2 different fourth-order cumulants are available.

1 Introduction

Assume the following basic statistical model:

$$Y = \mathbf{M}X + N, \tag{1.1}$$

in which $Y \in \mathbb{C}^I$ is referred to as the *observation vector*, $X \in \mathbb{C}^R$ is called the *source vector* and $N \in \mathbb{C}^I$ represents additive *noise*. $\mathbf{M} \in \mathbb{C}^{I\times R}$ is the *mixing matrix*. The goal of independent component analysis now consists of the estimation of the mixing matrix and/or the corresponding realizations of the source vector X, given only realizations of the observation vector Y. The key assumption is that the components of X are mutually statistically independent, as well as statistically independent from the noise components. This is a very strong hypothesis, but also quite natural in lots of applications. It means that the aim can often be rephrased as splitting the dataset into components "of a different nature", which contributed to the data in a linear way.

The solution of an independent component analysis is subject to two basic indeterminacies. First, it is impossible to determine the norm of the columns of $\mathbf{M}$ in eqn (1.1), since a rescaling of these vectors can be compensated by the inverse scaling of the source signal values. Similarly, the ordering of the source signals, having no physical meaning, cannot be identified.

Uniqueness is usually guaranteed by claiming that the columns of $\mathbf{M}$ are linearly independent and that the sources are non-Gaussian; it can be shown that, under these conditions, the two basic indeterminacies are indeed the only way in which the solution is not unique 1; 2. However, in this paper we focus on the situation in which there are more sources than sensors, i.e., the mixing matrix has more columns than rows. Hence, in Section 4, we have to follow an other approach. For non-circular complex-valued data (like for instance in telecom applications) the condition of linear independence will be replaced by a weaker condition.

Independent component analysis is related to the classical problem of principal component analysis. Here, by exploiting that the source signals are uncorrelated, one is able to find the sources as well as the mixing matrix up to a unitary transformation. Let us define the covariance matrix $\mathbf{C}_Z^{(1,1)}$ of a random vector Z by the element-wise equation

$$(\mathbf{C}_Z^{(1,1)})_{i_1 i_2} = \mathtt{E}\{(z_{i_1} - \mathtt{E}\{z_{i_1}\})(z_{i_2} - \mathtt{E}\{z_{i_2}\})^*\}, \tag{1.2}$$

in which $\cdot^*$ denotes a complex conjugation and $\mathtt{E}$ the expectation. The covariance matrices of Y and X are related by

$$\mathbf{C}_Y^{(1,1)} = \mathbf{M} \cdot \mathbf{C}_X^{(1,1)} \cdot \mathbf{M}^H + \mathbf{C}_N^{(1,1)}, \tag{1.3}$$

in which $\mathbf{C}_X^{(1,1)}$ is positive diagonal ($\cdot^H$ denotes the complex conjugated transpose). Since the sources may be arbitrarily rescaled, we can assume that they have unit variance. Then we have (we omit the noise term at this point, for clarity):

$$\mathbf{C}_Y^{(1,1)} = \mathbf{M} \cdot \mathbf{M}^H. \tag{1.4}$$

Hence the mixing matrix can be estimated, up to a unitary transformation, as a square root of the observed covariance. In principal component analysis the solution is made essentially unique by selecting a matrix of which the columns are orthonormal. (This factor may of course also be found, in a numerically more reliable way, from the singular value decomposition of the observed dataset 3.)

The solution to the independent component analysis problem lies in the fact that the assumption of *statistical independence* is stronger than the notion of *uncorrelated* signals. Statistical independence is not only a claim on the second-order statistics of the signals, but also on their higher-order statistics 4. More precisely, it is not sufficient that the source covariance $\mathbf{C}_X^{(1,1)}$ is a diagonal matrix — in addition, the higher-order cumulants of the source vector should be diagonal higher-order tensors. (A higher-order tensor can intuitively be imagined as a multi-way matrix, of which the entries are characterized by more than two indices; its diagonal is defined as the entries for which all the indices are equal; the definition of a higher-order cumulant will be given in Section 2.) This leads to a sufficient amount of constraints to solve the problem. Actually, it is not necessary to resort to the second-order information to be able to find the solution:

the higher-order statistics alone contain enough information. The advantage of the latter approach is that one is able to work in a way that is conceptually blind to the presence of additive Gaussian noise, as we will explain in Section 2; the technique derived in this paper belongs to this class.

After a short introduction to higher-order cumulants in Section 2, we will show in Section 3 that our problem can be formulated as the simultaneous decomposition of two 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. In Section 4 we will consider the two cumulant tensors as the tensor equivalent of a matrix pencil; by associating an eigenvalue decomposition with it, a mixing matrix with $I(I+1)/2$ columns can be identified. For moderate values of I, this number is higher than for any other method published so far. Nevertheless, our contribution should be considered as a new line of approach that can be further improved (as we will explain in Section 4), rather than as an end result. Therefore we found it unnecessary to add a section on numerical simulations at this stage. In Section 5 we go into some more detail w.r.t. a numerical aspect of the technique.

Once the mixing matrix has been estimated, it is common to estimate the source values by premultiplying the dataset with its pseudo-inverse. However the rectangular size of the mixing matrix in our problem setting makes this approach impossible. Nevertheless, the source values may be reconstructed by exploiting prior knowledge on the source distributions, like e.g. the finite alphabet property 5; 6.

Let us end this introduction by referring to some related papers. In 7 the structure of a complex fourth-order cumulant is exploited to extract, for a generic mixing matrix, R columns, bounded by $2R(R-1) \leqslant (I(I-1))^2$. The special case of (2×3) independent component analysis was first addressed in 5. Further results can be found in 8. The idea of the current paper was first presented in 9; this paper also explains how 7 can be adapted to take two fourth-order cumulant tensors into account.

2 Cumulants

The basic quantities of higher-order statistics are higher-order moments and higher-order cumulants. Cumulants have a number of important properties, that are not shared with higher-order moments, such that in practice cumulants are more frequently used. In this section we will give a definition of cumulants up to order 4, and list some of the important properties. For a more comprehensive introduction we refer to 4; 10.

Definition 2.1. (Cumulant) *For a zero-mean random variables* x_1, x_2, x_3, x_4

the cumulants up to order 4 are defined by:

$$\texttt{Cum}(x_1) \stackrel{\text{def}}{=} \texttt{E}\{x_1\}, \tag{2.1}$$

$$\texttt{Cum}(x_1, x_2) \stackrel{\text{def}}{=} \texttt{E}\{x_1 x_2\}, \tag{2.2}$$

$$\texttt{Cum}(x_1, x_2, x_3) \stackrel{\text{def}}{=} \texttt{E}\{x_1 x_2 x_3\}, \tag{2.3}$$

$$\begin{aligned}\texttt{Cum}(x_1, x_2, x_3, x_4) \stackrel{\text{def}}{=} \; & \texttt{E}\{x_1 x_2 x_3 x_4\} \\ & - \texttt{E}\{x_1 x_2\}\texttt{E}\{x_3 x_4\} \\ & - \texttt{E}\{x_1 x_3\}\texttt{E}\{x_2 x_4\} \\ & - \texttt{E}\{x_1 x_4\}\texttt{E}\{x_2 x_3\}.\end{aligned} \tag{2.4}$$

For every variable x_i ($1 \leqslant i \leqslant 4$) that has a non-zero mean, x_i has to be replaced in these formulas, except eqn (2.1), by $x_i - \texttt{E}\{x_i\}$.
The cumulants of stochastic vectors $X^{(1)}$, ..., $X^{(K)}$ ($K \leqslant 4$) are defined by the element-wise equation

$$(\texttt{Cum}(X^{(1)}, \ldots, X^{(K)}))_{i_1, \ldots, i_K} \stackrel{\text{def}}{=} \texttt{Cum}(x^{(1)}_{i_1}, \ldots, x^{(K)}_{i_K}). \tag{2.5}$$

We have the following properties:

(1) *Multilinearity:* if the stochastic vectors $X^{(1)}$, ..., $X^{(K)}$ are transformed into the stochastic vectors $\tilde{X}^{(1)}$, ..., $\tilde{X}^{(K)}$ by the matrix multiplications $\tilde{X}^{(k)} = \mathbf{A}^{(k)} \cdot X^{(k)}$, with $\mathbf{A}^{(k)} \in \mathbb{C}^{J_k \times I_k}$ $(1 \leqslant k \leqslant K)$, then we have:

$$(\texttt{Cum}(\tilde{X}^{(1)}, \ldots, \tilde{X}^{(K)}))_{j_1, \ldots, j_K} = \sum_{i_1, \ldots, i_K} (\mathbf{A}^{(1)})_{j_1 i_1} \ldots (\mathbf{A}^{(K)})_{j_K i_K} (\texttt{Cum}(X^{(1)}, \ldots, X^{(K)}))_{i_1, \ldots, i_K}. \tag{2.6}$$

(2) *Partitioning of independent variables:* if a subset of K stochastic variables x_1, x_2, ..., x_K is independent of the other variables, then we have:

$$\texttt{Cum}(x_1, x_2, \ldots, x_K) = 0. \tag{2.7}$$

A consequence of this property is that, for a stochastic vector X having mutually independent components, the cumulant $\texttt{Cum}(X, \ldots, X)$ is a diagonal tensor. (The same holds if in the definition of the cumulant X is replaced a couple of times by X^*.) This very strong algebraic condition is the basis of all algebraic techniques for independent component analysis.

(3) *Sum of independent variables:* if the stochastic variables x_1, x_2, ..., x_K are mutually independent from the stochastic variables y_1, y_2, ..., y_K, then we have:

$$\begin{aligned}&\texttt{Cum}(x_1 + y_1, x_2 + y_2, \ldots, x_K + y_K) = \\ &\texttt{Cum}(x_1, x_2, \ldots, x_K) + \texttt{Cum}(y_1, y_2, \ldots, y_K).\end{aligned} \tag{2.8}$$

The cumulant tensor of a sum of independent random vectors is the sum of the individual cumulants. This property explains the term “cumulant”.

(4) *Non-Gaussianity:* higher-order cumulants of a Gaussian variable are 0. In combination with the multilinearity property, we conclude that higher-order cumulants have the interesting property to be blind for additive Gaussian noise. Cumulants of a random variable are some measure of its non-Gaussianity.

Generally speaking, it becomes harder to estimate higher-order statistics from sample data as the order increases, i.e., longer datasets are required to obtain the same accuracy 11; 12. Hence in practice the use of higher-order statistics is usually restricted to third- and fourth-order cumulants. In what follows, we will only consider fourth-order cumulants because in most applications the probability density functions of the random variables exhibit symmetries that cause the third-order cumulants to vanish.

3 Two expansions in rank-1 terms

We start from the observation that for non-circular complex-valued sources two types of fourth-order cumulants are available, even when the probability density is symmetric about the origin:

$$\begin{aligned} \mathcal{C}_Y^{(4,0)} &\stackrel{\text{def}}{=} \mathtt{Cum}(Y, Y, Y, Y), \\ \mathcal{C}_Y^{(2,2)} &\stackrel{\text{def}}{=} \mathtt{Cum}(Y, Y, Y^*, Y^*). \end{aligned}$$

Due to the multilinearity property, the observation cumulants are related to the source cumulants $\kappa_r^{(4,0)}$ and $\kappa_r^{(2,2)}$ $(1 \leqslant r \leqslant R)$ in the following way:

$$(\mathcal{C}_Y^{(4,0)})_{ijkl} = \sum_{r=1}^{R} \kappa_r^{(4,0)} \, m_{ir} m_{jr} m_{kr} m_{lr}, \tag{3.1}$$

$$(\mathcal{C}_Y^{(2,2)})_{ijkl} = \sum_{r=1}^{R} \kappa_r^{(2,2)} \, m_{ir} m_{jr} m_{kr}^* m_{lr}^*. \tag{3.2}$$

for all index values. Note that eqns (3.1) and (3.2) form a clear fourth-order counterpart of eqn (1.3). Moreover, there is no explicit noise contribution if the noise is assumed to be Gaussian. Of course, when dealing with sample cumulants, eqns (3.1) and (3.2) are only valid up to some perturbation.

Each of the terms in eqns (3.1) and (3.2) corresponds to a generalization of an outer product of two vectors; actually we are dealing with outer products of four vectors; in multilinear algebraic terminology we say that $\mathcal{C}_Y^{(4,0)}$ and $\mathcal{C}_Y^{(2,2)}$ are decomposed in a sum of fourth-order rank-1 terms. Higher-order tensors have the advantage, compared to matrices, that for the decomposition to be unique, 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 problem of independent component analysis in a context where the number of sources exceeds the number of sensors.

4 A generalized eigenvalue problem

Consider $\mathcal{C}^{(4,0)}$ and $\mathcal{C}^{(2,2)}$ as linear matrix-to-matrix mappings, according to

$$(\mathcal{C}(\mathbf{A}))_{ij} = \sum_{p,q=1}^{N} c_{ijpq} a_{pq},$$

and let $\mathbf{C}^{(4,0)}, \mathbf{C}^{(2,2)} \in \mathbb{C}^{N^2 \times N^2}$ be the corresponding matrix representations. In this format, eqns (3.1) and (3.2) can be expressed as follows:

$$\mathbf{C}_Y^{(4,0)} = \tilde{\mathbf{M}} \cdot \mathbf{C}_X^{(4,0)} \cdot \tilde{\mathbf{M}}^T, \tag{4.1}$$

$$\mathbf{C}_Y^{(2,2)} = \tilde{\mathbf{M}} \cdot \mathbf{C}_X^{(2,2)} \cdot \tilde{\mathbf{M}}^H, \tag{4.2}$$

in which $\mathbf{C}_X^{(4,0)}, \mathbf{C}_X^{(2,2)} \in \mathbb{C}^{N^2 \times N^2}$ are diagonal matrices, containing the source autocumulants, and in which $\tilde{\mathbf{M}} = \mathbf{M} \odot \mathbf{M}$, with $\odot$ the Kathri-Rao product (if $\mathbf{M} = [M_1 M_2 \dots M_R]$ then $\mathbf{M} \odot \mathbf{M} \stackrel{\text{def}}{=} [M_1 \otimes M_1\ M_2 \otimes M_2\ \dots\ M_R \otimes M_R]$, with $\otimes$ the Kronecker product).

One can associate a generalized eigenvalue problem with the "tensor pencil" $(\mathcal{C}_Y^{(4,0)}, \mathcal{C}_Y^{(2,2)})$ as follows:

$$\mathbf{C}_Y^{(4,0)} (\mathbf{C}_Y^{(2,2)})^{\dagger *} = \tilde{\mathbf{M}} \cdot \left(\mathbf{C}_X^{(4,0)} (\mathbf{C}_X^{(2,2)})^{\dagger *} \right) \cdot \tilde{\mathbf{M}}^{\dagger *}, \tag{4.3}$$

in which $\cdot^{\dagger}$ denotes the pseudo-inverse. This is the unsymmetric version of the Tagaki eigenvalue decomposition 13; as far as we know, it has not been studied in the literature yet.

The decomposition can be reduced to an ordinary eigenvalue decomposition by multiplying each side of the equation with its complex conjugate:

$$\mathbf{C}_Y^{(4,0)} \cdot (\mathbf{C}_Y^{(2,2)})^{\dagger *} \cdot (\mathbf{C}_Y^{(4,0)})^{*} \cdot (\mathbf{C}_Y^{(2,2)})^{\dagger} = \tilde{\mathbf{M}} \cdot \mathbf{D} \cdot \tilde{\mathbf{M}}^{\dagger}, \tag{4.4}$$

in which

$$\mathbf{D} \stackrel{\text{def}}{=} \mathbf{C}_X^{(4,0)} \cdot (\mathbf{C}_X^{(2,2)})^{\dagger *} \cdot (\mathbf{C}_X^{(4,0)})^{*} \cdot (\mathbf{C}_X^{(2,2)})^{\dagger} \tag{4.5}$$

is diagonal. If there are no mutually collinear pairs in the set $\{(|\kappa_r^{(4,0)}|, |\kappa_r^{(2,2)}|)\}_{1 \leqslant r \leqslant R}$ and if the matrices $\{M_r M_r^T\}_{1 \leqslant r \leqslant R}$ are linearly independent, then the eigenvalue decomposition (4.4) is unique. The eigenvectors correspond to $\{M_r \odot M_r\}_{1 \leqslant r \leqslant R}$, from which the columns of the mixing matrix may be derived. If the sample cumulants are perturbed by noise then the eigenvectors may not exactly correspond to (vectorized representations of) rank-1 matrices. Best rank-1 approximations can then be used to initialize an optimization routine in which both sides of eqns (3.1) and (3.2) are matched.

Taking into account that the range of the mapping in (4.4) only contains symmetric matrices, we conclude that $R = I(I+1)/2$ factors can be identified in this way.

For generic mixing matrices the multiplicity condition on the eigenvalues of eqn (4.4) is not crucial. The problem then becomes the determination of a set of rank-1 matrices, given the vector space they span. This problem is addressed in 7; 14. One can allow a multiplicity S, bounded by $2S(S-1) \leqslant (I(I-1))^2$; then the rank-1 matrices in the eigenspace can still be identified in an essentially unique way. One can use the same technique to deal with the possible ill-conditioning of the eigenvectors when eigenvalues are close.

Note that, apart from the remark in the preceding paragraph, we did not exploit the fact that the eigenvectors correspond to rank-1 matrices. Incorporating this knowledge will lead to improved techniques. This forms a topic of future research.

5 Complex eigenvalues

Equation 4.5 shows that the eigenvalues in decomposition (4.4) should be real. On the other hand, not every matrix can be decomposed as in eqn (4.3), which may lead to the presence of complex eigenvalues in eqn (4.4).

If the noise level is so high that complex eigenvalues do occur in eqn (4.4), then one can still proceed in an approximate sense as follows. Instead of the eigenvalue decomposition (4.4), let us consider the Schur decomposition

$$\mathbf{B} \cdot \mathbf{B}^* = \mathbf{Q} \cdot \mathbf{R} \cdot \mathbf{Q}^H, \tag{5.1}$$

in which $\mathbf{R}$ is upper triangular and $\mathbf{B}$ is a short-hand notation for $\mathbf{C}_Y^{(4,0)} \cdot (\mathbf{C}_Y^{(2,2)})^{\dagger *}$. From this unitary decomposition one can find a least-squares approximation of $\mathbf{BB}^*$ that has real eigenvalues, by simply setting the diagonal of $\mathbf{R}$ equal to its real part; we denote the result by $\tilde{\mathbf{R}}$.

In terms of the QR-decomposition $\tilde{\mathbf{M}} = \mathbf{QU}$, the remaining task in the estimation of $\tilde{\mathbf{M}}$ is then the computation of the upper triangular eigenmatrix $\mathbf{U}$ of $\tilde{\mathbf{R}}$:

$$\tilde{\mathbf{R}} = \mathbf{U} \cdot \tilde{\mathbf{\Lambda}} \cdot \mathbf{U}^{\dagger}, \tag{5.2}$$

in which $\tilde{\mathbf{\Lambda}}$ is the diagonal part of $\tilde{\mathbf{R}}$. Complex conjugate eigenvalues of $\mathbf{BB}^*$ yield eigenvalues of multiplicity 2 in eqn (5.2). In the application under consideration, the indeterminacy in the eigenspace can be met by claiming that the corresponding columns of $\tilde{\mathbf{M}}$ represent rank-1 matrices, as in the previous section.

6 Conclusion

In this paper a multilinear algebraic technique for more-sources-than-sensors independent component analysis was developed. The technique was based on a tensor generalization of an unsymmetric variant of the Tagaki eigenvalue decomposition.

Acknowledgements

Lieven De Lathauwer holds a permanent research position with the French Centre National de la Recherche Scientifique (C.N.R.S.); he is also a post-doctoral researcher 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. Part of this work was supported by the Flemish Government (Research Council K.U.Leuven (GOA-Mefisto-666,IDO), FWO (G.0240.99,G.0256.97, Research Communities ICCoS and ANMMM), Federal State (IUAP P4-02, IUAP P4-24), project BIL 98 with South Africa. The scientific responsibility is assumed by the authors.

Bibliography

Comon, P. (1994). Independent component analysis, a new concept? *Signal Processing*, **36**, 287–314.

Tong, L., et al. (1991). Indeterminacy and identifiability of blind identification. *IEEE Trans. Circuits and Systems*, **38**, 499–509.

Golub, G. and Van Loan, C. (1996). *Matrix computations* (3rd edn). Johns Hopkins University Press, Baltimore, Maryland.

Nikias, C. and Mendel, J. (1993). Signal processing with higher-order spectra. *IEEE Signal Proc. Mag.*, 10–37.

Comon, P. (1998). Blind channel identification and extraction of more sources than sensors. *Proc. SPIE Conf.*, San Diego, CA, USA, July 19-24, 2–13.

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

Cardoso, J.-F. (1991). Super-symmetric decomposition of the fourth-order cumulant tensor. Blind identification of more sources than sensors. *Proc. ICASSP-91*, **5**, 3109–3112.

De Lathauwer, L., et al. (1999). ICA algorithms for 3 sources and 2 sensors. *Proc. IEEE Signal Processing Workshop on Higher-Order Statistics (HOS'99)*, Caesarea, Israel, June 14-16, 116–120.

De Lathauwer, L., et al. (1999). ICA techniques for more sources than sensors. *Proc. IEEE Signal Processing Workshop on Higher-Order Statistics (HOS'99)*, Caesarea, Israel, June 14-16, 121–124.

De Lathauwer, L., et al. (2000). An introduction to independent component analysis. *Journal of Chemometrics*, **14**, 123–149.

Kendall, M. and Stuart, A. (1977). *The advanced theory of statistics*, **1**. Griffin, London.

Bourin, C. and Bondon, P. (1995). Efficiency of high-order moment estimates. *Proc. IEEE Signal Processing / ATHOS Workshop on Higher-Order Statistics*, Girona, Spain, 186–190.

Horn, R. and Johnson, C. (1991). *Topics in matrix analysis.* Cambridge University Press, N.Y.

van der Veen, A.-J. and Paulraj, A. (1996). An analytical constant modulus algorithm. *IEEE Trans. Signal Processing*, **44**, 1136–1155.