Independent Process Analysis without A Priori Dimensional Information*

Barnabás Póczos, Zoltán Szabó, Melinda Kiszlinger, and András Lőrincz**

Department of Information Systems
Eötvös Loránd University, Budapest, Hungary
{pbarn,szzoli}@cs.elte.hu, {kmelinda,andras.lorincz}@elte.hu

Abstract. Recently, several algorithms have been proposed for independent subspace analysis where hidden variables are i.i.d. processes. We show that these methods can be extended to certain AR, MA, ARMA and ARIMA tasks. Central to our paper is that we introduce a cascade of algorithms, which aims to solve these tasks without previous knowledge about the number and the dimensions of the hidden processes. Our claim is supported by numerical simulations. As an illustrative application where the dimensions of the hidden variables are unknown, we search for subspaces of facial components.

1 Introduction

Independent Subspace Analysis (ISA) [1] is a generalization of Independent Component Analysis (ICA). ISA assumes that certain sources depend on each other, but the dependent groups of sources are still independent of each other, i.e., the independent groups are multidimensional. The ISA task has been subject of extensive research [1,2,3,4,5,6,7,8]. In this case, one assumes that the hidden sources are independent and identically distributed (i.i.d.) in time. Temporal independence is, however, a gross oversimplification of real sources including acoustic or biomedical data. One may try to overcome this problem, by assuming that hidden processes are, e.g., autoregressive (AR) processes. Then we arrive to the AR Independent Process Analysis (AR-IPA) task [9,10]. Another method to weaken the i.i.d. assumption is to assume moving averaging (MA). This direction is called Blind Source Deconvolution (BSD) [11], in this case the observation is a temporal mixture of the i.i.d. components.

The AR and MA models can be generalized and one may assume ARMA sources instead of i.i.d. ones. As an additional step, these models can be extended to non-stationary integrated ARMA (ARIMA) processes, which are important, e.g., for modelling economic processes [12].

* M.E. Davies et al. (Eds.): ICA 2007, LNCS 4666, pp. 252–259, 2007.

http://dx.doi.org/10.1007/978-3-540-74494-8_32.

** Corresponding author

In this paper, we formulate the AR-, MA-, ARMA-, ARIMA-IPA generalizations of the ISA task, when (i) one allows for multidimensional hidden components and (ii) the dimensions of the hidden processes are not known. We show that in the undercomplete case, when the number of 'sensors' is larger than the number of 'sources', these tasks can be reduced to the ISA task.

2 Independent Subspace Analysis

The ISA task can be formalized as follows:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{e}(t), \text{ where } \mathbf{e}(t) = \left[\mathbf{e}^1(t); \ldots; \mathbf{e}^M(t)\right] \in \mathbb{R}^{D_e} \tag{1}$$

and $\mathbf{e}(t)$ is a vector concatenated of components $\mathbf{e}^m(t) \in \mathbb{R}^{d_e^m}$. The total dimension of the components is $D_e = \sum_{m=1}^M d_e^m$. We assume that for a given m, $\mathbf{e}^m(t)$ is i.i.d. in time t, and sources $\mathbf{e}^m$ are jointly independent, i.e., $I(\mathbf{e}^1, \ldots, \mathbf{e}^M) = 0$, where $I(.)$ denotes the mutual information (MI) of the arguments. The dimension of the observation $\mathbf{x}$ is D_x. Assume that $D_x > D_e$, and $\mathbf{A} \in \mathbb{R}^{D_x \times D_e}$ has rank D_e. Then, one may assume without any loss of generality that both the observed ($\mathbf{x}$) and the hidden ($\mathbf{e}$) signals are white. For example, one may apply Principal Component Analysis (PCA) as a preprocessing stage. Then the ambiguities of the ISA task are as follows [13]: Sources can be determined up to permutation and up to orthogonal transformations within the subspaces.

2.1 The ISA Separation Theorem

We are to uncover the independent subspaces. Our task is to find orthogonal matrix $\mathbf{W} \in \mathbb{R}^{D_e \times D_x}$ such that $\mathbf{y}(t) = \mathbf{W}\mathbf{x}(t)$, $\mathbf{y}(t) = \left[\mathbf{y}^1(t); \ldots; \mathbf{y}^M(t)\right]$, $\mathbf{y}^m = [y_1^m; \ldots; y_{d_e^m}^m] \in \mathbb{R}^{d_e^m}$, $(m = 1, \ldots, M)$ with the condition that components $\mathbf{y}^m$ are independent. Here, y_i^m denotes the i^{th} coordinate of the m^{th} estimated subspace. This task can be solved by means of a cost function that aims to minimize the mutual information between components:

$$J_1(\mathbf{W}) \doteq I(\mathbf{y}^1, \ldots, \mathbf{y}^M). \tag{2}$$

One can rewrite $J_1(\mathbf{W})$ as follows:

$$J_2(\mathbf{W}) \doteq I(y_1^1, \ldots, y_{d_e^M}^M) - \sum_{m=1}^M I(y_1^m, \ldots, y_{d_e^m}^m). \tag{3}$$

The first term of the r.h.s. is an ICA cost function; it aims to minimize mutual information for all coordinates. The other term is a kind of *anti-ICA* term; it aims to maximize mutual information within the subspaces. One may try to apply a heuristics and to optimize (3) in order: (1) Start by any 'infomax' ICA algorithm and minimize the first term of the r.h.s. in (3). (2) Apply only permutations to the coordinates such that they optimize the second term. Surprisingly, this heuristics leads to the global minimum of (2) in many cases. In other words, in many cases, ICA that minimizes the first term of the r.h.s. of (3) solves the ISA task apart from the grouping of the coordinates into subspaces. This feature was observed by Cardoso, first [1]. The extent of this feature is still an open issue. Nonetheless, we call it '*Separation Theorem*', because for elliptically symmetric sources and for some other distribution types one can prove that it is rigorously true [14]. (See also, the result concerning local minimum points [15]). Although there is no proof for general sources as of yet, a number of algorithms apply this heuristics with success [1,3,15,16,17,18].

2.2 ISA with Unknown Components

Another issue concerns the computation of the second term of (3). If the d_e^m dimensions of subspaces $\mathbf{e}^m$ are known then one might rely on multi-dimensional entropy estimations [8], but these are computationally expensive. Other methods deal with implicit or explicit pair-wise dependency estimations [16,15]. Interestingly, if the observations are indeed from an ICA generative model, then the minimization of the pair-wise dependencies is sufficient to get the solution of the ICA task according to the Darmois-Skitovich theorem [19]. This is not the case for the ISA task, however. There are ISA tasks, where the estimation of pair-wise dependencies is insufficient for recovering the hidden subspaces [8]. Nonetheless, such algorithms seem to work nicely in many practical cases.

A further complication arises if the d_e^m dimensions of subspaces $\mathbf{e}^m$ are not known. Then the dimension of the entropy estimation becomes uncertain. Methods that try to apply pair-wise dependencies were proposed to this task. One can find a block-diagonalization method in [15], whereas [16] makes use of kernel estimations of the mutual information.

Here we shall assume that the separation theorem is satisfied. We shall apply ICA preprocessing. This step will be followed by the estimation of the pair-wise mutual information of the ICA coordinates. These quantities will be considered as the weights of a weighted graph, the vertices of the graph being the ICA coordinates. We shall search for clusters of this graph. In our numerical studies, we make use of Kernel Canonical Correlation Analysis [4] for the MI estimation. A variant of the Ncut algorithm [20] is applied for clustering. As a result, the mutual information within (between) cluster(s) becomes large (small).

The problem is that this ISA method requires i.i.d. hidden sources. Below, we show how to generalize the ISA task to more realistic sources. Finally, we solve this more general problem when the dimensions of the subspaces are not known.

3 ISA Generalizations

We need the following notations: Let z stand for the time-shift operation, that is $(z\mathbf{v})(t) := \mathbf{v}(t-1)$. The N order polynomials of $D_1 \times D_2$ matrices are denoted as $\mathbb{R}[z]_N^{D_1 \times D_2} := \{\mathbf{F}[z] = \sum_{n=0}^{N} \mathbf{F}_n z^n, \mathbf{F}_n \in \mathbb{R}^{D_1 \times D_2}\}$. Let $\nabla^r[z] := (\mathbf{I} - \mathbf{I}z)^r$ denote the r^{th} order difference operator, where $\mathbf{I}$ is the identity matrix, $r \geq 0$, $r \in \mathbb{Z}$.

Now, we are to estimate unknown components $\mathbf{e}^m$ from observed signals $\mathbf{x}$. We always assume that $\mathbf{e}$ takes the form like in (1) and that $\mathbf{A} \in \mathbb{R}^{D_x \times D_s}$ is of full column rank.

1. AR-IPA: The AR generalization of the ISA task is defined by the following equations: $\mathbf{x} = \mathbf{A}\mathbf{s}$, where $\mathbf{s}$ is a multivariate AR(p) process i.e, $\mathbf{P}[z]\mathbf{s} = \mathbf{Q}\mathbf{e}$, $\mathbf{Q} \in \mathbb{R}^{D_s \times D_e}$, and $\mathbf{P}[z] := \mathbf{I}_{D_s} - \sum_{i=1}^{p} \mathbf{P}_i z^i \in \mathbb{R}[z]_p^{D_s \times D_s}$. We assume that $\mathbf{P}[z]$ is stable, that is $\det(\mathbf{P}[z] \neq 0)$, for all $z \in \mathbb{C}$, $|z| \leq 1$. For $d_e^m = 1$ this task was investigated in [9]. Case $d_e^m > 1$ is treated in [10]. The special case of $p = 0$ is the ISA task.
2. MA-IPA or Blind Subspace Deconvolution (BSSD) task: The ISA task is generalized to blind deconvolution task (moving average task, MA(q)) as follows: $\mathbf{x} = \mathbf{Q}[z]\mathbf{e}$, where $\mathbf{Q}[z] = \sum_{j=0}^{q} \mathbf{Q}_j z^j \in \mathbb{R}[z]_q^{D_x \times D_e}$.
3. ARMA-IPA task: The two tasks above can be merged into a model, where the hidden $\mathbf{s}$ is multivariate ARMA(p,q): $\mathbf{x} = \mathbf{A}\mathbf{s}$, $\mathbf{P}[z]\mathbf{s} = \mathbf{Q}[z]\mathbf{e}$. Here $\mathbf{P}[z] \in \mathbb{R}[z]_p^{D_s \times D_s}$, $\mathbf{Q}[z] \in \mathbb{R}[z]_q^{D_s \times D_e}$. We assume that $\mathbf{P}[z]$ is stable. Thus the ARMA process is stationary.
4. ARIMA-IPA task: In practice, hidden processes $\mathbf{s}$ may be non-stationary. ARMA processes can be generalized to the non-stationary case. This generalization is called integrated ARMA, or ARIMA(p,r,q). The assumption here is that the r^{th} difference of the process is an ARMA process. The corresponding IPA task is then

$$\mathbf{x} = \mathbf{A}\mathbf{s}, \text{ where } \mathbf{P}[z]\nabla^r[z]\mathbf{s} = \mathbf{Q}[z]\mathbf{e}. \tag{4}$$

4 Reduction of ARIMA-IPA to ISA

We show how to solve the above tasks by means of ISA algorithms. We treat the ARIMA task. Others are special cases of this one. In what follows, we assume that: (i) $\mathbf{P}[z]$ is stable, (ii) the mixing matrix $\mathbf{A}$ is of full column rank, and (iii) $\mathbf{Q}[z]$ has left inverse. In other words, there exists a polynomial matrix $\mathbf{W}[z] \in \mathbb{R}[z]^{D_e \times D_s}$ such that $\mathbf{W}[z]\mathbf{Q}[z] = \mathbf{I}_{D_e}$. [1]

The route of the solution is elaborated here. Let us note that differentiating the observation $\mathbf{x}$ of the ARIMA-IPA task in Eq. (4) in r^{th} order, and making use of the relation $z\mathbf{x} = \mathbf{A}(z\mathbf{s})$, the following holds:

$$\nabla^r[z]\mathbf{x} = \mathbf{A}\left(\nabla^r[z]\mathbf{s}\right), \text{ and } \mathbf{P}[z]\left(\nabla^r[z]\mathbf{s}\right) = \mathbf{Q}[z]\mathbf{e}. \tag{5}$$

That is taking $\nabla^r[z]\mathbf{x}$ as observations, one ends up with an ARMA-IPA task. Assume that $D_x > D_e$ (undercomplete case). We call this task uARMA-IPA. Now we show how to transform the uARMA-IPA task to ISA. The method is similar to that of [22] where it was applied for BSD.

[1] One can show for $D_s > D_e$ that under mild conditions $\mathbf{Q}[z]$-has an inverse with probability 1 [21]; e.g., when the matrix $[\mathbf{Q}_0, \ldots, \mathbf{Q}_q]$ is drawn from a continuous distribution.

Theorem. *If the above assumptions are fulfilled then in the uARMA-IPA task, observation process* $\mathbf{x}(t)$ *is autoregressive and its innovation* $\tilde{\mathbf{x}}(t) := \mathbf{x}(t) - E[\mathbf{x}(t)|\mathbf{x}(t-1), \mathbf{x}(t-2), \ldots] = \mathbf{A}\mathbf{Q}_0\mathbf{e}(t)$*, where* $E[\cdot|\cdot]$ *denotes the conditional expectation value. Consequently, there is a polynomial matrix* $\mathbf{W}_{\mathrm{AR}}[z] \in \mathbb{R}[z]^{D_x \times D_x}$ *such that* $\mathbf{W}_{\mathrm{AR}}[z]\mathbf{x} = \mathbf{A}\mathbf{Q}_0\mathbf{e}$.

Due to lack of space the proof is omitted here. Thus, AR fit of $\mathbf{x}(t)$ can be used for the estimation of $\mathbf{A}\mathbf{Q}_0\mathbf{e}(t)$. This innovation corresponds to the observation of an undercomplete ISA model $(D_x > D_e)$[2], which can be reduced to a complete ISA $(D_x = D_e)$ using PCA. Finally, the solution can be finished by any ISA procedure. The reduction procedure implies that hidden components $\mathbf{e}^m$ can be recovered only up to the ambiguities of the ISA task: components of (identical dimensions) can be recovered only up to permutations. Within each subspaces, unambiguity is warranted only up to orthogonal transformations.

The steps of our algorithm are summarized in Table 1.

Table 1: Pseudocode of the undercomplete ARIMA-IPA algorithm

Input of the algorithm
Observation: $\{\mathbf{x}(t)\}_{t=1,\ldots,T}$
Optimization
Differentiating: for observation $\mathbf{x}$ calculate $\mathbf{x}^* = \nabla^r[z]\mathbf{x}$
AR fit: for $\mathbf{x}^*$ estimate $\hat{\mathbf{W}}_{\mathrm{AR}}[z]$
Estimate innovation: $\tilde{\mathbf{x}} = \hat{\mathbf{W}}_{\mathrm{AR}}[z]\mathbf{x}^*$
Reduce uISA to ISA and whiten: $\tilde{\mathbf{x}}' = \hat{\mathbf{W}}_{\mathrm{PCA}}\tilde{\mathbf{x}}$
Apply ICA for $\tilde{\mathbf{x}}'$: $\mathbf{e}^* = \hat{\mathbf{W}}_{\mathrm{ICA}}\tilde{\mathbf{x}}'$
Estimate pairwise dependency e.g., as in [16] on $\mathbf{e}^*$
Cluster $\mathbf{e}^*$ **by Ncut**: the permutation matrix is $\mathcal{P}$
Estimation
$\hat{\mathbf{W}}_{\text{ARIMA-IPA}}[z] = \mathcal{P}\hat{\mathbf{W}}_{\mathrm{ICA}}\hat{\mathbf{W}}_{\mathrm{PCA}}\hat{\mathbf{W}}_{\mathrm{AR}}[z]\nabla^r[z]$
$\hat{\mathbf{e}} = \hat{\mathbf{W}}_{\text{ARIMA-IPA}}[z]\mathbf{x}$

5 Results

In this section we demonstrate the theoretical results by numerical simulations.

5.1 ARIMA Processes

We created a database for the demonstration: Hidden sources $\mathbf{e}^m$ are 4 pieces of 2D, 3 pieces of 3D, 2 pieces of 4D and 1 piece of 5D stochastic variables, i.e., $M = 10$. These stochastic variables are independent, but the coordinates of each

[2] Assumptions made for $\mathbf{Q}[z]$ and $\mathbf{A}$ in the uARMA-IPA task implies that $\mathbf{A}\mathbf{Q}_0$ is of full column rank and thus the resulting ISA task is well defined.

stochastic variable $\mathbf{e}^m$ depend on each other. They form a 30 dimensional space together ($D_e = 30$). For the sake of illustration, the 3D (2D) sources emit random samples of uniform distributions defined on different 3D geometrical forms (letters of the alphabet). The distributions are depicted in Fig. 1a (Fig. 1b). 30,000 samples were drawn from the sources and they were used to drive an ARIMA(2,1,6) process defined by (4). Matrix $\mathbf{A} \in \mathbb{R}^{60\times 60}$ was randomly generated and orthogonal. We also generated the polynomial $\mathbf{Q}[z] \in \mathbb{R}[z]_5^{60\times 30}$ and the stable polynomial $\mathbf{P}[z] \in \mathbb{R}[z]_1^{60\times 60}$ randomly. The visualization of the 60 dimensional process is hard: a typical 3D projection is shown in Fig. 1c. The task is to estimate original sources $\mathbf{e}^m$ using these non-stationary observations. r^{th}-order differencing of the observed ARIMA process gives rise to an ARMA process. Typical 3D projection of this ARMA process is shown Fig. 1d. Now, one can execute the other steps of Table 1 and these steps provide the estimations of the hidden components $\hat{\mathbf{e}}^m$. Here, we estimated the AR process and its order by the methods detailed in [23]. Estimations of the 3D (2D) components are provided in Fig. 1e (Fig. 1f). In the ideal case, the product of matrix $\mathbf{AQ}_0$ and the matrices provided by PCA and ISA, i.e., $\mathbf{G} := (\mathcal{P}\hat{\mathbf{W}}_{\text{ICA}}\hat{\mathbf{W}}_{\text{PCA}})\mathbf{AQ}_0 \in \mathbb{R}^{D_e \times D_e}$ is a block permutation matrix made of $d_e^m \times d_e^m$ blocks. This is shown in Fig. 1g.

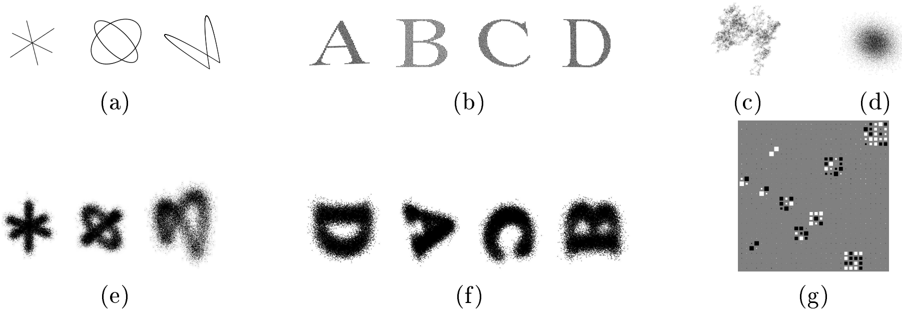


Fig. 1: (a-b) components of the database. (a): 3 pieces of 3D geometrical forms, (b): 4 pieces of 2D letters. Hidden sources are uniformly distributed variables on these objects. (c): typical 3D projection of the observation. (d): typical 3D projection of the r^{th}-order difference of the observation, (e): estimated 3D components, (f): estimated 2D components, (g): Hinton diagram of $\mathbf{G}$, which – in case of perfect estimation – becomes a block permutation matrix.

5.2 Facial Components

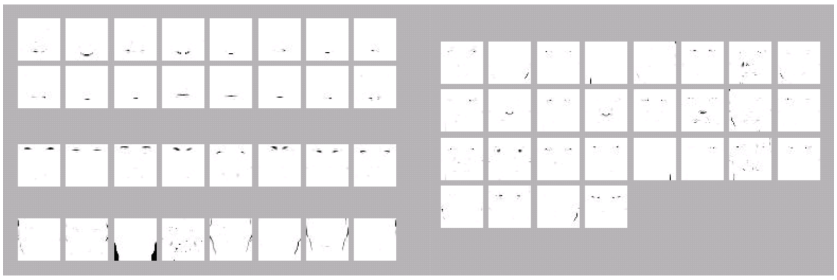
We were interested in the components that our algorithm finds when independence *is a crude approximation* at best. We have generated another database using the FaceGen[3] animation software. In our database we had 800 different front view faces with the 6 basic facial expressions. We had thus 4,800 images

[3] http://www.facegen.com/modeller.htm

in total. All images were sized to 40×40 pixel. Figure 2a shows samples of the database. A large $\mathbf{X} \in \mathbb{R}^{4800 \times 1600}$ matrix was compiled; rows of this matrix were 1600 dimensional vectors formed by the pixel values of the individual images. The *columns* of this matrix were considered as mixed signals. This treatment replicates the experiments in [24]: Bartlett et al., have shown that in such cases, undercomplete ICA finds components resembling to what humans consider facial components. We were interested in seeing the components grouped by undercomplete ISA algorithm. The observed 4800 dimensional signals were compressed by PCA to 60 dimensions and we searched for 4 pieces of ISA subspaces using the algorithm detailed in Table 1.

The 4 subspaces that our algorithm found are shown in Fig. 2b. As it can be seen, the 4 subspaces embrace facial components which correspond mostly to mouth, eye brushes, facial profiles, and eyes, respectively. Thus, ICA finds interesting components and MI based ISA groups them sensibly. The generalization up to ARIMA-IPA processes is straightforward.

(a)



(b)

Fig. 2: (a) Samples from the database. (b) Four subspaces of the components. Distinct groups correspond mostly to mouth, eye brushes, facial profiles, and eyes, respectively.

6 Conclusions

We have extended the ISA task in two ways. (1) We solved problems where the hidden components are AR, MA, ARMA, or ARIMA processes. (2) We suggested partitioning of the graph defined by pairwise mutual information to identify the hidden ISA subspaces under certain conditions. The algorithm does not require previous knowledge about the dimensions of the subspaces. An artificially generated ARIMA process was used for demonstration. The algorithm provided sensible grouping of the estimated components for facial expressions.

Acknowledgment

This material is based upon work supported by the EC FET and NEST grants, 'New Ties' (No. 003752) and 'PERCEPT' (No. 043261). Any opinions, findings and conclusions expressed in this material are those of the authors and do not necessarily reflect the views of the EC, or other members of the EC projects.

References

1. J. F. Cardoso. Multidimensional independent component analysis. In *Proc. of ICASSP*, volume 4, pages 1941–1944, 1998.
2. R. Vollgraf and K. Obermayer. Multi-dimensional ICA to separate correlated sources. In *Proc. of NIPS*, volume 14, pages 993–1000. MIT Press, 2001.
3. H. Stögbauer, A. Kraskov, S. A. Astakhov, and P. Grassberger. Least dependent component analysis based on mutual information. *Phys. Rev. E*, 70, 2004.
4. F. R. Bach and M. I. Jordan. Beyond independent components: Trees and clusters. *Journal of Machine Learning Research*, 4:1205–1233, 2003.
5. F. J. Theis. Blind signal separation into groups of dependent signals using joint block diagonalization. In *Proc. of ISCAS.*, pages 5878–5881, 2005.
6. A. Hyvärinen and U. Köster. FastISA: A fast fixed-point algorithm for independent subspace analysis. In *Proc. of ESANN*, Evere, Belgium, 2006.
7. G. Nolte, F. C. Meinecke, A. Ziehe, and K. R. Müller. Identifying interactions in mixed and noisy complex systems. *Physical Review E*, 73, 2006. doi: 051913.
8. B. Póczos and A. Lőrincz. Independent subspace analysis using geodesic spanning trees. In *Proc. of ICML*, pages 673–680, New York, NY, USA, 2005. ACM Press.
9. A. Hyvärinen. Independent component analysis for time-dependent stochastic processes. In *Proc. of ICANN*, pages 541–546, Berlin, 1998. Springer-Verlag.
10. B. Póczos, B. Takács, and A. Lőrincz. Independent subspace analysis on innovations. In *Proc. of ECML*, pages 698–706, Berlin, 2005. Springer-Verlag.
11. S. Choi, A. Cichocki, H.-M. Park, and S.-Y. Lee. Blind source separation and independent component analysis. *Neural Inf. Proc. - Lett. Rev.*, 6:1–57, 2005.
12. T. C. Mills. *Time Series Techniques for Economists.* Cambridge University Press, Cambridge, 1990.
13. F. J. Theis. Uniqueness of complex and multidimensional independent component analysis. *Signal Processing*, 84(5):951–956, 2004.
14. Z. Szabó, B. Póczos, and A. Lőrincz. Separation theorem for $\mathbb{K}$-independent subspace analysis with sufficient conditions. Technical report, 2006. http://arxiv.org/abs/math.ST/0608100.
15. F. J. Theis. Towards a general independent subspace analysis. In *Proc. of NIPS*, volume 19, 2007.
16. F. R. Bach and M. I. Jordan. Finding clusters in Independent Component Analysis. In *Proc. of ICA2003*, pages 891–896, 2003.
17. Z. Szabó, B. Póczos, and A. Lőrincz. Cross-entropy optimization for independent process analysis. In *Proc. of ICA2006, LNCS 3889*, pages 909–916. Springer, 2006.

18. K. Abed-Meraim and A. Belouchrani. Algorithms for joint block diagonalization. In *Proc. of EUSIPCO*, pages 209–212, 2004.
19. P. Comon. Independent Component Analysis, a new concept? *Signal Processing, Elsevier*, 36(3):287–314, April 1994. Special issue on Higher-Order Statistics.
20. S. Yu and J. Shi. Multiclass spectral clustering. In *Proc. of ICCV*, 2003.
21. R. Rajagopal and L. C. Potter. Multivariate MIMO FIR inverses. *IEEE Transactions on Image Processing*, 12:458 – 465, 2003.
22. A. Gorokhov and P. Loubaton. Blind identification of MIMO-FIR systems: A generalized linear prediction approach. *Signal Processing*, 73:105–124, 1999.
23. A. Neumaier and T. Schneider. Estimation of parameters and eigenmodes of multivariate autoregressive models. *ACM Trans. on Math. Soft.*, 27(1):27–57, 2001.
24. M. Bartlett, J. Movellan, and T. Sejnowski. Face recognition by independent component analysis. *IEEE Trans. on Neural Networks*, 13(6):1450–1464, 2002.