Acta Cybernetica **19** (2009) 177–190.

Complex Independent Process Analysis*

Zoltán Szabó† and András Lőrincz‡

Abstract

We present a general framework for the search of hidden independent processes in the complex domain. The task is to estimate the hidden independent multidimensional complex-valued components observing only the mixture of the processes driven by them. In our model (i) the hidden independent processes can be multidimensional, they may be subject to (ii) moving averaging, or may evolve in an autoregressive manner, or (iii) they can be non-stationary. These assumptions are covered by integrated autoregressive moving average processes and thus our task is to solve their complex extensions. We show how to reduce the undercomplete version of complex integrated autoregressive moving average processes to real independent subspace analysis that we can solve. Simulations illustrate the working of the algorithm.

1 Introduction

Our task is to find multidimensional independent components for complex variables. This task joins complex-valued neural networks [6] with independent component analysis (ICA) [4]. Although (i) complex-valued neural networks have several successful applications and (ii) there is a natural tendency to apply complex-valued computations for the analysis of biomedical signals (see, e.g., [2] and [3] for the analysis of EEG and fMRI data, respectively) the methodology of searching complex-valued independent components is barely treated in the literature. There are existing methods for the simplest ICA and blind source deconvolution tasks, but — to the best of our knowledge — there has been no study on non-i.i.d multidimensional hidden variables for the complex case. We provide the tools for this important problem family.

The paper is structured as follows: We treat the simplest complex-valued independent subspace analysis (complex ISA) problem and its solution in Section 2. In

*This research has been supported by the EC NEST 'PERCEPT' Grant FP6-043261. Opinions and errors in this manuscript are the author's responsibility, they do not necessarily reflect those of the EC or other project members.

†Department of Information Systems, Eötvös Loránd University, Pázmány Péter sétány 1/C, H-1117 Budapest, Hungary; E-mail: szzoli@cs.elte.hu

‡Corresponding author; Department of Information Systems, Eötvös Loránd University, Pázmány Péter sétány 1/C, H-1117 Budapest, Hungary; E-mail: andras.lorincz@elte.hu

DOI: 10.14232/actacyb.19.1.2009.12

the next section, the more general task, the complex-valued integrated autoregressive moving average independent subspace task is formulated. Section 4 contains our numerical illustrations. Conclusions are drawn in Section 5. The Appendix elaborates on the reduction technique: we show the series of transcriptions that reduce this task family to real independent subspace analysis.

2 The ISA Model

Below, in Section 2.1 we introduce the independent subspace analysis (ISA) problem. We show, how to reduce the complex-valued case to the real-valued one in Section 2.2.

2.1 The ISA Equations

We provide a joined formalism below for both the real and the complex-valued ISA models. To do so, let $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ may stand for either real or for complex numbers and $\mathbb{K}^{D_1 \times D_2}$ denote the set of $D_1 \times D_2$ matrices over $\mathbb{K}$. The definition of the ISA task is as follows. We assume M pieces of hidden independent multidimensional random variables (*components*).[1] Only the linear mixture of these variables is available for observation. Formally,

$$\mathbf{x}(t) = \mathbf{A}\mathbf{e}(t), \tag{1}$$

where $\mathbf{e}(t) = \left[\mathbf{e}^1(t); \ldots; \mathbf{e}^M(t)\right] \in \mathbb{K}^{D_e}$ ($D_e = Md$) is a vector concatenated of the independent components $\mathbf{e}^m(t) \in \mathbb{R}^d$, where – for the sake of notational simplicity – we used identical dimension for each components. The dimensions of observation $\mathbf{x}$ and hidden source $\mathbf{e}$ are D_x and D_e, respectively. $\mathbf{A} \in \mathbb{K}^{D_x \times D_e}$ is the so-called *mixing matrix*. The goal of the ISA task is to estimate the original source $\mathbf{e}(t)$ from observations $\mathbf{x}(t)$. Our ISA assumptions are the followings:

1. For a given m, $\mathbf{e}^m(t)$ is i.i.d. in time t.
2. $I(\mathbf{e}^1, \ldots, \mathbf{e}^M) = 0$, where I stands for the mutual information of the arguments.
3. $\mathbf{A} \in \mathbb{K}^{D_x \times D_s}$ has full column rank, so it has a left inverse.

If $\mathbb{K} = \mathbb{C}$, then one can talk about complex-valued ISA (complex ISA). For the case of $\mathbb{K} = \mathbb{R}$, the ISA task is real-valued. The particular case of $d = 1$ gives rise to the ICA task.

[1] An excellent review can be found in [5] on complex random variables. Throughout this paper all complex variables are assumed to be full, i.e., they are not concentrated in any lower dimensional complex subspace.

2.2 Reduction of Complex-valued ISA to Real-valued ISA

In what follows, the complex ISA task is reduced to a real-valued one. To do so, consider the mappings

$$\varphi_v : \mathbb{C}^L \ni \mathbf{v} \mapsto \mathbf{v} \otimes \begin{bmatrix} \Re(\cdot) \\ \Im(\cdot) \end{bmatrix} \in \mathbb{R}^{2L}, \tag{2}$$

$$\varphi_M : \mathbb{C}^{L_1 \times L_2} \ni \mathbf{M} \mapsto \mathbf{M} \otimes \begin{bmatrix} \Re(\cdot) & -\Im(\cdot) \\ \Im(\cdot) & \Re(\cdot) \end{bmatrix} \in \mathbb{R}^{2L_1 \times 2L_2}, \tag{3}$$

where $\otimes$ is the Kronecker product, $\Re$ stands for the real part, $\Im$ for the imaginary part, subscript 'v' ('M') for vector (matrix). Known properties of mappings φ_v, φ_M are as follows [8]:

$$\det[\varphi_M(\mathbf{M})] = |\det(\mathbf{M})|^2 \qquad (\mathbf{M} \in \mathbb{C}^{L \times L}), \tag{4}$$

$$\varphi_M(\mathbf{M}_1\mathbf{M}_2) = \varphi_M(\mathbf{M}_1)\varphi_M(\mathbf{M}_2) \qquad (\mathbf{M}_1 \in \mathbb{C}^{L_1 \times L_2}, \mathbf{M}_2 \in \mathbb{C}^{L_2 \times L_3}), \tag{5}$$

$$\varphi_v(\mathbf{M}\mathbf{v}) = \varphi_M(\mathbf{M})\varphi_v(\mathbf{v}) \qquad (\mathbf{M} \in \mathbb{C}^{L_1 \times L_2}, \mathbf{v} \in \mathbb{C}^{L_2}), \tag{6}$$

$$\varphi_M(\mathbf{M}_1 + \mathbf{M}_2) = \varphi_M(\mathbf{M}_1) + \varphi_M(\mathbf{M}_2) \qquad (\mathbf{M}_1, \mathbf{M}_2 \in \mathbb{C}^{L_1 \times L_2}), \tag{7}$$

$$\varphi_M(c\mathbf{M}) = c\varphi_M(\mathbf{M}) \qquad (\mathbf{M} \in \mathbb{C}^{L_1 \times L_2}, c \in \mathbb{R}). \tag{8}$$

In words: (4) describes transformation of determinant, while (5), (6), (7) and (8) expresses preservation of operation for matrix-matrix multiplication, matrix-vector multiplication, matrix addition, real scalar-matrix multiplication, respectively.

Now, one may apply φ_v to the complex ISA equation ((1) with $\mathbb{K} = \mathbb{C}$) and use (6). The result is as follows:

$$\varphi_v(\mathbf{x}) = \varphi_M(\mathbf{A})\varphi_v(\mathbf{e}). \tag{9}$$

Given that (i) the independence of $\mathbf{e}^m \in \mathbb{C}^d$ is equivalent to that of $\varphi_v(\mathbf{e}^m) \in \mathbb{R}^{2d}$, and (ii) the existence of the left inverse of $\varphi_M(\mathbf{A})$ is inherited from $\mathbf{A}$ (see Eq. (5)), we end up with a real-valued ISA task with observation $\varphi_v(\mathbf{x})$ and M pieces of $2d$-dimensional hidden components $\varphi_v(\mathbf{e}^m)$.

3 Complex-valued Integrated Autoregressive Moving Average Independent Subspace Analysis

The solution of the complex-valued ISA task is important, because a series of transcriptions enables one to reduce much more general processes to it. We elaborate on this transcription series in the Appendix. Here, we provide the end result of this series, the model for complex-valued integrated autoregressive moving average independent subspace analysis. This will be the subject of our illustrations in the next section.

The complex-valued autoregressive moving average independent subspace task is this:

$$\mathbf{s}(t) = \sum_{i=1}^{p} \mathbf{P}_i \mathbf{s}(t-i) + \sum_{j=0}^{q} \mathbf{Q}_j \mathbf{e}(t-j), \tag{10}$$

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t). \tag{11}$$

These equations can be written in a more compact form by introducing the notations $\mathbf{P}[z] := \mathbf{I} - \sum_{i=1}^{p} \mathbf{P}_i z^i \in \mathbb{K}[z]^{D_s \times D_s}$ and $\mathbf{Q}[z] := \sum_{j=0}^{q} \mathbf{Q}_j z^j \in \mathbb{K}[z]^{D_s \times D_e}$. Here, polynomial matrices $\mathbf{P}[z]$ and $\mathbf{Q}[z]$ represent the autoregressive (AR) and the moving average (MA) parts, respectively. Now, we can simply write the complex-valued autoregressive moving average independent subspace task as this:

$$\mathbf{P}[z]\mathbf{s} = \mathbf{Q}[z]\mathbf{e}, \tag{12}$$

$$\mathbf{x} = \mathbf{A}\mathbf{s}. \tag{13}$$

Now, we provide the definition of the complex-valued *integrated* autoregressive moving average independent subspace task. This means that the difference process is complex-valued autoregressive moving average process. For the sake of notational transparency, let $\nabla^r[z] := (\mathbf{I} - \mathbf{I}z)^r$ denote the operator of the r^{th} order difference ($0 \leq r \in \mathbb{Z}$), where $\mathbf{I}$ is the identity matrix. Then, the general integrated task as is follows. We assume M pieces of hidden independent random variables (*components*). Only the linear mixture of $ARIMA(p,r,q)$ ($0 \leq p, r \in \mathbb{Z};\ -1 \leq q \in \mathbb{Z}$) processes driven by these hidden components is available for observation. Formally,

$$\mathbf{P}[z]\nabla^r[z]\mathbf{s} = \mathbf{Q}[z]\mathbf{e}, \tag{14}$$

$$\mathbf{x} = \mathbf{A}\mathbf{s}, \tag{15}$$

where $\mathbf{e}(t) = \left[\mathbf{e}^1(t); \ldots; \mathbf{e}^M(t)\right] \in \mathbb{K}^{D_e}$ ($D_e = Md$) is a vector concatenated of the independent components $\mathbf{e}^m(t) \in \mathbb{R}^d$. Observation $\mathbf{x} \in \mathbb{K}^{D_x}$, hidden source $\mathbf{s} \in \mathbb{K}^{D_s}$, mixing matrix $\mathbf{A} \in \mathbb{K}^{D_x \times D_s}$, polynomial matrices $\mathbf{P}[z] := \mathbf{I} - \sum_{i=1}^{p} \mathbf{P}_i z^i \in \mathbb{K}[z]^{D_s \times D_s}$ and $\mathbf{Q}[z] := \sum_{j=0}^{q} \mathbf{Q}_j z^j \in \mathbb{K}[z]^{D_s \times D_e}$. The task is to estimate the original source $\mathbf{e}(t)$ from observations $\mathbf{x}(t)$.

The conditions, when we can reduce the complex-valued *integrated* autoregressive moving average independent subspace task to an ISA task are as follows:

1. For a given m, $\mathbf{e}^m(t)$ is i.i.d. in time t.
2. $I(\mathbf{e}^1, \ldots, \mathbf{e}^M) = 0$.
3. $\mathbf{A} \in \mathbb{C}^{D_x \times D_s}$ has full column rank.
4. Polynomial matrix $\mathbf{P}[z]$ is stable ($det(\mathbf{P}[z])$ has no roots within the closed unit circle).
5. The task is undercomplete: $D_x > D_e$.

The case of $r = 0$ corresponds to the complex-valued autoregressive moving average independent subspace task. Details of the series of transcriptions can be found in the Appendix. The interested reader may find further details and a number of references about multidimensional independent component analysis in [13].

4 Illustrations

The complex-valued integrated autoregressive moving average independent subspace analysis problem can be reduced to a real ISA task as it is detailed in Appendix B. Here we illustrate the performance of the algorithm based on those reductions. To evaluate the solutions we use a performance measure given in Section 4.1. Our test database is described in Section 4.2. Numerical results are summarized in Section 4.3.

4.1 Performance Index

Using the reduction principle of Section B, in the ideal case, the product of matrix $\varphi_M(\mathbf{A})\varphi_M(\mathbf{Q}_0)$ and the matrices provided by PCA (principal component analysis), ISA, i.e., $\mathbf{G} := (\hat{\mathbf{W}}_{\mathrm{ISA}}\hat{\mathbf{W}}_{\mathrm{PCA}})\varphi_M(\mathbf{A})\varphi_M(\mathbf{Q}_0) \in \mathbb{R}^{2D_e \times 2D_e}$ is a block-permutation matrix made of $2d \times 2d$ blocks. This block-permutation structure can be measured by the normalized version of the Amari-error [1] adapted to the ISA task [19]. Let us decompose matrix $\mathbf{G} \in \mathbb{R}^{2D_e \times 2D_e}$ into blocks of size $2d \times 2d$: $\mathbf{G} = [\mathbf{G}_{ij}]_{i,j=1,\ldots,M}$. Let g_{ij} denote the sum of the absolute values of matrix $\mathbf{G}_{ij} \in \mathbb{R}^{2d \times 2d}$. Now, the following term [15]

$$r(\mathbf{G}) := \frac{1}{2M(M-1)} \left[\sum_{i=1}^{M} \left(\frac{\sum_{j=1}^{M} g_{ij}}{\max_j g_{ij}} - 1 \right) + \sum_{j=1}^{M} \left(\frac{\sum_{i=1}^{M} g_{ij}}{\max_i g_{ij}} - 1 \right) \right] \tag{16}$$

denotes the Amari-index that takes values in [0,1]: for an ideal block-permutation matrix $\mathbf{G}$ it takes 0; for the worst case it takes 1.

4.2 Test Database

We created a database for the illustration, which is scalable in dimension d. The hidden sources $\mathbf{e}^m$ were defined by geometrical forms in $\mathbb{C}^d$. Using that $\varphi_v : \mathbb{C}^d \to \mathbb{R}^{2d}$ is a bijection, variables $\mathbf{e}^m$ were created in $\mathbb{R}^{2d}$. Namely, we used geometrical forms in $\mathbb{R}^{2d}$, applied uniform sampling on these and the φ_v^{-1} derived image of the samples $\mathbb{R}^{2d}$ was taken as $\mathbf{e}^m \in \mathbb{C}^d$. Geometrical forms were chosen as follows. We used: (i) the surface of the unit ball, (ii) the straight lines that connect the opposing corners of the unit cube, (iii) the broken line between $2d + 1$ points $\mathbf{0} \to \mathbf{e}_1 \to \mathbf{e}_1 + \mathbf{e}_2 \to \ldots \to \mathbf{e}_1 + \ldots + \mathbf{e}_{2d}$ (where $\mathbf{e}_i$ is the i canonical basis vector in $\mathbb{R}^{2d}$, i.e., all of its coordinates are zero except the i, which is 1), and (iv) the skeleton of the unit square. Thus in our numerical studies the number of components M was equal to 4. For illustration, see Fig 1.

(a) (b) (c) (d)

Figure 1: Illustration of our test database. Hidden components $\mathbf{e}^m \in \mathbb{C}^d$ are defined as variables uniformly distributed on geometrical forms, shown here. For this, bijection $\varphi_v : \mathbb{C}^d \to \mathbb{R}^{2d}$ was used. The figure serves illustrative purposes only, because $2d$ is even.

4.3 Simulations

We present our simulation results here. We focus on 2 distinct issues:

1. How does the estimation error scale with the number of samples? Sample number T ranged between $2,000 \leq T \leq 30,000$ and the orders of the AR and MA processes were kept low: $p = 1$, $q = 1$ (precisely, MA order: $q + 1 = 2$).

2. We assumed that polynomial matrix $\mathbf{P}[z]$ of Eq. (14) is stable. In the case of $r = 0$ this means that process $\mathbf{s}$ is stationary. For $r > 1$ the model describes non-stationary processes. It is expected that if the roots of $\mathbf{P}[z]$ are close to the unit circle then our estimation will deteriorate. We investigated this by generating polynomial matrix $\mathbf{P}[z]$ as follows:

$$\mathbf{P}[z] = \prod_{i=1}^{p} (\mathbf{I} - \lambda \mathbf{U}_i z) \quad (|\lambda| < 1, \lambda \in \mathbb{R}) \tag{17}$$

   Matrices $\mathbf{U}_i \in \mathbb{C}^{D_s \times D_s}$ were random unitary matrices and the $\lambda \to 1$ limit was studied. Now, sample number was set to $T = 20,000$. For the 'small task' ($p = 1$, $q = 1$) we could not see relevant performance drops even for $\lambda = 0.99$, therefore we increased parameters p and q to 5 and 10, respectively.

In our simulations: (i) the measure of undercompleteness was 2 ($D_x = D_s = 2D_e$), (ii) the Amari-index was used to measure the precision of our method. For all values of parameters (T, p, r, q), the average performances upon 20 random initializations of $\mathbf{e}, \mathbf{Q}[z], \mathbf{P}[z]$ and $\mathbf{A}$ were taken. In economic computations, the value of r is typically ≤ 2, we investigated values between $1 \leq r \leq 3$. The coordinates of matrices $\mathbf{Q}_j$ in the MA part (see Eq. (14)) were chosen independently and uniformly from the $\{\mathbf{v} = v_1 + iv_2 \in \mathbb{C} : -\frac{1}{2} \leq v_1, v_2 \leq \frac{1}{2}\}$ complex unit square. The mixing matrix $\mathbf{A}$ (see, Eq. (15)) was drawn randomly from the unitary group. Polynomial matrix $\mathbf{P}[z]$ was generated according to Eq. (17). The choice of λ is detailed later. The order of the AR estimation (see Fig. 4) was constrained from above as follows $deg(\hat{\mathbf{W}}_{\mathrm{AR}}[z]) \leq 2(q + 1) + p$ (i.e., two times the MA length + the AR length). We used the technique of [9] with the Schwarz's Bayesian Criterion to determine

the optimal order of the AR process. We applied the method of [14] to solve the associated ISA task.

In our first test the order of AR (p) and the order of MA processes (q) were set to the minimal meaningful values, 1. We investigated the estimation error as a function of the sample number. Parameter r of the process was set to $r = 1, 2$ and 3 in the different computations. Sample number varied as $T = 2, 5, 10, 20, 30 \cdot 10^3$. Scaling properties of the algorithm were studied by changing the value of the dimension of the components d between 1 and 15. The value of λ was 0.9 (see, Eq. (17)). Our results are summarized in Fig. 2(e), with an illustrative example given in Fig. 2(a)-(d).[2] According to Fig. 2(e), our method could recover the hidden components with high precision. The Amari-index $r(T)$ follows power law $r(T) \propto T^{-c}$ ($c > 0$). The power law is manifested by straight lines on loglog scales. The slope of the lines are about the same for different d values. The actual values of the Amari-index can be found in Table 1 for sample number $T = 30,000$.

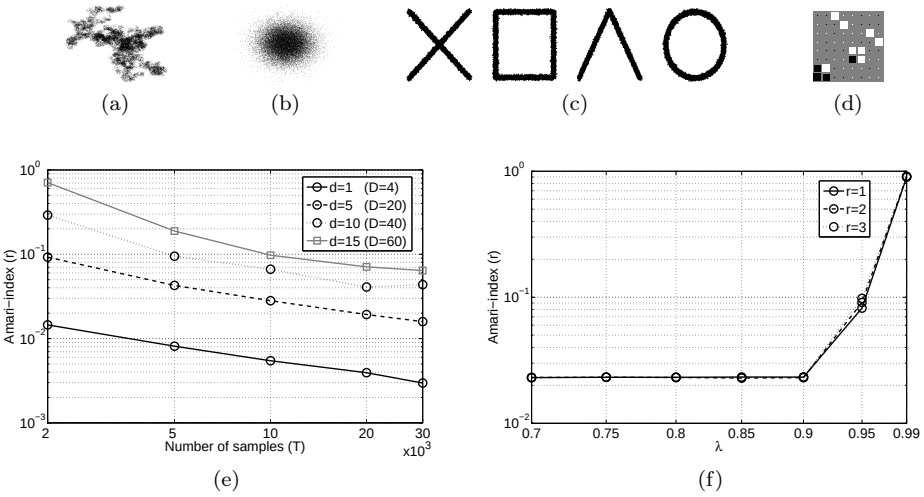

Figure 2: Illustration of our method. (a)-(d): AR order $p = 1$, MA order $q = 1$, order of integration $r = 1$, sample number $T = 30,000$. (a)-(b): typical 2D projection of the observed mixed $\mathbf{x}$ signal, and its r^{th}-order difference. (c): estimated components $[\varphi_v(\mathbf{e}^m)]$. (d): Hinton-diagram of $\mathbf{G}$, ideally block-permutation matrix with 2×2 blocks. (e): average Amari-index as a function of the sample size on loglog scale for different dimensional (d) components; $\lambda = 0.9$, $p = 1$, $q = 1$, $r = 1$ ($r \leq 3$). For $T = 30,000$, the exact errors are shown in Table 1. (f): Estimation error on log scale as a function of the magnitude of the roots of polynomial matrix $\mathbf{P}[z]$. (If $\lambda = 1$ then the roots are on the unit circle.) Parameters: $r = 1, 2$ and 3; AR order: $p = 5$; MA order: $q = 10$.

[2]The $r = 1$ case is illustrated, results are similar in the studied $r \leq 3$ domain.

In our other test we investigated what happens if the roots of polynomial matrix $\mathbf{P}[z]$ move towards the unit circle from the outside. In these simulations, parameter λ of Eq. (17) was varied. Our question was the following: How does our method behave when λ is close to 1? The sample number was set to $T = 20,000$ and simultaneously the AR order p, and MA order q were increased to 5 and 10, respectively. Dimension d of components $\mathbf{e}^m$ was 5. Parameter r took values on $1, 2$ and 3. Results are shown in Fig. 2(f). According to this figure, there is a sudden change in the performance at around $\lambda = 0.9 - 0.95$. Estimations for parameters $r = 1, 2$ and 3 have about the same errors. We note that for $p = 1$ and $q = 1$ we did not experience any degradation of performance up to $\lambda = 0.99$.

$d = 1$	$d = 5$	$d = 10$	$d = 15$
0.29% ($\pm$0.05)	1.59% ($\pm$0.05)	4.36% ($\pm$2.61)	6.40% ($\pm$3.10)

Table 1: Amari-index as a function of the dimension of the components d: average $\pm$ std. Sample size: $T = 30,000$. For other sample numbers, see Fig. 2(e).

5 Conclusions

We have given a general framework for the search of hidden independent components in the complex domain. This integrated autoregressive moving average subspace problem formulation can cover several distinct assumptions. The hidden processes (i) may be multidimensional, (ii) can be autoregressive or moving average processes, and (iii) may be non-stationary processes, too. We have shown that the undercomplete version of this task can be reduced to real-valued ISA problem. We investigated the efficiency of our method by means of numerical simulations. We experienced that (i) the estimation error decreases and follows a power law as a function of the sample number and (ii) the estimation is robust if the AR term is stable.

References

[1] Amari, Shun-ichi, Cichocki, Andrzej, and Yang, Howard H. A new learning algorithm for blind signal separation. *Advances in Neural Information Processing Systems*, 8:757–763, 1996.

[2] Anemüller, Jörn, Sejnowski, Terrence J., and Makeig, Scott. Complex ICA of frequency-domain electroencephalographic data. *Neural Networks*, 16:1311–1323, 2003.

[3] Calhoun, Vince D. and Adali, Tülay. Complex infomax: Convergence and approximation of infomax with complex nonlinearities. *VLSI Signal Processing Systems for Signal, Image, and Video Technology*, 44(1/2):173–190, 2006.

[4] Cichocki, Andrzej and Amari, Shun-ichi. *Adaptive blind signal and image processing*. John Wiley & Sons, 2002.

[5] Eriksson, Jan. Complex random vectors and ICA models: Identifiability, uniqueness and separability. *IEEE Transactions on Information Theory*, 52(3), 2006.

[6] Hirose, Akira. *Complex-Valued Neural Networks: Theories and Applications*, volume 5 of *Series on Innovative Intelligence*. World Scientific Publishing Co. Pte. Ltd., 2004.

[7] Hyvärinen, Aapo. Independent component analysis for time-dependent stochastic processes. In *Proceedings of ICANN*, pages 541–546, Berlin, 1998. Springer-Verlag.

[8] Krishnaiah, P. and Lin, Jugan. Complex elliptically symmetric distributions. *Communications in Statistics*, 15(12):3693–3718, 1986.

[9] Neumaier, Arnold and Schneider, Tapio. Estimation of parameters and eigenmodes of multivariate autoregressive models. *ACM Transactions on Mathematical Software*, 27(1):27–57, 2001.

[10] Póczos, Barnabás, Szabó, Zoltán, Kiszlinger, Melinda, and Lőrincz, András. Independent process analysis without a priori dimensional information. In *Proceedings of ICA*, volume 4666 of *LNCS*, pages 252–259. Springer-Verlag, 2007.

[11] Póczos, Barnabás, Takács, Bálint, and Lőrincz, András. Independent subspace analysis on innovations. In *Proceedings of ECML*, pages 698–706. Springer-Verlag, 2005.

[12] Rajagopal, Ravikiran and Potter, Lee C. Multivariate MIMO FIR inverses. *IEEE Transactions on Image Processing*, 12:458 – 465, 2003.

[13] Szabó, Zoltán. *Separation Principles in Independent Process Analysis*. PhD thesis, Eötvös Loránd University, Budapest, 2008. (submitted; available at `http://nipg.inf.elte.hu/index.php?option=com_remository&Itemid=27&func=fileinfo&id=153`).

[14] Szabó, Zoltán and Lőrincz, András. Real and complex independent subspace analysis by generalized variance. In *Proceedings of ICARN*, pages 85–88, Liverpool, U.K., 2006. (Available at `http://arxiv.org/abs/math.ST/0610438`).

[15] Szabó, Zoltán, Póczos, Barnabás, and Lőrincz, András. Cross-entropy optimization for independent process analysis. In *Proceedings of ICA*, volume 3889 of *LNCS*, pages 909–916. Springer, 2006.

[16] Szabó, Zoltán, Póczos, Barnabás, and Lőrincz, András. Undercomplete blind subspace deconvolution. *Journal of Machine Learning Research*, 8:1063–1095, 2007.

[17] Szabó, Zoltán, Póczos, Barnabás, and Lőrincz, András. Undercomplete blind subspace deconvolution via linear prediction. In *Proceedings of ECML*, volume 4701 of *LNAI*, pages 740–747. Springer-Verlag, 2007.

[18] Theis, Fabian J. Uniqueness of complex and multidimensional independent component analysis. *Signal Processing*, 84(5):951–956, 2004.

[19] Theis, Fabian J. Multidimensional independent component analysis using characteristic functions. In *Proceedings of EUSIPCO*, Antalya, Turkey, 2005.

Appendix

In this Appendix we elaborate on the details of the general $\mathbb{K}$-ARIMA-IPA model. Section A: we describe special cases, going step-by-step to more general process models. In Section B we reduce the complex-valued ARIMA-IPA task to the real-valued case. This reduction is analogous to the main lines of Section 2.2.

A The $\mathbb{K}$-ARIMA-IPA Equations

We defined the ISA task in Section 2.1. In case of ISA, one assumes that the hidden sources are independent and identically distributed (i.i.d.) in time. Temporal independence is, however, a gross oversimplification of sources. Temporal dependencies can be diminished, e.g., by an

- autoregressive (AR) assumption for the hidden sources. This is the AR independent process analysis (AR-IPA) task [7, 11]:

$$\mathbf{s}(t) = \sum_{i=1}^{p} \mathbf{P}_i \mathbf{s}(t-i) + \mathbf{Q}_0 \mathbf{e}(t), \tag{18}$$

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t). \tag{19}$$

  Here, we assume the i.i.d. property for driving noise $\mathbf{e}(t)$, but not for hidden source $\mathbf{s}(t)$. The state equation ((18)) and the observation ((19)) can be written compactly using the polynomial matrix formalism: let z stand for the time-shift operation, that is $(z\mathbf{v})(t) := \mathbf{v}(t-1)$ and polynomials of $D_1 \times D_2$ matrices are denoted as $\mathbb{K}[z]^{D_1 \times D_2} := \{\mathbf{F}[z] = \sum_{n=0}^{N} \mathbf{F}_n z^n, \mathbf{F}_n \in \mathbb{K}^{D_1 \times D_2}\}$. Then, Eqs. (18)-(19) take the forms:

$$\mathbf{P}[z]\mathbf{s} = \mathbf{Q}_0 \mathbf{e}, \tag{20}$$

$$\mathbf{x} = \mathbf{A}\mathbf{s}, \tag{21}$$

  where $\mathbf{P}[z] := \mathbf{I} - \sum_{i=1}^{p} \mathbf{P}_i z^i \in \mathbb{K}[z]^{D_s \times D_s}$ represents the AR part of order p. For $p = 0$, the $\mathbb{K}$-valued ISA ($\mathbb{K}$-ISA) task is recovered.

- moving average (MA) assumption. The observation in the $\mathbb{K}$-MA-IPA task (which could also be called $\mathbb{K}$-blind subspace deconvolution (BSSD) [16, 17]) task) is as follows:

$$\mathbf{x}(t) = \sum_{j=0}^{q} \mathbf{Q}_j \mathbf{e}(t-j). \tag{22}$$

  In polynomial matrix form

$$\mathbf{x} = \mathbf{Q}[z]\mathbf{e}, \tag{23}$$

  where $\mathbf{Q}[z] := \sum_{j=0}^{q} \mathbf{Q}_j z^j \in \mathbb{K}[z]^{D_s \times D_e}$ represents the MA part of order q. Here:

  - for $q = 0$ the $\mathbb{K}$-ISA task appears.
  - If $d = 1$ holds, then we end up with the $\mathbb{K}$-BSD ($\mathbb{K}$-blind source deconvolution) problem.

Combining the AR and the MA assumptions the $\mathbb{K}$-ARMA-IPA task emerges:

$$\mathbf{s}(t) = \sum_{i=1}^{p} \mathbf{P}_i \mathbf{s}(t-i) + \sum_{j=0}^{q} \mathbf{Q}_j \mathbf{e}(t-j), \tag{24}$$

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t), \tag{25}$$

which can be written compactly as

$$\mathbf{P}[z]\mathbf{s} = \mathbf{Q}[z]\mathbf{e}, \tag{26}$$

$$\mathbf{x} = \mathbf{A}\mathbf{s}, \tag{27}$$

where $\mathbf{P}[z] := \mathbf{I} - \sum_{i=1}^{p} \mathbf{P}_i z^i \in \mathbb{K}[z]^{D_s \times D_s}$ and $\mathbf{Q}[z] := \sum_{j=0}^{q} \mathbf{Q}_j z^j \in \mathbb{K}[z]^{D_s \times D_e}$. For the general ARMA process the condition is that polynomial matrix $\mathbf{P}[z]$ is *stable*, that is $\det(\mathbf{P}[z]) \neq 0$, for all $z \in \mathbb{C}, |z| \leq 1$. We note that the stability of $\mathbf{P}[z]$ implies the stationarity of ARMA process $\mathbf{s}$.

Using temporal differences, we enter the domain of non-stationary processes. In such case the ARMA property is assumed for the first order difference process $\mathbf{s}(t) - \mathbf{s}(t-1)$, or similarly for higher order difference processes. For the general order r, let $\nabla^r[z] := (\mathbf{I} - \mathbf{I}z)^r$ denote the operator of the r^{th} order difference ($0 \leq r \in \mathbb{Z}$), where $\mathbf{I}$ is the identity matrix. Then, the definition of the $\mathbb{K}$-ARIMA-IPA task as is follows. We assume M pieces of hidden independent random variables (*components*). Only the linear mixture of $ARIMA(p, r, q)$ ($0 \leq p, r \in \mathbb{Z}$; $-1 \leq q \in \mathbb{Z}$) processes driven by these hidden components is available for observation. Formally,

$$\mathbf{P}[z]\nabla^r[z]\mathbf{s} = \mathbf{Q}[z]\mathbf{e}, \tag{28}$$

$$\mathbf{x} = \mathbf{A}\mathbf{s}, \tag{29}$$

where $\mathbf{e}(t) = \left[\mathbf{e}^1(t); \ldots; \mathbf{e}^M(t)\right] \in \mathbb{K}^{D_e}$ ($D_e = Md$) is a vector concatenated of the independent components $\mathbf{e}^m(t) \in \mathbb{R}^d$. Observation $\mathbf{x} \in \mathbb{K}^{D_x}$, hidden source

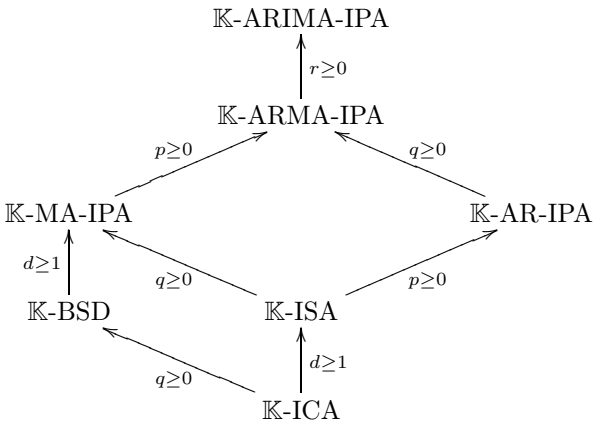

Figure 3: The $\mathbb{K}$-ARIMA-IPA model. Arrows show the direction of generalization. The labels of the arrows explain the method of the generalization. For example: '$\mathbb{K}$-ICA $\xrightarrow{d\geq 1}$$\mathbb{K}$-ISA' means that the $\mathbb{K}$-ISA task is the generalization of the $\mathbb{K}$-ICA task such that the hidden independent sources may be multidimensional, i.e., $d \geq 1$.

$\mathbf{s} \in \mathbb{K}^{D_s}$, mixing matrix $\mathbf{A} \in \mathbb{K}^{D_x \times D_s}$, polynomial matrices $\mathbf{P}[z] := \mathbf{I} - \sum_{i=1}^{p} \mathbf{P}_i z^i \in \mathbb{K}[z]^{D_s \times D_s}$ and $\mathbf{Q}[z] := \sum_{j=0}^{q} \mathbf{Q}_j z^j \in \mathbb{K}[z]^{D_s \times D_e}$. The goal of the $\mathbb{K}$-ARIMA-IPA task is to estimate the original source $\mathbf{e}(t)$ from observations $\mathbf{x}(t)$.

Our $\mathbb{K}$-ARIMA-IPA assumptions are listed below:

1. For a given m, $\mathbf{e}^m(t)$ is i.i.d. in time t.
2. $I(\mathbf{e}^1, \ldots, \mathbf{e}^M) = 0$.
3. $\mathbf{A} \in \mathbb{K}^{D_x \times D_s}$ has full column rank.
4. Polynomial matrix $\mathbf{P}[z]$ is stable.

The $\mathbb{K}$-ARMA-IPA task corresponds to the $r = 0$ case.

The relations amongst the different tasks are summarized in Fig. 3.

B Decomposition of the $\mathbb{C}$-uARIMA-IPA Model

Here, we reduce the $\mathbb{C}$-ARIMA-IPA task to $\mathbb{R}$-ISA for the *undercomplete* case ($D_x > D_e$; $\mathbb{C}$-uARIMA-IPA; letter 'u' is to show the restriction for the undercomplete case). The reduction takes two steps:

1. In Section B.1, the $\mathbb{C}$-uARIMA-IPA task is reduced to the $\mathbb{R}$-uARIMA-IPA task.

2. The $\mathbb{R}$-uARIMA-IPA task can be solved following the route suggested in [10], because it can be reduced to the $\mathbb{R}$-ISA task. The undercomplete assumption is used in the second step only. For the sake of completeness, we also provide a description of the second step (Section B.2).

In addition to the conditions of the ARIMA-IPA task, we assume that $\mathbf{Q}[z]$ 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 $\mathbf{W}[z]\mathbf{Q}[z] = \mathbf{I}_{D_e}$ (thus $D_s > D_e$)[3].

B.1 Reducing the Complex ARIMA-IPA Task to Real Variables

Here we reduce the tasks of Fig. 3, which have complex variables to real variables. In particular, we reduce the $\mathbb{C}$-uARIMA-IPA problem to the $\mathbb{R}$-uARIMA-IPA task.

One may apply φ_v to the (28)-(29) $\mathbb{C}$-ARIMA-IPA equations (with $\mathbb{K} = \mathbb{C}$) and use (6)-(8). The result is as follows:

$$\varphi_M(\mathbf{P}[z])\nabla^r[z]\varphi_v(\mathbf{s}) = \varphi_M(\mathbf{Q}[z])\varphi_v(\mathbf{e}), \tag{30}$$

$$\varphi_v(\mathbf{x}) = \varphi_M(\mathbf{A})\varphi_v(\mathbf{s}). \tag{31}$$

Given that (i) the independence of $\mathbf{e}^m \in \mathbb{C}^d$ is equivalent to that of $\varphi_v(\mathbf{e}^m) \in \mathbb{R}^{2d}$, and (ii) the stability of $\varphi_M(\mathbf{P}[z])$ and the existence of the left inverse of $\varphi_M(\mathbf{Q}[z])$ are inherited from $\mathbf{P}[z]$ and $\mathbf{Q}[z]$, respectively (see Eqs. (4) and (5)), we end up with an $\mathbb{R}$-ARIMA-IPA task with (p, r, q) parameters and M pieces of $2d$-dimensional hidden components $\varphi_v(\mathbf{e}^m)$.

B.2 Reduction of $\mathbb{R}$-uARIMA-IPA to $\mathbb{R}$-ISA

We ended up with a $\mathbb{R}$-uARIMA-IPA task in Section B.1. This task can be reduced to a $\mathbb{R}$-ISA task as it has been shown in [10]. The reduction requires two steps: (i) temporal differencing and (ii) linear prediction. These steps are formalized by the following lemmas:

Lemma 1. *Differentiating the observation* $\mathbf{x}$ *of an* $\mathbb{R}$*-(u)ARIMA-IPA task in* r^{th} *order one obtains an* $\mathbb{R}$*-(u)ARMA-IPA task:*

$$\mathbf{P}[z]\left(\nabla^r[z]\mathbf{s}\right) = \mathbf{Q}[z]\mathbf{e}, \tag{32}$$

$$\nabla^r[z]\mathbf{x} = \mathbf{A}\left(\nabla^r[z]\mathbf{s}\right), \tag{33}$$

(where, the relation $z\mathbf{x} = \mathbf{A}(z\mathbf{s})$ *has been used).*

We note that polynomial matrix $\varphi_M(\mathbf{Q}[z])$ derived from the $\mathbb{C}$-uARIMA-IPA task has a left inverse (see Section B.1). Thus, we can apply the above quoted linear prediction based result:

[3] One can show for $D_s > D_e$ that under mild conditions $\mathbf{Q}[z]$ has left inverse with probability 1 [12]; e.g., when the matrix $[\mathbf{Q}_0, \ldots, \mathbf{Q}_q]$ is drawn from a continuous distribution.

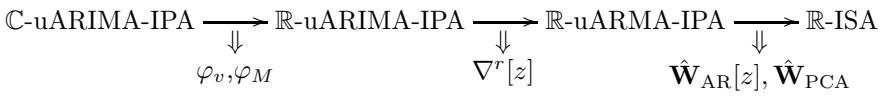

Figure 4: Reduction of ℂ-uARIMA-IPA to ℝ-ISA. Prefix 'u': undercomplete case. Double arrows: transformations of the reduction steps. Estimated ℝ-ISA separation matrix: $\hat{\mathbf{W}}_{\mathrm{ISA}}$. $\hat{\mathbf{W}}_{\mathbb{R}-\mathrm{ARIMA}}[z] = \hat{\mathbf{W}}_{\mathrm{ISA}}\hat{\mathbf{W}}_{\mathrm{PCA}}\hat{\mathbf{W}}_{\mathrm{AR}}[z]\nabla^r[z]$. Estimated source: $\hat{\mathbf{W}}_{\mathbb{R}-\mathrm{ARIMA}}[z]\varphi_v(\mathbf{x})$, or after transforming back to the complex space $\hat{\mathbf{e}} = \varphi_v^{-1}[\hat{\mathbf{W}}_{\mathbb{R}-\mathrm{ARIMA}}[z]\varphi_v(\mathbf{x})]$.

Lemma 2. *In the ℝ-uARMA-IPA task, observation process* $\mathbf{x}(t)$ *is autoregressive and its innovation* $\tilde{\mathbf{x}}(t) := \mathbf{x}(t) - E[\mathbf{x}(t)|\mathbf{x}(t-1), \mathbf{x}(t-2), \ldots]$ *is* $\mathbf{AQ}_0\mathbf{e}(t)$*, 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* $\mathbf{W}_{\mathrm{AR}}[z]\mathbf{x} = \mathbf{AQ}_0\mathbf{e}$.

Thus, AR fit of $\nabla^r[z]\,(\varphi_v[\mathbf{x}(t)])$ can be used for the estimation of $\varphi_M(\mathbf{AQ}_0)\varphi_v[\mathbf{e}(t)]$. This innovation corresponds to the observation of an undercomplete ℝ-ISA model ($D_x > D_e$), which can be reduced to a complete ℝ-ISA ($D_x = D_e$) using principal component analysis (PCA). Finally, the solution can be finished by any ℝ-ISA procedure. The steps of our algorithm are summarized in Fig. 4.

The reduction procedure implies that the derived hidden components $\varphi_v(\mathbf{e}^m)$ can be recovered only up to the ambiguities of the ℝ-ISA task [18]: components of (identical dimensions) can be recovered only up to permutations. Within each subspaces, unambiguity is warranted only up to linear transformations that can be reduced to orthogonal transformations provided that both the hidden source ($\mathbf{e}$) and the observation are white; their expectation values are $\mathbf{0}$ and the covariance matrices are identity matrices. These conditions make no loss to the generality of our solution. Notice that the unitary property of matrix $\mathbf{M}$ is equivalent to the orthogonality of matrix $\varphi_M(\mathbf{M})$ [8]. Thus, apart from a permutation of the components, we can reproduce components $\mathbf{e}^m$ only up to an unitary transformation.

Received 18th July 2007