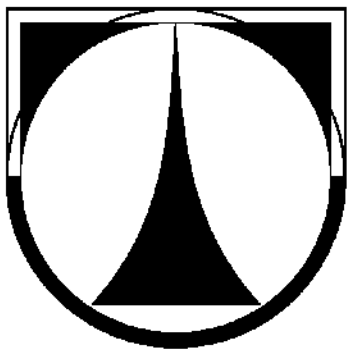
Faculty of Mechatronics, Informatics and
Interdisciplinary Studies,
Technical University of Liberec

Institute of Information Technology and Electronics



Blind Separation of Multichannel Signals by Independent Components Analysis

Habilitation Thesis

Ing. Zbyněk Koldovský, Ph.D.

November 2010

Acknowledgements

The subject matter of this thesis has been evolving for several years and is not a product of one person only. It has arisen thanks to the effort of many people, and in fact, my own contribution is only a part of it. Therefore, I want to thank all the people, my friends, family and colleagues.

My biggest thanks go to my colleague Petr Tichavský from the Institute of Information Theory and Automation. He was the supervisor of my master and doctoral studies, and I learned a lot from him. This encompasses most of my scientific experience that is important to my profession. I have also learned many things for life from him, especially, about the most important matter, which is a question of faith. I am happy that I met him some years ago, which was probably a fluke (or wasn't it a coincidence?). I like and admire his intellectual honesty. It is a pleasure to work with him, and I hope that our collaboration will continue in the future.

I would like to express my thanks to all my colleagues from the Institute of Information Technology and Electronics at the Technical University of Liberec. They kindly took me to their team when I started my work position in Liberec, and now they are my friends. My special thanks belong to professors Jan Nouza and Ondřej Novák, who have done a great job in taking care and supporting our institute. They have helped me many times, even personally. Likewise I would like to thank to the present head of our institute Zdeněk Plíva, who has never refused to help or give advice.

There are two important institutions that have played an important role in my professional growth. It is the Faculty of Nuclear Science and Physical Engineering at the Czech Technical University in Prague and the Institute of Information Theory and Automation at the Academy of Sciences of the Czech Republic. I would like to thank all the teachers and colleagues of these two institutions. Their commitment and enthusiasm motivated me in many aspects.

I had the great chance to collaborate with big names like Erkki Oja, Arie Yeredor, and Yannick Deville. It was pleasant and beneficial experience that I highly appreciate. I also thank to Germán Gómez Herrero from the Department of Signal Processing at the Tampere University of Technology and to Shahram

Hosseini from Laboratoire d'Astrophysique de Toulouse-Tarbes, Université Paul Sabatier.

A na závěr děkuji svým miláčkům domácím Zuzaně a Františkovi a také našim rodičům za lásku a trpělivost a vůbec prostě za všechno.

My apologies if I have omitted anyone to whom acknowledgement is due.

Abstract

The thesis presents a survey of the main achievements in the Blind Source Separation (BSS) area to which the author has been contributing since 2006, that is, since he finished his PhD studies. It is organized as a proceedings of six important articles, which were published in journals or in proceedings of distinguished conferences, accompanied by introductory parts and overviews putting each paper in the right context.

The thesis consists of three parts. The first part is devoted to the underlying problem of Independent Component Analysis (ICA), that is, to the instantaneous mixture model where the number of sensors is equal to the number of original signals. A special emphasis is put on Cramér-Rao bounds that limit the performance of given ICA models. The bounds reveal strengths and limitations of ICA and provide thus a general theoretical insight into its applicability within BSS.

The second part of the thesis deals with the blind separation of mixtures of real-world audio signals based on the ICA. The mixtures are described through the convolutive model, which needs to be transformed into one or more instantaneous problems so that standard ICA algorithms can be applied to it (them). The main result in this context is a time-domain method named T-ABCD. The method is introduced by a brief description while the articles attached to the part describe the method and its variants, including a subband variant, in full details.

The thesis is concluded by the third part that outlines present and future research topics. This encompasses further extensions of T-ABCD in audio BSS applications such as speech enhancement and dereverberation. Next, a brief introduction to the problem of tensor decompositions is given, and their application for underdetermined BSS is described briefly. The Appendix summarizes the main achievements of the author.

Contents

List of Figures

List of Abbreviations

ALS	Alternating least squares
AR	Autoregressive
BSS	Blind Source Separation
CANDECOMP	Canonical Decomposition
CP	Canonical Decomposition
CRLB	Cramér-Rao Lower Bound
DFT	Discrete Fourier Transform
DTFT	Discrete-Time Fourier Transform
ELS	Enhanced line search
FD	Frequency-domain
FFT	Fast Fourier Transform
FIM	Fisher Information Matrix
ICA	Independent Component Analysis
ISA	Independent Subspace Analysis
ISR	Interference-to-Signal Ratio
JAD	Joint Approximate Diagonalization
MDP	Minimal Distortion Principle
MIMO	Multi-input multi-output
MISO	Multi-input single-output
MLE	Maximum Likelihood Estimation
NMF	Non-negative Matrix Factorization
NTF	Non-negative Tensor Factorization

PARAFAC	Parallel Factor Analysis
PCA	Principal Component Analysis
PDF	Probability density function
SCA	Sparse Component Analysis
SDR	Signal-to-Distortion Ratio
SINR	Signal-to-Interference-plus-Noise Ratio
SIR	Signal-to-Interference Ratio
SISO	Single-input single-output
SNR	Signal-to-Noise Ratio
STFT	Short-Time Fourier Transform
TD	Time-domain
TDOA	Time-difference of arrival

Part I

Instantaneous Mixtures

Chapter 1

Blind Source Separation

1.1 Introduction

Blind Source Separation (BSS) represents a wide class of models and algorithms that have one goal in common: to retrieve unknown original signals from their mixtures when only the mixtures are observed, and the mixing system, whose input are the original signals, is also unknown [1]. In other words, the methods aim at separating mixed signals with as little knowledge about the signals and the mixing process as possible. This is, in fact, a general goal of the Signal Processing field: To retrieve signal or information from a signal that is distorted by noise, interferences or other unwanted artifacts, so it is observed as a mixture of the signals. The target signals and artifacts cannot be observed independently nor predicted in most real-world situations, and very little can be assumed about their properties. BSS therefore provides powerful tools since it requires very weak assumptions that are satisfied in various scenarios.

Such general assumptions or properties of signals, which are focused by recent research, are the independence, non-negativity and sparseness. Figure 1.1 shows an example of three recordings of drum instruments. These three signals can be assumed to be independent as they originate from different instruments that behave as independent sound generators (we can neglect the fact that the instruments play according to notes that are organized to a rhythm, which means certain dependence). They are sparse in the time-domain but also in the frequency-domain as can be seen from Fig. 1.2 where an amplitude spectrogram of the single-channel composition of signals is shown. The amplitude spectrogram is an example of a non-negative signal.

Now consider the situation when only the composed signal is available and should be separated into individual instruments without any further knowledge. This is a BSS problem. Since the instruments do not overlap much in time, the

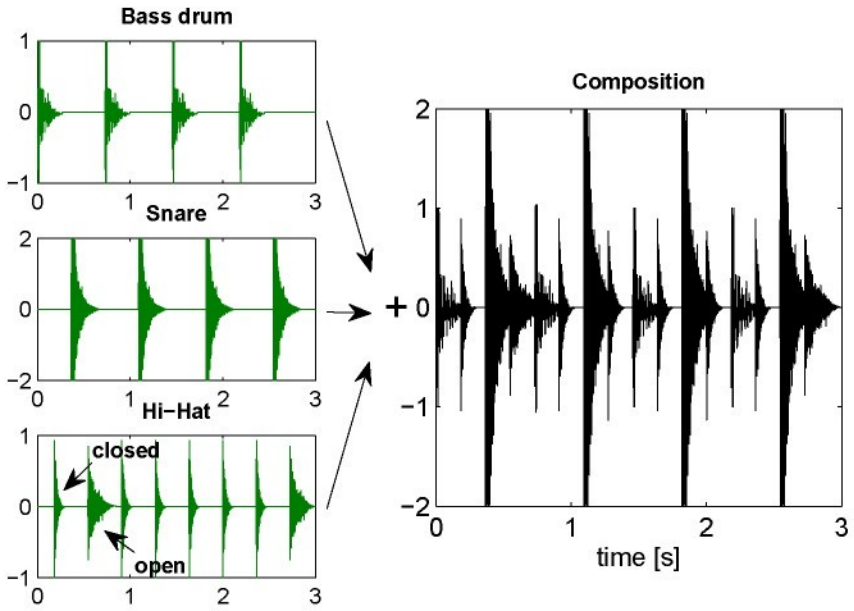


Figure 1.1: Signals of three drum instruments mixed into a single-channel composition.

simplest way would be to split the signal into particular hits and then to assign the separated hits to instruments and reconstruct their signals. This approach relies on the sparsity of the original signals in time. Similarly, the sparsity of the signals in time-frequency domain could be exploited if they were not sparse in time. A discipline that encompasses methods decomposing signals based on their sparsity is called Sparse Component Analysis (SCA) and serves as a tool for BSS.

The amplitude spectrogram in Fig. 1.2, taken as a matrix with non-negative elements, can be decomposed by means of an algorithm for Non-Negative Matrix Factorization (NMF). NMF is recognized as a further method for BSS [2], because it has the favourable ability to blindly identify individual parts of signals or images (the spectrogram can be considered for an image whose parts correspond to the hits of individual drum instruments).

The last property of the drum signals, their independence, can be utilized by Independent Component Analysis (ICA). ICA has become very popular in last two decades, because the independence assumption is physically plausible in many real-world scenarios and provide a strong ability to identify some hidden signals [3]. ICA provides a powerful tool for BSS, and is addressed by this thesis.

SCA, NMF and ICA may also be closely related in miscellaneous situations and applied together since many natural signals embody several of the desired

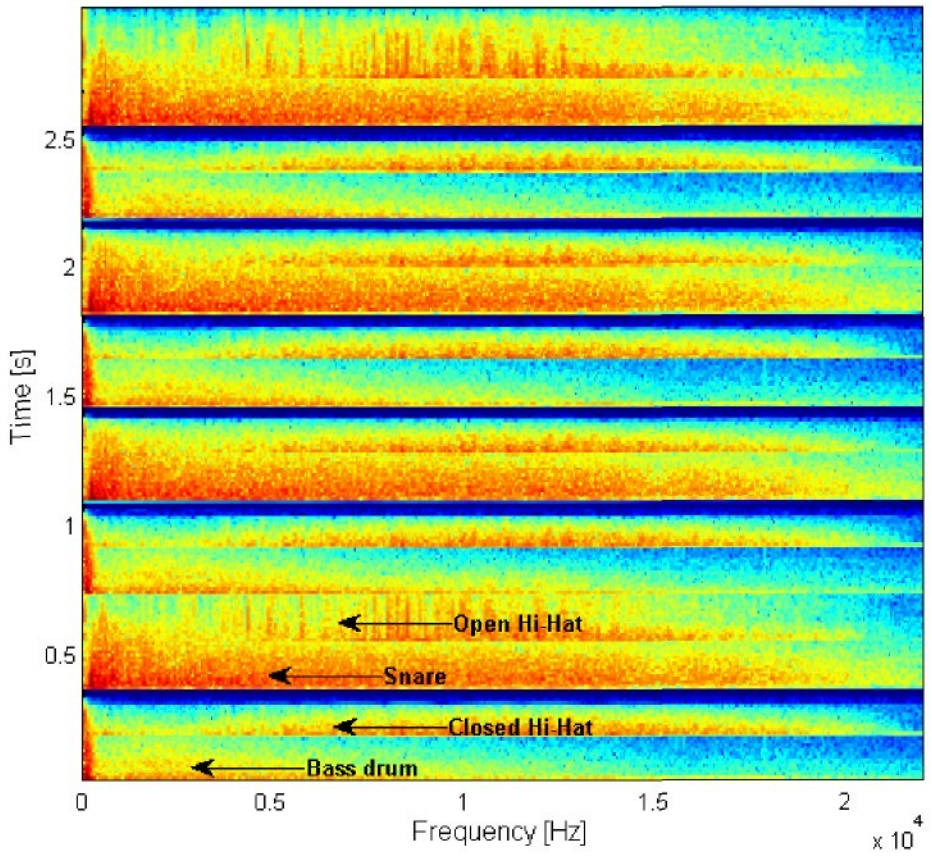


Figure 1.2: A spectrogram of the mixture of drum signals. The sampling rate was 44.1kHz, the length of the Discrete Fourier Transform was 1024 samples, and the length of window overlap was 512 samples.

properties at once.

In some situations, the measurement (recording) may be done by means of several sensors simultaneously, which gives rise to a multi-channel signal. For example, the individual drums can be mixed into a stereo signal in such way that each instrument is added to each channel with a different gain in order to imitate its location. The advantage of multi-channel signals consists in the diversity between signals obtained by sensors, which can be exploited for the separation.

1.2 The Mixing System and Notations

This section introduces basic mixing models and relating notations. A mixing model describes the mechanism how the original signals are mixed into the signals that are observed by sensors. The mixing process proceeds in a linear way in most real-world cases. Therefore, there are two basic linear models considered in the literature: the instantaneous and the convolutive model.

1.2.1 Instantaneous Model

The *instantaneous model* assumes the environment where the speed of propagation of signals is fast enough so that any delays and reflections of signals can be neglected. Each signal is just superposed on each sensor with an attenuation that bears relation to the positions of the source and the sensor. An example is the signal from Electroencephalogram (EEG), where the sampling frequency is very small compared to the speed of propagation of electromagnetic signals from the brain.

The model is described as

$$\mathbf{X} = \mathbf{AS} \tag{1.1}$$

where $\mathbf{X}$ is a $m \times N$ matrix whose rows contains samples of the observed signals from sensors. Since the signals are observed simultaneously, a column of $\mathbf{X}$ corresponds to a time instant. Similarly, $\mathbf{S}$ is a $d \times N$ matrix containing the original unknown signals. $\mathbf{A}$ is a $m \times d$ matrix of attenuations and is called the *mixing matrix* as it parametrizes the whole mixing system. N denotes the number of available samples of signals. The subspace of $\mathbb{R}^N$ spanned by rows of $\mathbf{X}$ will be called the *observation space*.

Individual signals of $\mathbf{X}$ will be denoted as $x_1(n), \ldots, x_m(n)$, so it holds that $x_i(n)$ corresponds to the inth element of $\mathbf{X}$, that is, $\mathbf{X}_{in}$. Similarly, $s_1(n), \ldots$ $\ldots, s_d(n)$ denote the original signals from $\mathbf{S}$.

1.2.2 Convolutive Model

The *convolutive model* is more general than the instantaneous one, because it takes the delays and reflections into account. The relation between the mixed and original signals is described by

$$x_i(n) = \sum_{j=1}^{d} \sum_{\tau=0}^{M_{ij}} h_{ij}(\tau) s_j(n-\tau). \tag{1.2}$$

The mixing system is a MIMO (multi-input multi-output) linear filter with source-sensor impulse responses h_{ij}'s each of length M_{ij}. If $M_{ij} = 0$ for all i and j, the simpler instantaneous model is obtained. The convolutive model will be considered in the second part of this thesis.

1.2.3 Number of Signals and Sensors

When solving the BSS problem, it is either possible to aim at finding the parameters characterizing the mixing system (the matrix $\mathbf{A}$) or to find the original signals themself. The former task is named the *system identification* and the latter names the *signal separation*. The tasks may be equivalent depending on the number of observed and original signals.

In general, two cases are distinguished under the assumption that $\mathbf{A}$ has full rank (otherwise, the system could be reduced). The most studied case is the *determined* one when $m = d$. Then, $\mathbf{A}$ is square, regular, its inversion exist, and it holds that $\mathbf{S} = \mathbf{A}^{-1}\mathbf{X}$. It follows that the finding of $\mathbf{A}$ and $\mathbf{S}$ are equivalent tasks, in this case.

The mixture of signals is called *underdetermined* when $m < d$. This task is more difficult, since the inversion of $\mathbf{A}$ does not exist. The identification of $\mathbf{A}$ and the retrieval of the original signal are no more equivalent, in this case.

1.2.4 Additive Noise

The model with additive noise is described by

$$\mathbf{X} = \mathbf{AS} + \mathbf{N} \tag{1.3}$$

where $\mathbf{N}$ has the same size as $\mathbf{X}$ and contains samples of noise signals. In fact, this model is an underdetermined one with $d + m$ original signals ($\mathbf{S}$ and $\mathbf{N}$) and m observations given by $\mathbf{X}$. However, the model is distinguished by the fact that the primary goal is to retrieve $\mathbf{S}$ only, not $\mathbf{N}$.

Here, the case where $m > d$ is sometimes considered, so there are more noisy observations than the original signals. Such case is sometimes called *overdetermined*, although the model itself is underdetermined. The most common approach here consists in reducing the number of observations to d, e.g. by Principal Component Analysis (PCA). The goal is to suppress the additive noise from the observed signals so that the remaining data can be approximated as a determined problem.

Chapter 2

Independent Component Analysis

In this chapter the determined instantaneous mixture model will be considered. It is the underlying model applied in many situations. For example, the solution of this task might be a building block for separation of more challenging convolutive mixtures addressed in the second part of this thesis.

2.1 Definition

ICA solves the BSS task based on the assumption that the original signals $\mathbf{S}$ are *independent* where the independence is understood as the term from probability theory. Since the original signals are mixed through $\mathbf{A}$, the observed signals $\mathbf{X}$ are dependent, in general. The ICA task thus can be formulated as to estimate the mixing matrix $\mathbf{A}$ or, equivalently, $\mathbf{W} \triangleq \mathbf{A}^{-1}$, called the *de-mixing matrix*, so that signals $\mathbf{A}^{-1}\mathbf{X} = \mathbf{W}\mathbf{X}$ are as independent as possible.

The beginnings of ICA can be dated to 1986 when Herault and Jutten published their paper [6] on a learning algorithm that was able to separate indepednent signals. Later, the concept of ICA was most clearly stated by Comon in [7], which is one of the most cited paper on ICA. Presently, there are several books and proceedings devoted to this important topic of signal processing [1, 3, 4, 5].

2.1.1 Indeterminacies

The solution of the ICA task is not uniquely determined. Any matrix $\mathbf{W}$ of the form

$$\mathbf{W} = \mathbf{\Lambda}\mathbf{P}\mathbf{A}^{-1}, \tag{2.1}$$

where $\mathbf{\Lambda}$ is a diagonal matrix with nonzero diagonal entries and $\mathbf{P}$ is a permutation matrix, separates the original signals from $\mathbf{X}$ up to their original order, scales, and

signs. These features cannot be retrieved without any further knowledge. To a certain extent, the indeterminacies entail a shortcoming of BSS approaches.

The scales of the original signals may be assumed to have arbitrary positive values. Without loss on generality, it will be assumed that the mean value and variance of each signal exist and are equal to zero and one, respectively[1].

Any parameter or statistic that is determined up to the mentioned indeterminacies will be said to be *essentially unique*.

2.1.2 Preprocessing

Let $\mathbf{Z} = \mathbf{C}\mathbf{X}$ where $\mathbf{C}$ is a preprocessing matrix to-be found. A necessary condition for the independence of $\mathbf{Z}$ is that they are not correlated. A consistent estimate of the matrix of correlations of $\mathbf{Z}$ is $\mathbf{Z}\mathbf{Z}^T/N$ (in most cases), so it may be required that $\mathbf{C}$ is such that

$$\mathbf{Z}\mathbf{Z}^T/N = \mathbf{I} \tag{2.2}$$

where $\mathbf{I}$ stands for the identity matrix. When searching for a matrix $\mathbf{U}$ that separates $\mathbf{Z}$, i.e. $\mathbf{U}\mathbf{Z} = \widehat{\mathbf{S}} \approx \mathbf{S}$, the transform should retain $\widehat{\mathbf{S}}$ uncorrelated since $\mathbf{S}$ are independent. This is ensured whenever $\mathbf{U}$ is orthogonal, i.e., $\mathbf{U}\mathbf{U}^T = \mathbf{I}$, and the constraint is called the *orthogonal constraint*.

Applying $\mathbf{C}$ to $\mathbf{X}$ such that (2.2) is fulfilled, is a common preprocessing step used by most ICA methods. Such $\mathbf{C}$ can be easily found as any orthogonalizing transform of $\mathbf{X}$; a particular one is obtained by PCA, where $\mathbf{Z}$ correspond to principal components of $\mathbf{X}$.

Then, the separating matrix $\mathbf{U}$ may be searched under the *orthogonal constraint*. However, such constraint limits the accuracy of separation, since (2.2) is an estimate of the true correlation matrix only. In other words, $\mathbf{S}\mathbf{S}^T/N$ is only approximately equal to $\mathbf{I}$ due to the estimation error, while the orthogonal constraint requires exactly that $\widehat{\mathbf{S}}\widehat{\mathbf{S}}^T/N = \mathbf{I}$. Therefore, algorithms designed to achieve the best possible accuracy try to avoid the constraint.

2.2 Performance Evaluation

To evaluate the accuracy of separation, the mixing matrix or the original signals must be known. Such situation arises, for example, when doing simulations in order to test some algorithms. Some original independent signals are mixed by a known mixing matrix, the tested algorithms are applied to the mixed signals, and the resulting separating matrices or separated signals can be compared with the

[1] In case of a nonstationary signal, its scale is assumed to be equal to one, which is the average variance of all of its samples.

mixing matrix or with the original signals, respectively. The major problem of the evaluation to be overcame is that the order of separated signals is random, which is called the *permutation problem.*

Let $\widehat{\mathbf{W}}$ be an estimated separating matrix produced by an ICA algorithm when the original mixing matrix was $\mathbf{A}$. The so-called *gain matrix* is defined as

$$\mathbf{G} = \widehat{\mathbf{W}}\mathbf{A}. \tag{2.3}$$

Owing to the indeterminacy of ICA, $\mathbf{G}$ should be ideally equal to $\mathbf{\Lambda P}$ from (2.1).

The Amari's index was introduced in [8]. It evaluates the accuracy of separation as a whole by yielding a non-negative value that is equal to zero if and only if $\mathbf{G} = \mathbf{\Lambda P}$. It is defined as

$$I = \sum_{i=1}^{d} \left(\frac{\sum_{j=1}^{d} |\mathbf{G}_{ij}|}{\max_k |\mathbf{G}_{ik}|} - 1 \right) + \sum_{j=1}^{d} \left(\frac{\sum_{i=1}^{d} |\mathbf{G}_{ij}|}{\max_k |\mathbf{G}_{kj}|} - 1 \right). \tag{2.4}$$

The criterion reflects the fact that $\mathbf{G}$ should contain one and only one dominant element in each row and column.

To evaluate each separated signal individually, it is possible to use standard measures such as Signal-to-Interference ratio etc. However, the separated signals must be reordered prior to the evaluation. This means to find the permutation matrix $\mathbf{P}$. A straightforward way is to match the separated and original signals based on dominant elements of $\mathbf{G}$ under the condition that the matched pairs of signals are disjoint. The most common approach, called *greedy*, finds the maximal element of $\mathbf{G}$ in absolute value, assigns the corresponding signals, and repeats the process until all signals are paired. A more sophisticated pairing based on the Kuhn-Munkres pairing algorithm was proposed in [j1].

Once the permutation matrix is found, the separated signals and rows of $\mathbf{G}$ can be reordered. Assume that the correct order was found. The kth separated signal, denoted as $\widehat{s}_k(n)$, is equal to

$$\widehat{s}_k(n) = \mathbf{G}_{k1}s_1(n) + \cdots + \mathbf{G}_{kk}s_k(n) + \cdots + \mathbf{G}_{kd}s_d(n).$$

Since $s_1(n), \ldots, s_d(n)$ are independent (uncorrelated) and are assumed to have zero mean and unit variance, the ratio of energies of the kth (target) signal and the other interfering signals in $\widehat{s}_k(n)$, known as the Signal-to-Interference Ratio (SIR), is equal to

$$\mathrm{SIR}_k = \frac{|\mathbf{G}_{kk}|^2}{\sum_{i=1, i\neq k}^{d} |\mathbf{G}_{ki}|^2}. \tag{2.5}$$

Its reciprocal value is named the Interference-to-Signal Ratio (ISR), that is

$$\mathrm{ISR}_k = \frac{\sum_{i=1, i\neq k}^{d} |\mathbf{G}_{ki}|^2}{|\mathbf{G}_{kk}|^2}. \tag{2.6}$$

In case that the additive noise is taken into account as in (1.3), Signal-to-Interference-plus-Noise Ratio (SINR) of the kth estimated signal is equal to

$$\text{SINR}_k = \frac{|\mathbf{G}_{kk}|^2}{\sum_{i=1,i\neq k}^{d} |\mathbf{G}_{ki}|^2 + \sigma^2 \sum_{i=1}^{d} |\widehat{\mathbf{W}}_{ki}|^2}, \tag{2.7}$$

where σ^2 denotes the variance of the noise signals $\mathbf{N}$ (assumed to be the same for each noise signal).

Monte Carlo Simulations

Many evaluation tests are based on repeated simulations with randomly selected or generated signals, mixing matrices, initializations etc. Each trial is evaluated by using some of the criteria mentioned above. The final criterion is usually the average or median value computed over the trials. It is worth to point out here that the statistical properties of the criteria might or might not be different.

As an example, consider an experiment where three tests (trials) are done. Artificially mixed signals are separated, and the first separated signal is evaluated in terms of SIR. Assume that the algorithm works well in two trials and yields very high SIR equal to A, but in the last trial it fails and gives very small SIR equal to ϵ, where $\epsilon \ll A$. The final average is equal to

$$\frac{A + A + \epsilon}{3} \approx \frac{2}{3}A.$$

Since A is large, the average SIR is large as well so one might conclude that the algorithm works well. However, when ISR is used as the main criterion, then it is equal to $1/A$ in two experiments, which is a small value, and to $1/\epsilon$ in the last experiment, which is a large value. The average ISR is equal to

$$\frac{1}{3}\left(\frac{1}{A} + \frac{1}{A} + \frac{1}{\epsilon}\right) = \frac{2\epsilon + A}{3A\epsilon} \approx \frac{1}{3\epsilon},$$

which is a large number revealing failures of the algorithm. From this point of view, SIR is a more optimistic criterion than ISR.

On the other hand, when using the median instead of the average as the final value, both criteria give the same conclusion (A in SIR and $1/A$ in ISR) showing a good performance of the algorithm. It is a well-known feature of the median that it avoids outlying observations such as the SIR or ISR achieved in the last trial of the hypothetical experiment.

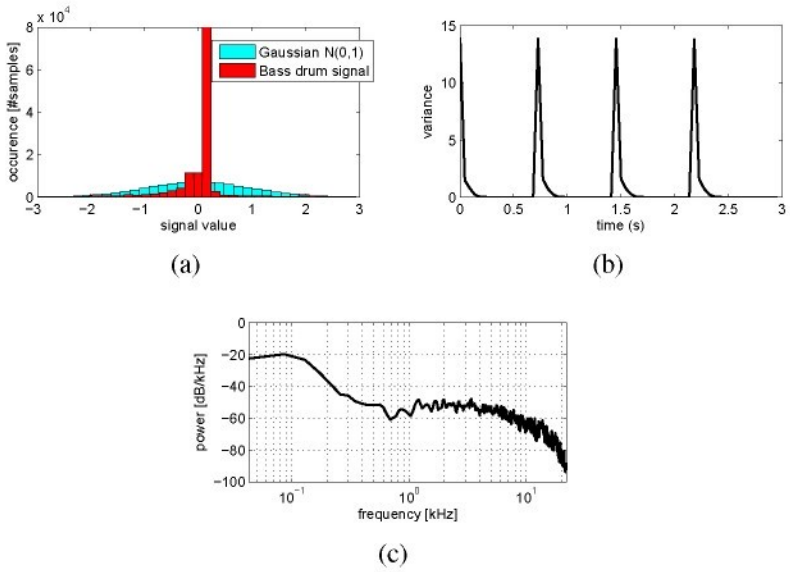


Figure 2.1: Characteristics of the bass drum signal. (a) Histogram compared with that of a standard Gaussian random variable. (b) Variances of the signal in partitioning to 70 blocks of equal length. (c) Power spectral density of the signal.

2.3 Models of Signals

When the original signals come from nature, it is often hard to characterize them. For example, if these signals are speech signals, it is possible to model them in several ways. For different models, the condition of their statistical independence has various forms, which leads to miscellaneous separation principles and methods. It is no wonder that signals possessing several features simultaneously may be separated on the basis of different principles.

There are three basic ICA approaches coming from different statistical models of signals [9]. The first one assumes that a signal is a sequence of identically and independently distributed random variables. The condition of separability of such signals requires that one signal is Gaussian at most, so the approach is said to be based on the *non-Gaussianity*. The second approach takes the *nonstationarity* of signals into account by modeling them as independently distributed Gaussian variables whose variance is changing in time. The third basic model considers weakly stationary Gaussian processes. These signals are separable if their spectra are distinct, therefore, it is said to be based on the *spectral diversity*.

For example, the basic three features, i.e., the non-Gaussianity, nonstationarity and spectral non-whiteness, can be observed on the bass drum signal from Fig. 1.1. This is demonstrated by diagrams in Fig. 2.3 where the histogram verifies the non-Gaussianity, the variance profile proves the nonstationarity, and the power spectral

density shows a non-uniform power spectrum.

The following subsections describe the basic three models and corresponding separation principles in details and provide surveys of several existing algorithms.

2.3.1 Non-Gaussianity

The non-Gaussianity-based model assumes that each original signal is a sequence of i.i.d. random variables [11]. It means that each sample of the ith original signal $s_i(n)$ has the probability density function (PDF) f_{s_i}. Since the signals are assumed to be independent, the joint density of $s_1(n), \ldots, s_d(n)$ is equal to the product of the corresponding marginals

$$f_{s_1,\ldots,s_d} = \prod_{i=1}^{d} f_{s_i}. \tag{2.8}$$

This equation gives a concrete form of the independence condition that the PDFs of separated signals should satisfy.

Most generally it can be said that the ICA algorithms based on the non-Gaussianity optimize a criterion that measures the degree of equality in (2.8) between the PDFs of separated signals. A common criterion used in theory is the *Kullback-Leibler divergence* that is equal, by definition, to the *mutual information* of the separated signals. Let the separated signals be denoted by $\mathbf{Y}$. The mutual information of $\mathbf{Y}$ is defined as

$$I(\mathbf{Y}) = \int_{\mathbb{R}^d} f_{y_1,\ldots,y_d}(\xi_1, \ldots, \xi_d) \ln \frac{f_{y_1,\ldots,y_d}(\xi_1, \ldots, \xi_d)}{\prod_{i=1}^{d} f_{y_i}(\xi_i)} d\xi_1, \ldots, d\xi_d. \tag{2.9}$$

If $\mathbf{Y}$ are not correlated and normalized to have variance equal to one (which is usually achieved by a preprocessing), the mutual information is equal to the sum of entropies of individual signals plus a constant, i.e.,

$$I(\mathbf{Y}) = \sum_{i=1}^{d} H(y_i) + \text{const.}, \tag{2.10}$$

where $H(y_i)$ is the entropy of the ith separated signal defined as

$$H(y_i) = -\int_{\mathbb{R}} f_{y_i}(\xi) \ln f_{y_i}(\xi) d\xi. \tag{2.11}$$

Hence, the minimization of (2.9) is equivalent to the minimization of the entropies of all signals. This clarifies why this approach is called "non-Gaussianity-based", because the entropy of the Gaussian distribution is maximal among all distributions.

In practice, the PDFs of signals are not known and should be either replaced by an appropriate nonlinear function or estimated from the available data. An analogous problem arises when the maximum likelihood estimation (MLE) approach is used, since the densities must be known to build up the likelihood function. Other approaches rely on some necessary (but not sufficient) conditions for the independence following, e.g., from the properties of cumulants.

2.3.2 Nonstationarity

Let the original signals and the mixture be partitioned into M blocks of the same length $N_1 = N/M$, where N_1 is an integer,

$$\mathbf{S} = [\mathbf{S}^{(1)}, \ldots, \mathbf{S}^{(M)}] \tag{2.12}$$

$$\mathbf{X} = [\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(M)}]. \tag{2.13}$$

Assume that each signal in each block $\mathbf{S}^{(\ell)}$ is Gaussian i.i.d., with zero mean and a variance $\sigma_k^{2(\ell)}$, where $k = 1, \ldots, d$ is the index of the signal and $\ell = 1, \ldots, M$ is the index of the block. It means that each signal is parameterized by M (unknown) parameters corresponding to variances on blocks. Since the variance may vary from block to block, the model embodies the nonstationarity of signals [10].

The received data are, by assumption, Gaussian distributed, so the sufficient statistics for estimating $\mathbf{A}$ and the variances is the set of sample covariance matrices

$$\widehat{\mathbf{R}}_\ell = \frac{1}{N_1}\mathbf{X}^{(\ell)}(\mathbf{X}^{(\ell)})^T, \qquad \ell = 1, \ldots, M. \tag{2.14}$$

Theoretical covariance matrices obey the relation

$$\mathbf{R}_\ell = \mathbf{A}\,\mathrm{diag}\left[\sigma_1^{2(\ell)}, \ldots, \sigma_d^{2(\ell)}\right]\mathbf{A}^T \tag{2.15}$$

where diag$[\cdot]$ is a diagonal square matrix containing the elements of argument on its diagonal. In fact, diag$[\sigma_1^{2(\ell)}, \ldots, \sigma_d^{2(\ell)}]$ is the covariance matrix of $\mathbf{S}^{(\ell)}$, because the signals are independent.

This gives rise to the idea how to separate such signals: The covariance matrices on blocks of transformed signals $\mathbf{Y} = \mathbf{WX}$ are equal to $\mathbf{WR}_\ell\mathbf{W}^T$, $\ell = 1, \ldots, M$. Hence, the separating matrix $\mathbf{W}$ can be found as such that provides a approximate joint diagonalization of the matrices $\{\widehat{\mathbf{R}}_\ell\}$, i.e. has the property that the matrices $\{\mathbf{W}\widehat{\mathbf{R}}_\ell\mathbf{W}^T\}$ are all approximately diagonal.

The joint approximate diagonalization (JAD) of a set of matrices can be performed in several ways, optimizing several possible criteria. A straightforward one is

$$C_{\mathrm{LS}}(\mathbf{W}) = \sum_{\ell=1}^{M} \|\mathrm{off}\left(\mathbf{W}\widehat{\mathbf{R}}_\ell\mathbf{W}^T\right)\|_F, \tag{2.16}$$

where the operator "off" nullifies the diagonal elements of a matrix, and $\|\cdot\|_F$ stands for the Frobenius norm. A minimization of this criterion must be constrained, so as to evade trivial minimization by $\mathbf{W} = 0$. For instance, the orthogonal constraint entails that $\mathbf{W}$ cannot be zero.

It can be shown, that the maximum likelihood estimator of the separating matrices is realized by minimizing the criterion

$$C_{\mathrm{LL}}(\mathbf{W}) = \sum_{\ell=0}^{M-1} \log \frac{\det \operatorname{ddiag}(\mathbf{W}\widehat{\mathbf{R}}_\ell \mathbf{W}^T)}{\det(\mathbf{W}\widehat{\mathbf{R}}_\ell \mathbf{W}^T)}, \tag{2.17}$$

where the operator "ddiag" nullifies the off-diagonal elements of a square matrix. This criterion is meaningful only for positive definite matrices $\{\widehat{\mathbf{R}}_\ell\}$.

The condition that $\mathbf{W}$ jointly diagonalizes $\{\mathbf{R}_\ell\}$ is a sufficient condition for the independence of separated signals, because the model considered here assumes their distribution to be Gaussian. However, in a more general manner, the condition is necessary only. Some authors therefore do not consider methods based on the joint diagonalization of correlation matrices for ICA methods, but rather for more general-purpose BSS methods. The following example reveals more in this respect.

Assume that r original signals (without loss on generality, let they be the first r signals) have the same dynamic profiles, so it holds that $\sigma_1^{2(\ell)} = \cdots = \sigma_r^{2(\ell)}$, for each $\ell = 1, \ldots, M$. It can be easily verified that any orthogonal transform of these signals keeps their covariance matrices diagonal on all blocks. It follows that the joint diagonalization of $\{\widehat{\mathbf{R}}_\ell\}$ is then essentially unique up to an orthogonal transform of the subspace spanned by the signals.

In summary, signals having the same dynamic profiles cannot be separated through the joint diagonalization of covariance matrices. Note, however, that the subspaces of signals having different dynamics can be separated each from the other, which is a fact addressed in the next chapter.

2.3.3 Spectral Diversity

The third signal model assumes that the original signals are independent weak stationary Gaussian processes [12]. As will be shown, it relies on the fact that such signals are distinguishable in the frequency domain. It is practical to limit the theoretical considerations to autoregressive (AR) processes with known order.

A sufficient statistic for joint estimation of the separating matrix and autoregressive parameters of the signals is the set of the estimated time-lagged correlation matrices,

$$\widehat{\mathbf{R}}[\tau] = \frac{1}{N-\tau} \sum_{n=1}^{N-\tau} \mathbf{x}[n]\mathbf{x}^T[n+\tau], \qquad \tau = 0, \ldots, R-1, \tag{2.18}$$

where $\mathbf{x}[n]$ denotes the nth column of $\mathbf{X}$, $R-1$ is the order of the AR model, and ^{T} denotes the matrix/vector transposition. The true covariance matrices are equal to

$$\mathbf{R}[\tau] = \mathbf{A}\,\mathrm{diag}\big[r_1[\tau], \ldots, r_d[\tau]\big]\mathbf{A}^T, \tag{2.19}$$

where $r_k[\tau]$ is the autocorrelation of $s_k(n)$. The time-lagged correlation matrix of the original signals is $\mathrm{diag}\big[r_1[\tau], \ldots, r_d[\tau]\big]$, which is diagonal thanks to their independence. Therefore, the approach here is the same as in the previous nonstationarity-based model. The separating matrix $\mathbf{W}$ is estimated as a joint diagonalizer of the set of matrices $\{\widehat{\mathbf{R}}[\tau]\}$.

Analogously to the previous model, signals having the same autocorrelations cannot be separated via the joint diagonalization of $\{\widehat{\mathbf{R}}[\tau]\}$. The autocorrelations determine the signal spectra and are the same if and only if the spectra are the same. Therefore, the model here is said to rely on the spectral diversity of signals.

2.3.4 Hybrid Models

Since natural signals usually embody several features such as the non-Gaussianity, nonstationarity and spectral non-whiteness simultaneously, it is meaningful to combine the above models in order to derive more flexible and accurate ICA methods. However, the complexity of models increases by combining more properties of signals, so only two models are usually put together to provide a reasonably complex theoretical background.

The other way is to derive an algorithm heuristically without deep theoretical considerations. For instance, it is possible to built up a separating criterion that is a weighted sum of particular contrasts which reflect various properties of signals. Similarly it is possible to do JAD of covariance and cross-covariance matrices and cumulant and cross-cumulant slices based on a weighted criterion of diagonality.

Piecewise Stationary Non-Gaussian Signals

Many audio signals, especially speech signals, are known to exhibit both the non-Gaussianity and nonstationarity. The composite model taking both features into account considers each sample of a signal as an independent zero-mean random variable that need not have a fixed distribution (nor a fixed variance) [10]. To allow the estimation of the distribution, it is practical to assume a piecewise stationarity so that the distribution is fixed within time intervals of a certain length. Let the number of intervals (blocks of data) be M and their length be N_1, which is the same partitioning as that in Section 2.3.2.

The model introduced here provides a generalization of the non-Gaussianity and nonstationarity-based models described in previous subsections. While the

former model coincides with the piecewise stationary one when the number of blocks is $M = 1$, the latter model corresponds with the case when signals are restricted to have Gaussian distributions on all blocks.

In notations a superscript in parentheses, e.g. $^{(\ell)}$, will be used to denote quantities, data or probability density functions that are related to the ℓth block.

Piecewise Stationary Autoregressive Processes

The model based on spectral diversity of signals can be generalized in an analogous way. It results in a hybrid model that involves the nonstationarity [25]. The data are partitioned into M blocks of the same length. In each block, the signals are modeled as Gaussian AR processes of the order R with zero mean and arbitrary variance. A sufficient statistic is the set of time-lagged correlation matrices computed on all blocks, that is,

$$\widehat{\mathbf{R}}^{(\ell)}[\tau] = \frac{1}{N_1 - \tau} \sum_{n=1}^{N_1-\tau} \mathbf{x}^{(\ell)}[n] \big(\mathbf{x}^{(\ell)}[n+\tau]\big)^T, \tag{2.20}$$

where $\mathbf{x}^{(\ell)}[n]$ is the nth column of $\mathbf{X}^{(\ell)}$, $\tau = 0, \ldots, R-1$ and $\ell = 1, \ldots, M$. The separating matrix $\mathbf{W}$ is searched as a joint approximate diagonalizer of these matrices.

2.4 Performance Bounds

2.4.1 Cramér-Rao-Induced Bounds for Residual Interference-to-Signal Ratio

Section 2.2 introduces performance measures that can be used to evaluate the accuracy of particular data separated by a given algorithm. To see how the algorithm performs in a more general manner, the experiment can be repeated (Monte Carlo trials). If the average of a selected criterion converges, its value is indicative of some statistical properties of the algorithm. On the other hand, a theoretical analysis of the algorithm can be done by deriving closed-form formulas of the true mean value of the criterion under a given model of signals.

The common criterion considered in this section is the mean ISR defined in (2.6). To this end, the ISR matrix is introduced as such that its elements are defined as

$$\mathrm{ISR}_{ij} = \mathrm{E}\left[\frac{|\mathbf{G}_{ij}|^2}{|\mathbf{G}_{ii}|^2}\right] \frac{\sigma_j^2}{\sigma_i^2}, \tag{2.21}$$

where E stands for the expectation operator, and $\widehat{\mathbf{W}}$ is the separating matrix estimated by the algorithm, $\mathbf{G} = \widehat{\mathbf{W}}\mathbf{A}$ is the gain matrix, and σ_i^2 is the variance of the ith signal, or, more generally, its scale (the average of variances over blocks, in case of piecewise stationary signals).

Cramér-Rao lower bound (CRLB) is a general bound for the variance of unbiased estimators [39]. Consider a vector of parameters $\boldsymbol{\theta}$ being estimated from a data vector $\mathbf{x}$, having probability density $f_{\mathbf{x}|\boldsymbol{\theta}}(\mathbf{x}|\boldsymbol{\theta})$. Let $\hat{\boldsymbol{\theta}}$ be some unbiased estimator of $\boldsymbol{\theta}$. If the following *Fisher information matrix* (FIM) exists

$$\mathbf{F}_{\boldsymbol{\theta}} = \mathrm{E}_{\boldsymbol{\theta}}\left[\frac{1}{f_{\mathbf{x}|\boldsymbol{\theta}}^2}\frac{\partial f_{\mathbf{x}|\boldsymbol{\theta}}(\mathbf{x}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\left(\frac{\partial f_{\mathbf{x}|\boldsymbol{\theta}}(\mathbf{x}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right)^T\right], \tag{2.22}$$

then under mild regularity conditions [39], it holds that

$$\operatorname{cov}\hat{\boldsymbol{\theta}} \geq \mathrm{CRLB}_{\boldsymbol{\theta}} = \mathbf{F}_{\boldsymbol{\theta}}^{-1},$$

where $\operatorname{cov}\hat{\boldsymbol{\theta}}$ is the covariance matrix of $\hat{\boldsymbol{\theta}}$.

The Cramér-Rao theory can be used to derive algorithm-independent bounds for elements of the ISR matrix (2.21). It is assumed that $\mathbf{G} = \mathbf{I} + \boldsymbol{\epsilon}$ where $\boldsymbol{\epsilon}$ is a "small" matrix of errors, so it is assumed that the indeterminacies of ICA were correctly resolved. Then the elements of ISR matrix can be approximated as

$$\mathrm{ISR}_{ij} \approx \mathrm{E}[|\boldsymbol{\epsilon}_{ij}|^2]\frac{\sigma_j^2}{\sigma_i^2}, \tag{2.23}$$

and the lower bound can be defined as the CRLB for $\boldsymbol{\epsilon}$. The problem of transforming the CRLB of $\boldsymbol{\epsilon}$ to the bound of ISR matrix is analyzed in more details in [36].

Piecewise Stationary Non-Gaussian Signals

The CRLB bounds were already derived in [10, j6] for the hybrid models described in Section 2.3.4, which gives also the bounds for the basic models as special cases. For the piecewise stationary non-Gaussian signals the CRLB bound says that

$$\mathrm{ISR}_{ij} \geq \frac{1}{N}\cdot\frac{A_{ij}}{A_{ij}A_{ji}-1}\cdot\frac{\sigma_j^2}{\sigma_i^2}, \qquad i \neq j, \tag{2.24}$$

where

$$A_{ij} = \frac{1}{M}\sum_{\ell=1}^{M}\frac{\sigma_i^{2(\ell)}}{\sigma_j^{2(\ell)}}\kappa_j^{(\ell)} \tag{2.25}$$

$$\sigma_i^2 = \frac{1}{M}\sum_{\ell=1}^{M}\sigma_i^{2(\ell)}. \tag{2.26}$$

$\kappa_i^{(\ell)}$ is defined as

$$\kappa_i^{(\ell)} = \mathrm{E}\left[\left(\psi_i^{(\ell)}(x)\right)^2\right] \tag{2.27}$$

where $\psi_i^{(\ell)} = -\left(\bar{f}_i^{(\ell)}\right)' / \bar{f}_i^{(\ell)}$ is the score function of the density $\bar{f}_i^{(\ell)}$, which is the PDF of the ith signal on the ℓth block, i.e. $f_i^{(\ell)}$, but normalized to the unit variance (the variance of $f_i^{(\ell)}$ is involved in $\sigma_i^{2(\ell)}$). It can be shown that $\kappa_i^{(\ell)} \geq 1$ where the equality holds if and only if $f_i^{(\ell)}$ is Gaussian.

Two key properties of the CRLB can be seen from the expression (2.24). First, the bound is independent of the mixing matrix $\mathbf{A}$, which could be expected, because ISR is also independent of $\mathbf{A}$. Second, ISR_{ij} depends on the characteristics of the ith and jth signal only, which means that the minimum residual interference of one source in another depends on properties of the two signals only, and is independent of the other signals in the mixture.

The bound for the basic non-Gaussianity based model [c1] is obtained by taking $M = 1$. The superscript signifying the index of block can be omitted, and then (2.24) reduces to

$$\mathrm{ISR}_{ij} \geq \frac{1}{N} \frac{\kappa_j}{\kappa_i \kappa_j - 1}, \qquad i \neq j. \tag{2.28}$$

The denominator of (2.28) becomes equal to zero only if both κ_i and κ_j are equal to one, which means that both the ith and jth signals have Gaussian distribution. In such case the bound signifies that the two signals cannot be separated by any unbiased estimator. This is in accordance with the primary requirement that one original signal has the Gaussian PDF at most so as the mixture be separable. It can also be seen, that the bound is minimized when $\kappa_i \to +\infty$ and $\kappa_j \to +\infty$, which can be interpreted as that the signals are non-Gaussian as much as possible.

The basic nonstationarity-based model assumes Gaussian distributions of signals on all blocks, which means that $\kappa_i^{(\ell)} = 1$ for all i and ℓ. This simplifies the expression (2.25), and the bound becomes the form

$$\mathrm{ISR}_{ij} \geq \frac{1}{N} \cdot \frac{\frac{1}{M}\sum_{\ell=1}^{M} \frac{\sigma_i^{2(\ell)}}{\sigma_j^{2(\ell)}}}{\left(\frac{1}{M}\sum_{\ell=1}^{M} \frac{\sigma_i^{2(\ell)}}{\sigma_j^{2(\ell)}}\right)\left(\frac{1}{M}\sum_{\ell=1}^{M} \frac{\sigma_j^{2(\ell)}}{\sigma_i^{2(\ell)}}\right) - 1} \cdot \frac{\sum_{\ell=1}^{M} \sigma_j^{2(\ell)}}{\sum_{\ell=1}^{M} \sigma_i^{2(\ell)}}, \qquad i \neq j. \tag{2.29}$$

Remind that the condition of separability requires that the variance profiles of signals are different. This can be seen from (2.29), because the denominator is equal to zero if $\sigma_i^{2(\ell)} = C \cdot \sigma_j^{2(\ell)}$, where $C > 0$, for all ℓ.

Piecewise Stationary Autoregressive Processes

The expression for the bound of the hybrid model connecting nonstationarity and spectral diversity principles is similar to (2.24). It was derived in [c16] in the form

$$\mathrm{ISR}_{ij} \geq \frac{1}{N} \cdot \frac{\bar{\phi}_{ij}}{\bar{\phi}_{ij}\bar{\phi}_{ji} - 1} \cdot \frac{\bar{r}_j[0]}{\bar{r}_i[0]} \tag{2.30}$$

where

$$\bar{\phi}_{ij} = \frac{1}{M} \sum_{\ell=1}^{M} \frac{\sigma_i^{2(\ell)}}{\sigma_j^{2(\ell)}} \phi_{ij}^{(\ell)} \tag{2.31}$$

$$\phi_{ij}^{(\ell)} = \frac{1}{\sigma_i^{2(\ell)}} \sum_{p,q=0}^{R} a_{j,p}^{(\ell)} a_{j,q}^{(\ell)} r_i^{(\ell)}[p-q] \tag{2.32}$$

$$\bar{r}_j[0] = \frac{1}{M} \sum_{\ell=1}^{M} r_j^{(\ell)}[0]. \tag{2.33}$$

The parameters $a_{i,p}^{(\ell)}$, $p = 1, \ldots, R$, are the AR coefficients of the ith signal in the ℓth block with $a_{i,0}^{(\ell)} = 1$, i.e., $s_i^{(\ell)}(n) = -\sum_{p=1}^{R} a_{i,p}^{(\ell)} s_i^{(\ell)}(n-p) + w_i^{(\ell)}(n)$. Note that here $\sigma_i^{2(\ell)}$ denotes the variance of the Gaussian innovation $w_i^{(\ell)}(n)$ that generates the AR process.

For the basic model purely based on the spectral diversity, the number of blocks is $M = 1$, and (2.30) simplifies to

$$\mathrm{ISR}_{ij} \geq \frac{1}{N} \cdot \frac{\phi_{ij}}{\phi_{ij}\phi_{ij} - 1} \cdot \frac{r_j[0]}{r_i[0]} \tag{2.34}$$

where

$$\phi_{ij} = \frac{1}{\sigma_i^2} \sum_{p,q=0}^{R} a_{j,p} a_{j,q} r_i[p-q]. \tag{2.35}$$

An alternative expression of (2.35) is

$$\phi_{ij} = \frac{1}{2\pi\iota} \oint \frac{A_j(z) A_j^*(\frac{1}{z^*}) z^{-1}}{A_i(z) A_i^*(\frac{1}{z^*})} dz \tag{2.36}$$

where $A_i(z) = \sum_{p=0}^{R} a_{i,p} z^{-p}$, * denotes the complex conjugation, and ι denotes the imaginary unit. Since $A_i(z)$ and $A_j(z)$ determine spectra of the signals, it is easily seen that $\phi_{ij} = 1$ if and only if $A_i(z) = A_j(z)$, i.e., the spectra are the same. It also follows that the denominator in (2.34) is zero if and only if the spectra of the ith and jth signal are proportional.

For $R = 0$, the model corresponds with the basic nonstationarity-based one. Then $\phi_{ij}^{(\ell)} = 1$ and $r_i^{(\ell)}[0] = \sigma_i^{2(\ell)}$ for all i, j, and ℓ, and the same expression as (2.29), which was derived from (2.24), is obtained.

2.4.2 Minimum Signal-to-Interference-plus-Noise Ratio

In this section, the mixing model with the additive noise (1.3) will be considered. The noise signals, i.e. the rows of $\mathbf{N}$, are assumed to be Gaussian and uncorrelated, which simultaneously means that they are independent.

The common criterion used for the evaluation of separated noisy signals [c6] is the SINR defined by (2.7). While SIR, which ignores the noise, can be arbitrarily large as the estimated separating matrix $\widehat{\mathbf{W}}$ approaches $\mathbf{A}^{-1}$ (up to indeterminacies), SINR is bounded provided that the variance of additive noise $\sigma^2 > 0$. The maximum SINR is achieved for

$$\widehat{\mathbf{W}} = \mathbf{W}^{\mathrm{MMSE}} = \mathbf{A}^T(\mathbf{A}\mathbf{A}^T + \sigma^2\mathbf{I})^{-1}, \tag{2.37}$$

which simultaneously minimizes the mean square distance between the original and separated signals, i.e.

$$\mathbf{W}^{\mathrm{MMSE}} = \arg\min_{\widehat{\mathbf{W}}} \mathrm{E}[\mathbf{S} - \widehat{\mathbf{W}}\mathbf{X}]^2. \tag{2.38}$$

By putting $\mathbf{W}^{\mathrm{MMSE}}$ into (2.7), the ultimate bound for the SINR of the kth signal is obtained as

$$\min \mathrm{SINR}_k = \frac{\mathbf{V}_{kk}^2}{\sum_{i\neq k}^{d} \mathbf{V}_{ki}^2 + \sigma^2 \sum_{i=1}^{d} (\mathbf{V}\mathbf{A}^{-1})_{ki}^2} \tag{2.39}$$

where $\mathbf{V} = (\mathbf{I} + \sigma^2(\mathbf{A}^T\mathbf{A})^{-1})^{-1}$. In contrast to the Cramér-Rao bounds where the additive noise was not considered, (2.39) depends on the mixing matrix $\mathbf{A}$ and, naturally, on the noise variance σ^2.

A special case is when $\mathbf{A}$ is an orthogonal matrix satisfying $\mathbf{A}\mathbf{A}^T = \mathbf{I}$. In such case, the model $\mathbf{X} = \mathbf{A}\mathbf{S} + \mathbf{N}$ can be rewritten as

$$\mathbf{X} = \mathbf{A}\mathbf{S} + \mathbf{A}\mathbf{N} = \mathbf{A}(\mathbf{S} + \mathbf{N}),$$

since the noisy signals $\mathbf{A}\mathbf{N}$ are statistically equivalent to $\mathbf{N}$. It follows that the model behaves like the ordinary instantaneous model without the additive noise where the original independent sources are $\widetilde{\mathbf{S}} = \mathbf{S} + \mathbf{N}$. Therefore, in this case the noise is sometimes called the *source noise*.

The asymptotic expansion of (2.39) for "small" σ^2 gives

$$\min \mathrm{SINR}_k = \frac{1}{\sigma^2 \|\mathbf{w}_k\|^2} - B + \mathcal{O}(\sigma^2), \tag{2.40}$$

where

$$B = 2 + \frac{1}{\|\mathbf{w}_k\|^4} \left(\sum_{i\neq k}^{d} (\mathbf{W}\mathbf{W}^T)_{ki}^2 - 2 \sum_{i=1}^{d} \mathbf{W}_{ki} (\mathbf{W}\mathbf{W}^T\mathbf{W})_{ki} \right),$$

$\mathbf{W} = \mathbf{A}^{-1}$, and $\mathbf{w}_k^T$ denotes the kth row of $\mathbf{W}$. The first term in the asymptotic expansion reveals that if the rows of $\mathbf{A}^{-1}$ have the same norm, the ultimate bound (2.39) is approximately the same for each signal (provided that $\mathbf{A}$ is well conditioned).

2.4.3 Bias

If the additive noise is present in the mixed signals as in (1.3), the estimation of $\mathbf{A}$ or $\mathbf{A}^{-1}$ may be biased. To get an unbiased estimate, the covariance matrix of the noise, which is $\sigma^2\mathbf{I}$, must be known a priori [43]. However, the optimum signals in terms of SINR (2.7), which is more natural criterion than ISR when the noise is present, are

$$\mathbf{S}^{\mathrm{MMSE}} = \mathbf{W}^{\mathrm{MMSE}}\mathbf{X}. \tag{2.41}$$

It is thus legitimate to aim at finding $\mathbf{W}^{\mathrm{MMSE}}$ rather than identifying $\mathbf{A}$ or $\mathbf{A}^{-1}$.

Consequently, the bias of an estimated separating matrix $\widehat{\mathbf{W}}$ is studied as the difference of

$$\mathrm{E}[\widehat{\mathbf{W}}](\mathbf{W}^{\mathrm{MMSE}})^{-1} \tag{2.42}$$

from a diagonal matrix. Owing to the indeterminable scale of separated signals, it is common practice to estimate $\widehat{\mathbf{W}}$ so that it yields normalized signals. Then, the bias should be defined as the difference between (2.42) and $\mathbf{D}$, which is the diagonal matrix that normalizes $\mathbf{S}^{\mathrm{MMSE}}$. It can be shown that

$$\mathbf{D} = \mathbf{I} + \frac{1}{2}\sigma^2 \mathrm{diag}[\mathbf{V}_{11}, \ldots, \mathbf{V}_{dd}] + \mathcal{O}(\sigma^3), \tag{2.43}$$

where $\mathbf{V} = (\mathbf{A}^T\mathbf{A})^{-1}$.

Bias of algorithms using the orthogonal constraint

The orthogonal constraint requires that

$$\mathrm{E}[\widehat{\mathbf{W}}\mathbf{X}(\widehat{\mathbf{W}}\mathbf{X})^T] = \widehat{\mathbf{W}}(\mathbf{A}\mathbf{A}^T + \sigma^2\mathbf{I})\widehat{\mathbf{W}}^T = \mathbf{I}. \tag{2.44}$$

It follows that the bias of all constrained algorithms is lower bounded by

$$\min_{\widehat{\mathbf{W}}(\mathbf{A}\mathbf{A}^T + \sigma^2\mathbf{I})\widehat{\mathbf{W}}^T = \mathbf{I}} \|\widehat{\mathbf{W}}(\mathbf{W}^{\mathrm{MMSE}})^{-1} - \mathbf{D}\|_F \tag{2.45}$$

where the minimization proceeds for $\widehat{\mathbf{W}}$. It was shown in [m6] that the minimizer $\widehat{\mathbf{W}}$ of (2.45) fulfills that $\widehat{\mathbf{W}}(\mathbf{W}^{\mathrm{MMSE}})^{-1} = \mathbf{I} + \sigma^2\mathbf{\Gamma} + \mathcal{O}(\sigma^3)$, where $\mathbf{\Gamma}$ is a nonzero matrix obeying $\mathbf{\Gamma} + \mathbf{\Gamma}^T = \mathbf{V}$. It follows that the average bias[2] of all ICA algorithms that use the orthogonal constraint has the asymptotic order $\mathcal{O}(\sigma^2)$.

[2] The average bias computed over all separated signals.

2.5 Algorithms

This section provides a brief survey of several ICA algorithms. It is organized into subsections each corresponding to a given model of signals.

2.5.1 Methods Based on the Non-Gaussianity

FastICA

A reasonably accurate and fast ICA algorithm can be obtained by minimizing a contrast function, which can be quite arbitrary nonlinear and non-quadratic statistic of the data, such as kurtosis. The statistic works as an approximation or a subrogate of the entropy. An example of this approach is the popular FastICA algorithm [15].

The finding of the kth row of the de-mixing matrix $\mathbf{W}$, denoted as $\mathbf{w}_k^T$, in FastICA proceeds by optimizing the contrast function

$$c(\mathbf{w}_k) = \hat{\mathrm{E}}[G(\mathbf{w}_k^T \mathbf{Z})], \tag{2.46}$$

where $\hat{\mathrm{E}}$ stands for the sample mean estimator, and $\mathbf{Z}$ is the preprocessed matrix $\mathbf{X}$ such that rows of $\mathbf{Z}$ are orthogonal. G is a properly chosen nonlinear function whose derivative will be denoted by g. The signals obtained by the current (and final) $\mathbf{W}$ will be denoted by $\mathbf{Y}$, i.e. $\mathbf{Y} = \mathbf{W}\mathbf{Z}$.

Ideally, $c(\mathbf{w}_k)$ should be an estimate of the entropy of y_k, which is achieved when G is equal to $-\ln f_{y_k}$, and g is the score function $\psi_k = -f'_{y_k}/f_{y_k}$. The original FastICA utilizes fixed choice of G, e.g., such that $g(x) = x^3$ or $g(x) = \tanh(x)$. The version of the algorithm that estimates one signal only is called *one-unit* FastICA. The estimation of the whole $\mathbf{W}$ proceeds by finding all local extrema of $c(\mathbf{w}_k)$ on the unit sphere. The deflation approach estimates $\mathbf{W}$ row by row so that each row must be orthogonal to the previous ones. Another approach, called *Symmetric* FastICA, orthogonalizes all rows of $\mathbf{W}$ after each iteration by means of the symmetric orthogonalization. It means that both the deflation and symmetric approach apply the orthogonal constraint.

From the analysis [j2] of the statistical and global convergence properties follows that the theoretical accuracy of the one-unit and symmetric algorithms, in terms of the residual ISR, is expressed, respectively, by

$$\mathrm{ISR}_{ij}^{1U} \approx \frac{1}{N}\frac{\gamma_i}{\tau_i^2} \tag{2.47}$$

$$\mathrm{ISR}_{ij}^{SYM} \approx \frac{1}{N}\frac{\gamma_i + \gamma_j + \tau_j^2}{(\tau_i + \tau_j)^2} \tag{2.48}$$

where

$$\mu_i = \mathrm{E}[s_i g_i(s_i)] \tag{2.49}$$
$$\rho_i = \mathrm{E}[g_i'(s_i)] \tag{2.50}$$
$$\beta_i = \mathrm{E}[g_i^2(s_i)] \tag{2.51}$$
$$\gamma_i = \beta_i - \mu_i^2 \tag{2.52}$$
$$\tau_i = |\mu_i - \rho_i|, \tag{2.53}$$

provided that the expectations exist. Here, s_i denotes the random variable of the same distribution as the ith original signal. g_i is the nonlinearity used to substitute the score function of the ith signal. In the original FastICA, a common nonlinearity, denoted without the index, is used for all signals.

The results of the analysis reveal several interesting facts. The accuracy of the one-unit algorithm given by (2.47) depends on the distribution of the signal being estimated (the ith one) but independent of the other signals. On the other hand, (2.48) depends on distributions of the ith and jth signal, where the same property have the Cramér-Rao bounds. The nonlinearity g has an influence on the residual ISR through the quantities μ_i, ρ_i and β_i defined above. It can be shown that both (2.47) and (2.48) are minimized when g is equal to the score function of the signals' PDF (assuming that all signals have the same distribution). Then, $\mu_i = 1$ and $\rho_i = \beta_i = \kappa_i = \kappa$, where κ is defined through (2.27), and

$$\mathrm{ISR}_{ij}^{1U} \approx \frac{1}{N}\frac{1}{\kappa - 1} \tag{2.54}$$
$$\mathrm{ISR}_{ij}^{SYM} \approx \frac{1}{N}\left(\frac{1}{4} + \frac{1}{2}\frac{1}{\kappa - 1}\right). \tag{2.55}$$

Remind that the corresponding CRLB (2.28) when all signals have the same PDF says that

$$\mathrm{ISR}_{ij} \geq \frac{1}{N}\frac{\kappa}{\kappa^2 - 1}, \tag{2.56}$$

from which it follows that none of the versions of FastICA achieves the bound. A comparison of the expressions is given by Fig. 2.2. Finally, it is worth to point to the first term in (2.55). It is a constant independent of κ limiting the ISR, which is caused by the orthogonal constraint applied within Symmetric FastICA.

Infomax

The Infomax algorithm was first published in [35] and then its extended version in [14]. It is derived on the basis of MLE. A model-log-likelihood function is maximized, which, in principle, also leads to the minimization of approximated

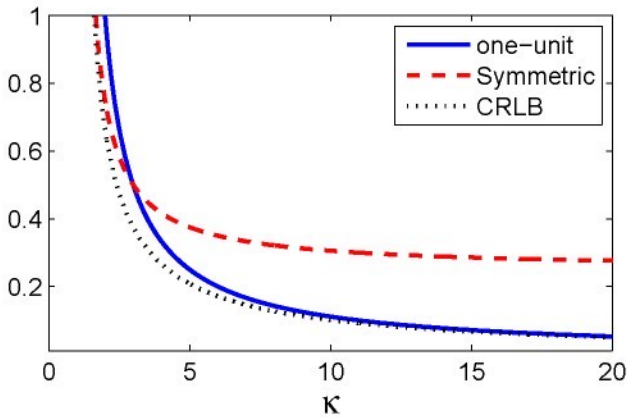


Figure 2.2: A comparison of the best theoretical accuracy of one-unit FastICA, Symmetric FastICA and the corresponding CRLB given by expressions, respectively, (2.54), (2.55) and (2.56). The expressions are plotted as functions of $\kappa \geq 1$ with $N = 1$.

entropies of signals like in FastICA. The main difference is that the optimization proceeds by a different iterative algorithm. There are two nonlinearities $g(x) = x \pm \tanh(x)$ approximating the score functions. They differ in the sign before the term $\tanh(x)$, where one is for sub-Gaussian and the other for super-Gaussian sources. The optimization rule switches the sign according to a stability criterion.

JADE

Joint cumulants of random variables of higher order than two provide higher-order statistics with appealing properties. For example, a cumulant of a set of random variables is equal to zero if the random variables can be divided in two independent non-empty sets. The cumulants are multilinear, which allows to derive such cumulant matrices that have analogous property like the correlation matrices in (2.15), which means that the separation can be based on the joint diagonalization of these matrices.

JADE was proposed in [13]. It is an algorithm based on the joint diagonalization of cumulant matrices involving all the cumulants of order two and four. The joint diagonalization algorithm utilizes Jacobi rotations and the orthogonal constraint. The residual ISR achieved by JADE was shown to be given by [38]

$$\mathrm{ISR}_{ij}^{\mathrm{JADE}} \approx \frac{1}{N} \frac{k_j^4 + l_i k_i^2 + l_j k_j^2}{(k_i^2 + k_j^2)^2}, \tag{2.57}$$

where

$$k_i = \mathrm{E}[s_i^4] - 3$$
$$l_i = \mathrm{E}[s_i^6] - \mathrm{E}^2[s_i^4].$$

Note that k_i is the kurtosis of the ith original source, which is equal to zero if the ith distribution is Gaussian. If all signals have the same distribution, (2.57) is lower bounded by $1/(4N)$, which is in accordance with the fact that JADE applies the orthogonal constraint (see (2.55) also).

An optimized variant of JADE, called OFORIA, was proposed in [37]. It relies on a non-orthogonal weighted joint approximate diagonalization of the cumulant matrices. The weights are optimized (the weights correspond to elements of the inverse covariance matrix of estimates of the cumulant matrices), and the resulting performance of OFORIA is

$$\mathrm{ISR}_{ij}^{\mathrm{OFORIA}} \approx \frac{1}{N} \frac{l_i l_j + l_i k_j^2}{l_i k_j^2 + l_j k_i^2 + k_i^2 k_j^2}. \tag{2.58}$$

Non-Parametric and Parametric Algorithms

A well known fact following from various theories and analyses is that the most accurate separation can be achieved if correct models of signal PDFs are used. For example, the best nonlinearity g that can be used in FastICA is the score function of the PDF of signals (assuming that all signals have the same distribution, since FastICA applies one nonlinearity only). Therefore, some methods aim at estimating the PDFs or score functions either in a parametric or non-parametric way.

Examples of non-parametric algorithms are NPICA [17] or RADICAL [34]. While NPICA estimates PDF using a Gaussian mixture model, RADICAL utilizes order statistics. The advantage of non-parametric methods is their flexibility, since they can separate signals of arbitrary non-Gaussian distribution. This is in contrast to parametric approaches, since for any fixed nonlinearity, it is quite easy to find a non-Gaussian distribution which FastICA (and likely other algorithms), endowed by the nonlinearity, fails to separate [c8].

Although the non-parametric separation methods are usually accurate, they are computationally very complex and cannot be used to separate more than few (10-20) signals, in practice.

Some other separation methods use a parametric modeling of the score functions of the separated signals. For instance, Pham et al. proposed mean square fitting of the score functions by linear combinations of fixed nonlinear functions in [18], to derive a blind separating algorithm. Another parametric estimator coming from the maximum entropy principle was proposed in [19].

EFICA

The analysis of FastICA gave rise to a new, more sophisticated, algorithm named EFICA [j3]. EFICA is initialized by the outcome of Symmetric FastICA. Then, a special technique called a *test of saddle points* is applied to make sure that the global minimum of the contrast function was found. The partly separated signals are used to select optimal nonlinearities g_i, $i = 1, \ldots, d$, for each separated signal, and used in a fine tuning of rows of $\mathbf{W}$. Finally, the whole $\mathbf{W}$ is refined from the version with fine-tuned rows. The refinement entails the avoidance of the orthogonal constraint. It is done in an optimum way so that the variance of the estimate of $\mathbf{W}$ is minimized.

Theoretical analysis based on [j2] shows that the residual ISR due to the separation by EFICA is

$$\mathrm{ISR}_{ij}^{EF} \approx \frac{1}{N} \frac{\gamma_i(\gamma_j + \tau_j^2)}{\tau_j^2 \gamma_i + \tau_i^2(\gamma_j + \tau_j^2)}. \tag{2.59}$$

If the nonlinearities g_i, $i = 1, \ldots, d$, match the score functions of the signals, the right side of (2.59) becomes equal to the CRLB (2.28), which means that EFICA is asymptotically efficient in the special case. In this respect, the adaptive (parametric) choice of the nonlinearities was designed to work efficiently with signals having a generalized Gaussian distribution[3].

On principle, EFICA does not differ much from FastICA in terms of computational complexity, so it retains its popular property, which is high speed. On the other hand, it outperforms FastICA in terms of accuracy and global convergence (stability), which was demonstrated by various experiments even with real-world signals. Some further improvements of EFICA in terms of speed and accuracy were proposed in [c8] and [c11].

1FICA

EFICA is an optimal estimator of the separating matrix in terms of the estimation variance when the mixed signals do not contain any noise. However, if the noise is present, the estimate by EFICA is biased and need not be optimal, e.g., in terms of SINR. The bias was discussed in Section 2.4.3. In [m6], the asymptotic bias of EFICA and Symmetric FastICA was derived, and it was shown that both the algorithms have bias of the asymptotic order $\mathcal{O}(\sigma^2)$, where $\sigma^2\mathbf{I}$ is the covariance matrix of the noise in (1.3). On the other hand, the bias of one-unit FastICA was shown to have the order $\mathcal{O}(\sigma^3)$, at least.

[3] PDF of the generalized Gaussian distribution is proportional to $\exp(-|x|^\alpha/\beta)$, where α is a shape parameter and β controls the variance. The distribution include standard normal distribution for $\alpha = 2$, Laplacean distribution for $\alpha = 1$ and a uniform distribution as a limit for $\alpha \to \infty$.

However, one-unit FastICA estimates one original signal only, and it depends on its initialization which of the d signals is being estimated. The 1FICA algorithm was designed to have the same performance as the one-unit algorithm, but the global convergence to all original signals is ensured. Under mild conditions, it follows that the bias of 1FICA has the same asymptotic order as that of one-unit FastICA.

The bias is one aspect of the quality of an estimator; the other one is its variance. When the noise is present, both should be taken into account to select the best algorithm for separation. In the noise-free case, it was shown that the variance is small when the distribution of signals is highly non-Gaussian. Such property have, for instance, finite-alphabet signals used in wireless communications. 1FICA was shown to separate these signals disturbed by noise better than the other FastICA-based variants (including EFICA) for large enough $\sigma^2 > 0$. This is thanks to the lower bias, because the variance is small and less important is such cases.

The 1FICA algorithm and its version for complex-valued signals was proposed in paper [c9], whose reprint is included as a part of this thesis.

2.5.2 Methods Based on Approximate Joint Diagonalization of (Cross-)Covariance Matrices

As explained in Sections 2.3.2, 2.3.3, and 2.3.4, the utilization of nonstationarity and/or spectral diversity of signals leads to an JAD of covariance and cross-covariance matrices computed either over the whole data or their blocks. In general, the methods differ in the criterion of the joint diagonality and in a way to optimize it. While the criterion has mainly the influence on statistical performance of the algorithm, the optimization approach affects its speed and stability. An important role plays the constraint that restrains the algorithm to converge to the trivial solution $\mathbf{W} = 0$.

Several algorithms for the JAD have been proposed and applied to derive ICA methods; a survey of JAD approaches is given, e.g., in [28]. Earlier papers proposing methods based on the nonstationarity are, for instance, [33] and [32]. These methods apply the orthogonal constraint, especially, due to the computational and implementation simplicity. More recent methods apply different constraints to improve the accuracy of separation, e.g. [30] and [31]. The methods in [28] incorporate weights into the criterion to optimize statistical properties of the resulting joint diagonalizer for given models of signals.

BGSEP

In [28], a method for JAD was proposed, which is asymptotically equivalent to the Pham's estimator from [29], but is more appealing computationally. It bears the name BGWEDGE (Block Gaussian Weighted Exhaustive Diagonalization with Gauss itErations), and the corresponding separation algorithm is called BGSEP (Block Gaussian separation). The performance of BGSEP was shown to attain the CRLB (2.29) when signals obey the given statistical model.

Although theoretical computational complexity of Pham's algorithm and BGWEDGE is the same, $O(d^2M)$ operations per iteration, the latter algorithm is easier to parallelize. In matlab implementation, BGWEDGE is realized with fewer embedded "for" cycles, and, therefore, it is faster in higher dimensions.

WASOBI

One of the first algorithms assuming the spectral diversity of signals was AMUSE proposed in [27]. It is based on the diagonalization of one cross-correlation matrix $\widehat{\mathbf{R}}[\tau]$ ($\tau > 0$) under the orthogonal constraint. Later, a popular algorithm doing the JAD of a set of matrices based on Jacobi rotations was proposed in [12] and is known under the acronym SOBI (Second Order Blind Identification). The algorithm became quite popular in biomedical applications.

SOBI, however, is not statistically efficient, if the original signals obey the assumed AR model. Statistically efficient estimators of the mixing or separating matrix attaining the CRLB (2.34) were independently proposed by Pham, Degerine and Zaidi, and Tichavský and Yeredor. The latter algorithm is called WASOBI (weight adjusted SOBI) [26]. The weights in WASOBI are derived from AR modeling of partially separated signals. WASOBI was shown to allow an approximately efficient separation even in high (100+) dimensional datasets.

BARBI

The abbreviation BARBI [c16] stands for the Block AutoRegressive Blind Identification. It is a hybrid method that relies both on the nonstationarity and spectral diversity. Like BGSEP, the method assumes that the mixture can be partitioned into M blocks, and in each of them, the separated signals are Gaussian stationary and autoregressive of order $R-1$. Therefore, it can be viewed as an extension of BGSEP and WASOBI.

Similarly to the SONS algorithm from [25], it does an JAD of the cross-covariance matrices computed at each block separately. The number of these matrices is $M \times R$. Unlike other ICA algorithms that are based on an JAD of some matrices, the JAD in BARBI incorporates a data dependent weighting, which reflects the statistical model of the separated data. Therefore BARBI outperforms

other separation methods in terms of accuracy, if the assumed model is in accordance with the reality. Then it achieves the CRLB bound given by (2.30). The JAD method is similar to that in BGSEP and WASOBI, so BARBI is computationally effective and allows separation of high-dimensional data.

2.5.3 Block EFICA: A Method Combining the Nonstationarity and Non-Gaussianity

Block EFICA is an ICA/BSS algorithm that relies both on the non-Gaussianity and nonstationarity [j6]. Like the BGSEP algorithm, Block EFICA assumes that the separated signal can be partitioned in a set of M blocks, so that the signals are stationary in each block. The signals may have different variances and even different distributions on distinct blocks. This model of data is also considered by Pham in [24], who suggests an asymptotically efficient (CRLB attaining) estimator based on the maximum likelihood approach. The latter algorithm, however, appeared not to be as stable as Block EFICA, and its performance is problematic in difficult scenarios and in high dimensions.

The concept of Block EFICA is similar to that of EFICA. The main difference consists in that the optimal nonlinearities approximating score functions are estimated separately in each block of signals. The Pham's parametric estimator from [18] is used for adaptive selection of the best linear combination of the functions from [c8]. The second main difference is that the optimum weights for the refinement of the final estimate of $\mathbf{W}$ are computed accordingly, respecting the piecewise stationary model.

Block EFICA asymptotically approaches CRLB (2.24) under common assumptions when *variance of the signals is constant over blocks*, which is expressed by

$$\mathrm{ISR}_{ij}^{\mathrm{BEF}} \approx \frac{1}{N} \frac{\overline{\kappa}_j}{\overline{\kappa}_i \overline{\kappa}_j - 1}, \qquad i \neq j, \tag{2.60}$$

where $\overline{\kappa}_i = \frac{1}{M} \sum_{\ell=1}^{M} \kappa_i^{(\ell)}$. In case that the variance of signals is changing, the algorithm is not optimal in theory, but its performance is close to the CRLB (2.24) in practice. This was demonstrated by experiments with both synthetic and real-world signals.

The reprint of [j6] is included as a part of this thesis.

2.5.4 Towards Combining the Non-Gaussianity and Spectral Diversity: the MULTICOMBI algorithm

MULTICOMBI is an algorithm that combines EFICA and WASOBI to separate mixtures of signals that are either non-Gaussian or can be well resolved in the

spectral domain. It is based on the fact that these algorithms allow to estimate not only the separating matrix, but also the separation performance measured in terms of the estimated ISR matrix.

The ISR matrix is estimated by examining the statistical properties of the separated signals using the formulas (2.59) for EFICA and (2.34) for WASOBI. For instance, if some separated component is highly non-Gaussian, ISR of EFICA with respect to other components will be low, and vice versa: If there is a group of components that have nearly Gaussian distribution and cannot be well separated one from the other, the corresponding ISR submatrix will have large entries. Similarly WASOBI produces an estimated ISR matrix which reveals the structure of the mixture, i.e. which components have mutually similar spectra (and therefore they are hard to separate one from the other) and vice versa.

MULTICOMBI applies both algorithms to the input data, which gives two different sets of independent components. The clusters of components are identified separately for each algorithm by a clustering method in which a similarity of clusters is defined through the estimated ISR matrices. MULTICOMBI then accepts the clusters of one algorithm only such that yield better estimated ISR than the best cluster of the other algorithm. The other clusters of the first algorithm are accepted as one merged cluster, unless it is empty. The procedure is applied recursively to each non-singleton cluster until all clusters are singletons, i.e., contain one component only and provide the output of MULTICOMBI.

In simulations, MULTICOMBI was shown to outperform other existing methods that rely on non-Gaussianity and spectral diversity, for instance, JADE_{TD} [21], JCC [20] and ThinICA [22]. These methods are mostly based on a joint approximate diagonalization of either cross-covariance, cumulant and cross-cumulant matrices. The (cross-)cumulants represent higher-order statistics taking the non-Gaussianity into account. Neither of these methods optimize the separation criterion to achieve the statistical efficiency given by the combined model.

The reprint of the paper [j5] where MULTICOMBI was proposed is included as a part of this thesis.

Chapter 3

Independent Subspace Analysis

3.1 Definition

One possible extension of the ICA is the so-called *Independent Subspace Analysis* (ISA). The goal here is to decompose the signal space spanned by rows of $\mathbf{X}$ to a direct sum of independent subspaces. It means that an element of a given subspace should be statistically independent of elements of the other subspaces but need not be independent of elements within the subspace. Due to proper definition of the solution (uniqueness) it is required that no independent subspace can be further decomposed into independent subspaces (maximal decomposition). The exception could be a subspace of Gaussian components that can always be made uncorrelated, thus, independent.

ICA can be seen as a special case of ISA where the dimension of each subspace is equal to one. Therefore, ISA is also known under the name *Multidimensional ICA* [40].

ISA can be applied in situations where not all original signals can be separated from each other. In many cases, the ISA problem can be approached by applying an ICA algorithm, because the algorithm aims at finding components that are mutually independent as much as possible. This is followed by a clustering of the obtained components according to some similarity measure that reflects, e.g., their mutual dependence. The identified clusters of components then should correspond to the desired subspaces.

The following section considers several important examples where the ISA problem arises.

3.2 Examples

3.2.1 Invisible Independent Sources

Some signals cannot be separated by a given ICA algorithm in spite of being independent, but can be separated by another algorithm. This occurs when not all signals meet the separability conditions of the given model, in particular, when

1. Non-Gaussianity-based model: there are signals whose marginal distributions are Gaussian,

2. Nonstationarity-based model: there are signals having the same variance profiles, or

3. Spectral diversity-based model: some signals have the same spectra.

A key question is how the given ICA algorithm behaves when applied to a mixture containing such signals. It is desired that the algorithm separates the subspace of undistinguishable signals from the other signals. Fortunately, the answer whether this is fulfiled is positive in most situations.

A theoretical justification is provided by the Cramér-Rao bounds introduced in Section 2.4.1. Under the given models, the attainable residual ISR between two separated signals was shown to be independent of properties of the other signals. The behavior of algorithms is expected to be the same, which is usually corroborated experimentally. The analyses of the FastICA-based methods, e.g. given by (2.47), (2.48) or (2.59), or of JADE in (2.57), have confirmed the property in theory. The MULTICOMBI algorithm described in Section 2.5.4 relies on this property of the EFICA and WASOBI algorithms.

3.2.2 Dependent Sources

An ISA problem is defined if some of the original signals are assumed to be dependent. This is motivated by some natural signals such as recordings from electrocardiogram (ECG) from a pregnant woman, where the goal is to separate the maternal signal and the fetal signal [41]. It is believed that each individual ECG signal consists of several dependent components. Since the maternal and fetal signals should be independent, they form two independent subspaces in the mixed data.

General ISA algorithms aim at restoring the independence of subspaces while they leave out the dependence of signals within the subspaces. For example, this can be done by doing an approximate joint-block diagonalization of a set of covariance matrices [42]. It differs from the JAD task in that the covariance matrices

of separated signals need not be diagonal, but it suffices if they are block-diagonal. This means that some signals are allowed to be correlated (dependent) after the separation. The drawback is that the dimensions of subspaces (blocks) must be known or estimated. Furthermore, the criterion of block-diagonality might have local minima.

As pointed early in this chapter, the other approach is to separate the mixture into components by an ICA algorithm and perform a clustering of the components. To solve the ISA problem this way, it is necessary that components containing the dependent signals are separated from the independent ones. Since ICA algorithms aim at restoring the independence of all signals, it is highly expectable that the condition is fulfilled. The only redundant effect should be that the dependent signals are mutually transformed to be "as independent as possible". Some practical experiments confirming this feature of some ICA methods were done, e.g., in [c15].

The paper [c15] is included in the second part of this thesis.

3.2.3 Underdetermined and Noisy Mixtures

The underdetermined (more signals than sensors) or noisy mixtures can sometimes be regarded as such where the original signals are "slightly" dependent. Consider a case where each signal of a regular mixture of d signals is contaminated by an unwanted signal $s_{d+1}(n)$. The mixed signals can be written as

$$\mathbf{X} = \mathbf{AS} + \mathbf{b}\mathbf{s}_{d+1} \tag{3.1}$$

where $\mathbf{s}_{d+1}$ denotes a $1 \times N$ vector containing N samples of $s_{d+1}(n)$, and $\mathbf{b}$ is a $d \times 1$ vector of coefficients determining the level of contamination on channels. Assume that $\mathbf{b}$ is "small enough" and $s_{d+1}(n)$ is not a target signal to be separated (not a further original signal).

To separate the original signals $\mathbf{S}$ each from the other, the mixed signals should be multiplied by $\mathbf{A}^{-1}$ (up to the indeterminacies), and the resulting signals are

$$\widehat{\mathbf{S}} = \mathbf{A}^{-1}\mathbf{X} = \mathbf{S} + \mathbf{A}^{-1}\mathbf{b}\mathbf{s}_{d+1}. \tag{3.2}$$

It is seen that these signals are dependent, which is caused by the common additive term $\mathbf{A}^{-1}\mathbf{b}\mathbf{s}_{d+1}$. As discussed in Section 2.4.3, a contamination of mixture by additive noise (of arbitrary distribution) causes bias in the mixing matrix estimation provided that $\mathbf{b}$ is small enough. Consequently, the example here points to the behavior of ICA algorithms when applied to slightly dependent signals: They are biased but estimate the target signals correctly.

3.3 Pseudo-Convolutive Mixtures

A pseudo-convolutive mixture of signals is the instantaneous mixture $\mathbf{X} = \mathbf{AS}$ where the matrix of original signals $\mathbf{S}$ is defined in a special way as

$$\widetilde{\mathbf{S}} = \begin{bmatrix} s_1(N_1) & \dots & \dots & s_1(N_2) \\ s_1(N_1-1) & \dots & \dots & s_1(N_2-1) \\ \vdots & \vdots & \vdots & \vdots \\ s_1(N_1-L+1) & \dots & \dots & s_1(N_2-L+1) \\ s_2(N_1) & \dots & \dots & s_2(N_2) \\ s_2(N_1-1) & \dots & \dots & s_2(N_2-1) \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ s_d(N_1-L+1) & \dots & \dots & s_d(N_2-L+1) \end{bmatrix}, \tag{3.3}$$

where N_1 and N_2, $N_2 > N_1$, determine the block of samples of original signals $s_1(n), \dots, s_d(n)$, and L is a free integer parameter.

The point here is that the rows of $\widetilde{\mathbf{S}}$ corresponding to the same original signal are dependent unless each original signal is a sequence of independent samples (for example, an i.i.d. signal). Nevertheless, $s_1(n), \dots, s_d(n)$ are still assumed to be independent, so the blocks of $\widetilde{\mathbf{S}}$ form L-dimensional independent subspaces, and the separation of the instantaneous mixture of $\widetilde{\mathbf{S}}$ poses an ISA problem.

The mixture is called "pseudo-convolutive", because the matrix $\mathbf{X} = \mathbf{A}\widetilde{\mathbf{S}}$ has several features that are common to a matrix defined through

$$\mathbf{\Xi} = \begin{bmatrix} x_1(N_1) & \dots & \dots & x_1(N_2) \\ x_1(N_1-1) & \dots & \dots & x_1(N_2-1) \\ \vdots & \vdots & \vdots & \vdots \\ x_1(N_1-L+1) & \dots & \dots & x_1(N_2-L+1) \\ x_2(N_1) & \dots & \dots & x_2(N_2) \\ x_2(N_1-1) & \dots & \dots & x_2(N_2-1) \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ x_d(N_1-L+1) & \dots & \dots & x_d(N_2-L+1) \end{bmatrix}, \tag{3.4}$$

when $x_1(n), \dots, x_d(n)$ are convolutive mixtures of $s_1(n), \dots, s_d(n)$, i.e., they are mixed according to the model (1.2)[1]. In fact, the signals $\mathbf{X} = \mathbf{A}\widetilde{\mathbf{S}}$ corresponds

[1] The matrix defined through (3.4) is considered in Section 6.2 on time-domain methods for separation of convolutive mixtures of audio signals.

to dL outputs of MIMO filters of the length L whose inputs are $s_1(n), \ldots, s_d(n)$. Nevertheless, note that $\boldsymbol{\Xi}$ cannot be decomposed in the form $\boldsymbol{\Xi} = \mathbf{A}\widetilde{\mathbf{S}}$ where $\mathbf{A}$ is square. In other words, the matrices $\mathbf{A}\widetilde{\mathbf{S}}$ and $\boldsymbol{\Xi}$ are *not* equal, in general. A special case when a convolutive mixture matrix $\boldsymbol{\Xi}$ can be rearranged into an instantaneous (overdetermined) mixing problem is described, e.g., in [57].

The pseudo-convolutive mixtures can be used for testing and comparing ICA and ISA algorithms. It is the subject of article [c15], where the original signals are audio signals (speech) that have significant temporal structure. Hence, the signals are dependent within the subspaces but the subspaces are independent, which means an ISA problem. The conclusions of [c15] are important for the selection of appropriate ICA algorithm that is applied within the method described in Section 6.2.1.

The article [c15] is included as a part of this thesis.

Reprints

- Z. Koldovský and P. Tichavský, "Blind Instantaneous Noisy Mixture Separation with Best Interference-plus-noise Rejection," *Proceedings of 7th International Conference on Independent Component Analysis (ICA2007)*, Lecture Notes in Computer Science, vol. 4666, pp. 730-737, Springer Berlin / Heidelberg, Sept. 2007.

- P. Tichavský, Z. Koldovský, A. Yeredor, G. G. Herrero, E. Doron, "A Hybrid Technique for Blind Non-Gaussian and Time-Correlated Sources Using a Multicomponent Approach, " *IEEE Trans. on Neural Networks*, 19(3), 421–430, March 2008.

- Z. Koldovský, J. Málek, P. Tichavský, Y. Deville, and S. Hosseini, "Blind Separation of Piecewise Stationary NonGaussian Sources", *Signal Processing*, Volume 89, Issue 12, Pages 2570-2584, ISSN 0165-1684, December 2009.

- Z. Koldovský and P. Tichavský, "A Comparison of Independent Component and Independent Subspace Analysis Algorithms, " *Proc. of EUSIPCO 2009*, pp. 1447-1451, Glasgow, August 24-28, 2009.

Blind Instantaneous Noisy Mixture Separation with Best Interference-Plus-Noise Rejection*

Zbyněk Koldovský[1,2] and Petr Tichavský[1]

[1] Institute of Information Theory and Automation, Pod vodárenskou věží 4, P.O. Box 18, 182 08 Praha 8, Czech Republic
zbynek.koldovsky@tul.cz
[2] Faculty of Mechatronic and Interdisciplinary Studies
Technical University of Liberec, Hálkova 6, 461 17 Liberec, Czech Republic

Abstract. In this paper, a variant of the well known algorithm FastICA is proposed to be used for blind source separation in off-line (block processing) setup and a noisy environment. The algorithm combines a symmetric FastICA with test of saddle points to achieve fast global convergence and a one-unit refinement to obtain high noise rejection ability. A novel test of saddle points is designed for separation of complex-valued signals. The bias of the proposed algorithm due to additive noise can be shown to be asymptotically proportional to σ^3 for small σ, where σ^2 is the variance of the additive noise. Since the bias of the other methods (namely the bias of all methods using the orthogonality constraint, and even of recently proposed algorithm EFICA) is asymptotically proportional to σ^2, the proposed method has usually a lower bias, and consequently it exhibits a lower symbol-error rate, when applied to blind separation of finite alphabet signals, typical for communication systems.

1 Introduction

The noisy model of Independent Component Analysis (ICA) considered in this paper, is

$$\mathbf{X} = \mathbf{AS} + \sigma\mathbf{N}, \tag{1}$$

where $\mathbf{S}$ denotes a vector of d independent random variables representing the original signals, $\mathbf{A}$ is an unknown regular $d \times d$ mixing matrix, and $\mathbf{X}$ represents the observed mixed signals. The noise $\mathbf{N}$ denotes a vector of independent variables having the covariance matrix $\mathbf{\Sigma}$. Without loss of generality, we will further assume that $\mathbf{\Sigma}$ equals to the identity matrix $\mathbf{I}$. Consequently, σ^2 is the variance of the added noise to the mixed signals. All signals considered here are i.i.d. sequences, i.e., they are assumed to be white in the analysis.

It is characteristic for most ICA methods that they were derived for the noiseless case, so to solve the task of estimating the mixing matrix $\mathbf{A}$ or its inversion

* This work was supported by Ministry of Education, Youth and Sports of the Czech Republic through the project 1M0572 and through the Grant 102/07/P384 of the Grant Agency of the Czech Republic.

M.E. Davies et al. (Eds.): ICA 2007, LNCS 4666, pp. 730–737, 2007.

$\mathbf{W} = \mathbf{A}^{-1}$. Then, abilities to separate noised data are studied experimentally, and the non-vanishing estimation error as $N \to +\infty$, N being length of data, is taken for a bias caused by the noise. To compensate such bias, several techniques were proposed [6]. Unfortunately, these methods have a drawback that the covariance structure of the noise needs to be known *a priori* [4].

In accord with [7], we suggest to measure the separation quality not through accuracy of estimation of the mixing mechanism but through the achieved interference + noise to signal ratio (INSR) or its inverse SINR. In separating the finite alphabet signals, the ultimate criterion should be the symbol error rate (SER). Computation of the INSR or the SER assumes that the permutation, scale, and sign or phase ambiguities were resolved by minimizing the INSR.

In the case $\mathbf{\Sigma} = \mathbf{I}$, the INSR of a k-th estimated signal can be computed as

$$\mathrm{INSR}_k = \frac{\sum_{i \neq k}^{d} (\mathbf{BA})_{ki}^2 + \sigma^2 \sum_{i=1}^{d} \mathbf{B}_{ki}^2}{(\mathbf{BA})_{kk}^2}, \tag{2}$$

where $\mathbf{B}$ is the separating transformation [7]. The solutions that minimize (2) are known to be given by the MMSE separating matrix, denoted by $\mathbf{W}^{\mathrm{MMSE}}$, that takes the form

$$\mathbf{W}^{\mathrm{MMSE}} = \mathbf{A}^H (\mathbf{A}\mathbf{A}^H + \sigma^2 \mathbf{I})^{-1} \tag{3}$$

where ^{H} denotes the conjugate (Hermitian) transpose. Signals given by $\mathbf{W}^{\mathrm{MMSE}}\mathbf{X}$ will be further called the *MMSE solution*. Note that these signals may not be necessarily normalized to have unit variance, unlike outcome of common blind separation methods, that produce normalized components. For exact comparisons, we introduce a matrix $\mathbf{W}^{\mathrm{NMMSE}}$ such that $\mathbf{W}^{\mathrm{NMMSE}}\mathbf{X}$ are the normalized MMSE signals.

The paper is organized as follows. In section 2, we briefly describe several variants of algorithm FastICA and the proposed method, including a novel test of saddle points for separating complex-valued signals. Section 3 presents analytic expressions for an asymptotic bias of solutions obtained by real domain FastICA variants [5,8] from the MMSE solution. Specifically, we study the biases of estimates of de-mixing matrix $\mathbf{W}$ from $\mathbf{W}^{\mathrm{NMMSE}}$, and the one-unit FastICA and the proposed algorithm is shown to be less biased than the other methods. Simulations in Section 4 demonstrate drawbacks of the unbiased algorithm [6] (further referred to as *unbiased FastICA*) following from required knowledge of $\mathbf{\Sigma}$ and/or σ. Conversely, the proposed algorithm with one-unit FastICA-like performance is shown to be the best blind MMSE estimator when separating noisy finite-alphabet signals.

2 FastICA and Its Variants

Common FastICA algorithms work with the decorrelated data $\mathbf{Z} = \mathbf{C}^{-1/2}\mathbf{X}$, where $\mathbf{C} = \mathrm{E}[\mathbf{X}\mathbf{X}^H]$ is the data covariance matrix. Only the unbiased FastICA [6] that aims at unbiased estimation of $\mathbf{A}^{-1}$ assuming that the noise has a known covariance matrix $\mathbf{\Sigma}$, uses the preprocessing $\mathbf{Z} = (\mathbf{C} - \mathbf{\Sigma})^{-1/2}\mathbf{X}$.

One-unit FastICA in real domain [5] estimates one de-mixing vector $\mathbf{w}_k^{1\mathrm{U}}$ iteratively via the recursion

$$\mathbf{w}_k^+ \leftarrow \mathrm{E}[\mathbf{Z}g(\mathbf{w}_k^{1\mathrm{U}T}\mathbf{Z})]\} - \mathbf{w}_k^{1\mathrm{U}}\mathrm{E}\{g'(\mathbf{w}_k^{1\mathrm{U}T}\mathbf{Z}))\}, \qquad \mathbf{w}_k^{1\mathrm{U}} \leftarrow \mathbf{w}_k^+/\|\mathbf{w}_k^+\| \quad (4)$$

until convergence is achieved. Here $g(\cdot)$ is a smooth nonlinear function that approximates/surrogates the score function corresponding to the distribution of the original signals [11]. The theoretical expectation values in (4) are, in practice, replaced by their sample-based counterparts.

Similar recursion was proposed for one-unit FastICA in the complex domain [1]. The symmetric (real or complex) variant performs the one-unit iterations in parallel for all d separating vectors, but the normalization in (4) is replaced by a symmetric orthogonalization.

The algorithm EFICA [8] combines the symmetric approach with the test of saddle points, an adaptive choice of nonlinearity $g_k(\cdot)$ for each signal separately, and it does the refinement step that relaxes the orthogonal constraint introduced by the symmetric approach and is designed towards asymptotic efficiency.

The unbiased FastICA [6] uses the recursion

$$\mathbf{w}_k^+ \leftarrow \mathrm{E}[\mathbf{Z}g(\mathbf{w}_k^{\mathrm{unb}T}\mathbf{Z})] - (\mathbf{I} + \widetilde{\mathbf{\Sigma}})\mathbf{w}_k^{\mathrm{unb}}\,\mathrm{E}[g'(\mathbf{w}_k^{\mathrm{unb}T}\mathbf{Z})],$$

where $\widetilde{\mathbf{\Sigma}} = (\mathbf{C}-\mathbf{\Sigma})^{-1/2}\mathbf{\Sigma}(\mathbf{C}-\mathbf{\Sigma})^{-1/2}$. Both approaches (one-unit and symmetric) can be considered; in simulations, we use the one-unit variant, and the resulting de-mixing matrix will be denoted by $\mathbf{W}^{\mathrm{UNB}}$. In order to compare performance of the unbiased FastICA by means of (2) with the other techniques fairly, it is necessary to consider a MMSE estimate derived from $\mathbf{W}^{\mathrm{UNB}}$, namely

$$\mathbf{W}^{\text{MMSE-UNB}} = \mathbf{\Sigma}^{-1}(\mathbf{W}^{\mathrm{UNB}})^{-T} \times [(\mathbf{W}^{\mathrm{UNB}})^{-1}\mathbf{\Sigma}^{-1}(\mathbf{W}^{\mathrm{UNB}})^{-T} + \sigma^2\mathbf{I}]^{-1} \quad (5)$$

2.1 Proposed Algorithm

The proposed algorithm is a combination of symmetric FastICA, test of saddle points, and one-unit FastICA as a refinement. Usually, one unit FastICA is used in a deflation way, when the estimated components are subtracted from the mixture one by one. This is computationally effective method, but accuracy of the later separated components might be compromised. Therefore, we propose to initialize the algorithm using symmetric FastICA, that is known for having very good global convergence and allows equal separation precision for all components.

The test of saddle points was first proposed in [11] to improve probability of the symmetric FastICA to converge to the true global maximum of the cost function $[\mathrm{E}\{G(\mathbf{w}^T\mathbf{Z})\} - G_0]^2$ where $G(\cdot)$. is a primitive function of $g(\cdot)$ and $G_0 = \mathrm{E}\{G(\xi)\}$, where ξ is a standard Gaussian random variable.

In short, the test of saddle points consists in checking all pairs of the estimated components $(\mathbf{u}_k, \mathbf{u}_\ell)$, whether or not other pair of signals $(\mathbf{u}'_k, \mathbf{u}'_\ell)$ gives a higher value of the cost function

$$c(\mathbf{u}_k, \mathbf{u}_\ell) = [\mathrm{E}\{G(\mathbf{u}_k)\} - G_0]^2 + [\mathrm{E}\{G(\mathbf{u}_\ell)\} - G_0]^2, \quad (6)$$

where $\mathbf{u}'_k = (\mathbf{u}_k + \mathbf{u}_\ell)/\sqrt{2}$ and $\mathbf{u}'_\ell = (\mathbf{u}_k - \mathbf{u}_\ell)/\sqrt{2}$.

The motivation is that a random initialization of the algorithm may begin at a point of zero gradient of the cost function (a saddle point / an unstable point of the iteration) and terminate there, despite being not the desired stable solution. See [11] for details.

In the complex domain, the situation is a bit more tricky, because if $(\mathbf{u}_k, \mathbf{u}_\ell)$ is the pair of valid independent components in the mixture, not only their weighted sum and a difference represent a false (unstable) point of the iteration. *All* pairs $(\mathbf{u}'_k, \mathbf{u}'_\ell)$ of the form $\mathbf{u}'_k = (\mathbf{u}_k + e^{i\alpha}\mathbf{u}_\ell)/\sqrt{2}$ and $\mathbf{u}'_\ell = (\mathbf{u}_k - e^{i\alpha}\mathbf{u}_\ell)/\sqrt{2}$ are stationary but unstable for any phase factor $e^{i\alpha}$, $\alpha \in \mathcal{R}$.

Therefore we propose to do a phase shift of each separated component so that the real part and the imaginary part of the signal are as much independent each of other as possible before the test of the saddle points. This phase shift can be easily performed using a two-dimensional symmetric FastICA in the real domain applied to the real and imaginary part of the component. After this preprocessing, it is sufficient to perform the test of saddle points exactly as in the real-valued case, i.e. to check all pairs $(\mathbf{u}'_k, \mathbf{u}'_\ell)$ with $\mathbf{u}'_k = (\mathbf{u}_k + \mathbf{u}_\ell)/\sqrt{2}$ and $\mathbf{u}'_\ell = (\mathbf{u}_k - \mathbf{u}_\ell)/\sqrt{2}$, whether they give a higher value of the cost function (6) or not.

Validity of the above described complex domain test of the saddle points can be easily confirmed in simulations by starting the algorithm from the pairs $\mathbf{u}'_k = (\mathbf{u}_k + e^{i\alpha}\mathbf{u}_\ell)/\sqrt{2}$ and $\mathbf{u}'_\ell = (\mathbf{u}_k - e^{i\alpha}\mathbf{u}_\ell)/\sqrt{2}$ with an arbitrary $\alpha \in \mathcal{R}$ where $\mathbf{u}_k$ and $\mathbf{u}_\ell$ are the true independent sources. We have successfully tested this approach on separation of complex-valued finite alphabet sources known in communications (QAM, V27).

The resultant algorithm (symmetric FastICA + test of saddle points + one unit refinements) will be referred to as 1FICA.

3 Bias of the FastICA Variants

In this section, asymptotic expressions for bias of algorithms described in previous section working in the real domain will be presented. (The complex-domain FastICA exhibits a similar behavior in simulations.) For details of analysis, the reader is referred to [9] due to lack of space.

In brief, the theoretical analysis is done for "small" σ and infinite number of samples. Similarly to [11], for theoretical considerations, it is assumed that the analyzed method starts from the MMSE solution and stops after one iteration. This assumption is reasonable due to the following facts: (1) deviation of the global maximizer $\widehat{\mathbf{W}}$ of the FastICA cost function from $\mathbf{W}^{\mathrm{MMSE}}$ is of the order $O(\sigma^2)$, and (2) convergence of the algorithm is at least quadratic [10]. Therefore, after performing the one iteration, the deviation of the estimate from the global maximizer $\widehat{\mathbf{W}}$ is of the order $O(\sigma^4)$ and, hence, is negligible.

The bias of the algorithm will be studied in terms of the deviation of $\widehat{\mathbf{W}}(\mathbf{W}^{\mathrm{MMSE}})^{-1}$ from a diagonal matrix. More precisely, the bias is equal to the difference between $\mathrm{E}[\widehat{\mathbf{W}}](\mathbf{W}^{\mathrm{MMSE}})^{-1}$ and $\mathbf{D} = \mathbf{W}^{\mathrm{NMMSE}}[\mathbf{W}^{\mathrm{MMSE}}]^{-1}$, where $\mathbf{D}$ is the diagonal matrix that normalizes the MMSE signals $\mathbf{S}^{\mathrm{MMSE}} = \mathbf{W}^{\mathrm{MMSE}}\mathbf{X}$.

It holds that

$$\mathbf{D} = \mathbf{I} + \frac{1}{2}\sigma^2 \mathrm{diag}[\mathbf{V}_{11}, \ldots, \mathbf{V}_{dd}] + O(\sigma^3). \tag{7}$$

From here we use the notation $\mathbf{W} = \mathbf{A}^{-1}$ and $\mathbf{V} = \mathbf{W}\mathbf{W}^T$. Finally, for a matrix $\widehat{\mathbf{W}}$ that separates the data $\mathbf{S}^{\mathrm{MMSE}}$, the bias is $\mathrm{E}[\widehat{\mathbf{W}}] - \mathbf{D}$.

3.1 Bias of the One-Unit FastICA and 1FICA

It can be shown that the de-mixing vector $\mathbf{w}_k^{1\mathrm{U}}$ resulting from the one-unit FastICA (applied to the data $\mathbf{S}^{\mathrm{MMSE}}$), for $N \to +\infty$, is proportional to

$$\mathbf{w}_k^{1\mathrm{U}} = \tau_k \mathbf{e}_k + \frac{1}{2}\sigma^2 \mathbf{V}_{kk}(\tau_k + \delta_k)\mathbf{e}_k + O(\sigma^3) \tag{8}$$

where $\tau_k = \mathrm{E}[s_k g(s_k) - g'(s_k)]$, and δ_k is a scalar that depends on the distribution of s_k and on the nonlinear function g and its derivatives to the third order. Since (8) is a scalar multiple of $\mathbf{e}_k$ (the k-th column of the identity matrix), it follows that the asymptotic bias of the one-unit approach is $O(\sigma^3)$. Prospectively, the separating matrix $\mathbf{W}^{1\mathrm{F}}$ given by the proposed 1FICA has the same bias. Simulations confirm this expectation [9].

3.2 Bias of the Inversion Solution

It is interesting to compare the previous result with the solution that is given by exact inversion of the mixing matrix, i.e. $\mathbf{W}\mathbf{X} = \mathbf{S} + \sigma\mathbf{W}\mathbf{N}$; the signals will be called *the inversion solution*. From

$$\mathbf{W}(\mathbf{W}^{\mathrm{MMSE}})^{-1} = \mathbf{W}(\mathbf{A}\mathbf{A}^T + \sigma^2\mathbf{I})\mathbf{W}^T = \mathbf{I} + \sigma^2\mathbf{V}$$

it follows that the "bias" of the inversion solution is proportional to σ^2 and in general it is greater than that of 1FICA. In other words, **the algorithm 1FICA produces components that are asymptotically closer to the MMSE solution than to the inversion solution.**

3.3 Bias of Algorithms Using the Orthogonal Constraint

Large number of ICA algorithms (e.g. JADE [2], symmetric FastICA, etc.) use an orthogonal constraint, i.e., they enforce the separated components to have sample correlations equal to zero. Since the second-order statistics cannot be estimated perfectly, this constraint compromises the separation quality [3,11]. Here we show that the bias of all ICA algorithms that use the constraint has the asymptotic order $O(\sigma^2)$.

The orthogonality constraint can be written as

$$\mathrm{E}[\widehat{\mathbf{W}}\mathbf{X}(\widehat{\mathbf{W}}\mathbf{X})^T] = \widehat{\mathbf{W}}(\mathbf{A}\mathbf{A}^T + \sigma^2\mathbf{I})\widehat{\mathbf{W}}^T = \mathbf{I}\,. \tag{9}$$

It follows that the bias of all constrained algorithms is lower bounded by

$$\min_{\widehat{\mathbf{W}}(\mathbf{A}\mathbf{A}^T + \sigma^2\mathbf{I})\widehat{\mathbf{W}}^T = \mathbf{I}} \|\widehat{\mathbf{W}}(\mathbf{W}^{\mathrm{MMSE}})^{-1} - \mathbf{D}\|_F = O(\sigma^2) \tag{10}$$

where the minimization proceeds for $\widehat{\mathbf{W}}$. The matrix $\mathbf{D}$ in (10) is the same as in (7). For the minimizer $\widehat{\mathbf{W}}$ of (10) it holds that $\widehat{\mathbf{W}}(\mathbf{W}^{\mathrm{MMSE}})^{-1} = \mathbf{I}+\sigma^2\mathbf{\Gamma}+O(\sigma^3)$, where $\mathbf{\Gamma}$ is a nonzero matrix obeying $\mathbf{\Gamma} + \mathbf{\Gamma}^T = \mathbf{V}$; see [9] for details. This result can be interpreted in the way that the algorithms using the orthogonality constraint may prefer some of the separated components to give them a zero bias, but the total average bias for all components has the order $O(\sigma^2)$.

3.4 Bias of the Symmetric FastICA and EFICA

The biases of the algorithms can be expressed as

$$\mathrm{E}[\widehat{\mathbf{W}}](\mathbf{W}^{\mathrm{MMSE}})^{-1} - \mathbf{D} = \frac{1}{2}\sigma^2\mathbf{V} \odot (\mathbf{1}_{d\times d} - \mathbf{I} + \mathbf{H}) + O(\sigma^3), \qquad (11)$$

where $\mathbf{H}_{k\ell} = \frac{|\tau_\ell|-|\tau_k|}{|\tau_k|+|\tau_\ell|}$ for the symmetric FastICA, and $\mathbf{H}_{k\ell} = \frac{c_{k\ell}|\tau_\ell|-|\tau_k|}{|\tau_k|+c_{k\ell}|\tau_\ell|}$ for EFICA, where $c_{k\ell} = \frac{|\tau_\ell|\gamma_k}{|\tau_k|(\gamma_\ell+\tau_\ell^2)}$ for $k \neq \ell$ and $c_{kk} = 1$. Here, $\gamma_k = \mathrm{E}[g_k^2(s_k)] - \mathrm{E}^2[s_k g_k(s_k)]$, and g_k is the nonlinear function chosen for the k-th signal.

It can be seen that the bias of both of the algorithms has the order $O(\sigma^2)$.

4 Simulations

In this section, we present results of two experiments to demonstrate and compare the performance of the proposed algorithm 1FICA with competing methods: The symmetric FastICA (marked by SYMM), the unbiased FastICA (unbiased FICA), EFICA, and JADE [2]. Results given by "oracle" MMSE solution and the inversion solution are included as well. Examples with complex signals are not included due to lack of space.

In the first example, we separate 10 randomly mixed [7] BPSK signals with added Gaussian noise, first, for various length of data N (Fig. 1(a)) and, second, for varying input signal-to-noise ratio (SNR) defined as $1/\sigma^2$ (Fig. 1(b)). The experiment encompasses several extremal conditions: In the first scenario, where SNR=5dB ($\sigma \doteq 0.56$), N goes from 100, which is quite low for the dimension $d = 10$. The second situation examines $N = 200$ and SNR going down to 0dB.

Note that the bias may be less important than the estimation variance when the data length N is low. Therefore, in simulations, we have included two slightly changed versions of 1FICA and EFICA algorithm, denoted by "1FICA-biga" and "EFICA-biga", respectively. The modifications consist in that the used nonlinear function g is equal to the score function of marginal pdfs of the signals to-be estimated (i.e., noisy BPSK that have **bi**modal **Ga**ussian distribution, therefore, "biga" in the acronym). Adopted from the noiseless case [11], better performance of the modified algorithms may be expected.

Figure 1 shows superior performance of the proposed algorithm 1FICA and of its modified version. The same performance is achieved by the modified EFICA for $N \leq 200$, but it is lower due to the bias when N is higher. The unbiased FastICA achieves the same accuracy for $N \geq 500$ but is unstable when N is low.

The average performance of an algorithm is often spoiled due to poorer stability, which occurs in high dimensions and low N cases, mainly. In this issue, we highlight positive effect of the test of saddle points that is included in the proposed 1FICA or in EFICA. For instance, the results achieved by the symmetric FastICA would be significantly improved if the test was included in it.

The second example demonstrates conditions when the covariance of the noise is not exactly known or varying. To this end, the noise level was changed randomly from trial to trial. Five BPSK signals of the length $N = 50000$ were mixed with a random matrix and disturbed by Gaussian noise with covariance $\sigma^2\mathbf{I}$, where σ was randomly taken from interval $[0, 1]$, and then blindly separated. The mean value of the noise covariance matrix, i.e. $\mathbf{I}/3$, was used as the input parameter of the unbiased FastICA. Note that INSR and BER of this method were computed for solutions given by $\mathbf{W}^{\text{MMSE-UNB}}$ defined in (5).

The following table shows the average INSR and bit error rate (BER) that were achieved in 1000 trials. The performance of the proposed 1FICA is almost the same like that of "oracle" MMSE separator, because, here, N is very high, and the estimation error is caused by the bias only. The unbiased FastICA significantly suffers from inaccurate information about the noise intensity.

algorithm	average INSR [dB]	BER [%]
1FICA	**-5,98**	**3,19**
Symmetric FastICA	-5,68	3,55
unbiased FastICA	*6,79*	*5,25*
EFICA	-5,79	3,41
MMSE solution	**-5,98**	**3,19**
inversion solution	-4,76	4,71
JADE	-5,68	3,55

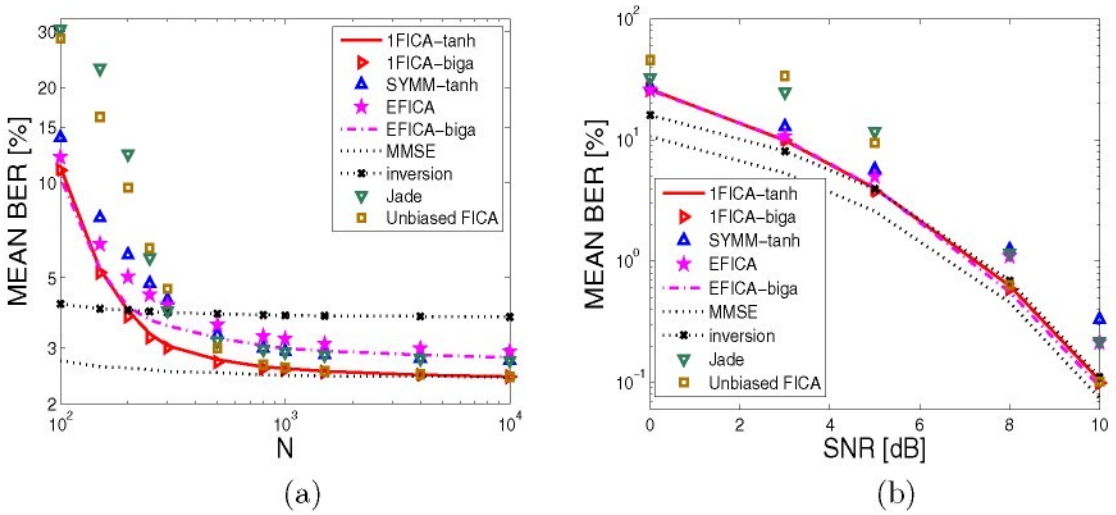


Fig. 1. Average BER of 10 separated BPSK signals when (a) SNR is fixed to 5dB and (b) a fixed number of data samples is $N = 200$. Averages are taken from 1000 independent trials for each settings.

5 Conclusions

This paper presents novel results from analysis of bias of several FastICA variants, whereby the one-unit FastICA was shown to be minimally biased from the MMSE solution, i.e., it achieves the best interference-plus-noise rejection rate for $N \to +\infty$.

By virtue of the theoretical results, a new variant of FastICA algorithm, called 1FICA, was derived to have the same global convergence as symmetric FastICA with the test of saddle points, and a noise rejection like the one-unit FastICA. Computer simulations show superior performance of the method when separating binary (BPSK) signals. Unlike the unbiased FastICA, it does not require prior knowledge of covariance of the noise to achieve the best MMSE separation.The Matlab codes for 1FICA in real and in complex domains can be downloaded from the first author's homepage, http://itakura.kes.tul.cz/zbynek/downloads.htm.

References

1. Bingham, E., Hyvärinen, A.: A fast fixed-point algorithm for independent component analysis of complex valued signals. Int. J. Neural Systems 10, 1–8 (2000)
2. Cardoso, J.-F., Souloumiac, A.: Blind Beamforming from non-Gaussian Signals. IEE Proc.-F 140, 362–370 (1993)
3. Cardoso, J.-F.: On the performance of orthogonal source separation algorithms. In: Proc. EUSIPCO, Edinburgh, pp. 776–779 (1994)
4. Davies, M.: Identifiability Issues in Noisy ICA. IEEE Signal Processing Letters 11, 470–473 (2005)
5. Hyvärinen, A., Oja, E.: A fast fixed-point algorithm for independent component analysis. Neural Computation 9, 1483–1492 (1997)
6. Hyvärinen, A.: Gaussian Moments for Noisy Independent Component Analysis. IEEE Signal Processing Letters 6, 145–147 (1999)
7. Koldovský, Z., Tichavský, P.: Methods of Fair Comparison of Performance of Linear ICA Techniques in Presence of Additive Noise. In: Proc. of ICASSP 2006, Toulouse, pp. 873–876 (2006)
8. Koldovský, Z., Tichavský, P., Oja, E.: Efficient Variant of Algorithm FastICA for Independent Component Analysis Attaining the Cramér-Rao Lower Bound. IEEE Tr. Neural Networks 17, 1265–1277 (2006)
9. Koldovský, Z., Tichavský, P.: Asymptotic analysis of bias of FastICA-based algorithms in presence of additive noise, Technical report no 2181, ÚTIA, AV ČR (2007), Available at `http://itakura.kes.tul.cz/zbynek/publications.htm`
10. Oja, E., Yuan, Z.: The FastICA algorithm revisited: Convergence analysis. IEEE Trans. on Neural Networks 17, 1370–1381 (2006)
11. Tichavský, P., Koldovský, Z., Oja, E.: Performance Analysis of the FastICA Algorithm and Cramér-Rao Bounds for Linear Independent Component Analysis. IEEE Tr. on Signal Processing 54, 1189–1203 (2006)

Signal Processing 89 (2009) 2570–2584

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Blind separation of piecewise stationary non-Gaussian sources

Zbyněk Koldovský [a,b,*], Jiří Málek [a], Petr Tichavský [b], Yannick Deville [c], Shahram Hosseini [c]

[a] Faculty of Mechatronics, Informatics and Interdisciplinary Studies, Technical University of Liberec, Studentská 2, 461 17 Liberec, Czech Republic
[b] Institute of Information Theory and Automation, Pod vodárenskou věží 4, P.O. Box 18, 182 08 Praha 8, Czech Republic
[c] Laboratoire d'Astrophysique de Toulouse-Tarbes, Université de Toulouse, CNRS, 14 Av. Edouard Belin, 31400 Toulouse, France

ARTICLE INFO

Article history:
Received 15 September 2008
Received in revised form
25 March 2009
Accepted 16 April 2009
Available online 24 April 2009

Keywords:
Independent component analysis
Blind source separation
Cramér–Rao lower bound
FastICA algorithm
Piecewise stationary model

ABSTRACT

We address independent component analysis (ICA) of piecewise stationary and non-Gaussian signals and propose a novel ICA algorithm called Block EFICA that is based on this generalized model of signals. The method is a further extension of the popular non-Gaussianity-based FastICA algorithm and of its recently optimized variant called EFICA. In contrast to these methods, Block EFICA is developed to effectively exploit varying distribution of signals, thus, also their varying variance in time (nonstationarity) or, more precisely, in time-intervals (piecewise stationarity). In theory, the accuracy of the method asymptotically approaches Cramér–Rao lower bound (CRLB) under common assumptions when variance of the signals is constant. On the other hand, the performance is practically close to the CRLB even when variance of the signals is changing. This is demonstrated by comparing our algorithm with various methods that are asymptotically efficient within ICA models based either on the non-Gaussianity or the nonstationarity. The benefit of our algorithm is demonstrated by examples with real-world audio signals.

1. Introduction

The instantaneous linear mixture model is the basic configuration considered in blind source separation (BSS) [1]. The relation between unobserved original signals and observed measured signals is here given by equation

$$\mathbf{X} = \mathbf{AS}, \tag{1}$$

where $\mathbf{X}$ and $\mathbf{S}$ are matrices with N columns, each of which represents samples of the measured and the original signals, respectively. We will consider the regular case, thus, the number of rows of $\mathbf{X}$ and $\mathbf{S}$ is the same and equal to d, and the *mixing* matrix $\mathbf{A}$ is a $d \times d$ regular matrix.

Estimating the mixing matrix $\mathbf{A}$ or equivalently the original signals $\mathbf{S}$ from the data $\mathbf{X}$ is the general task of BSS. To solve this problem, a principle giving some assumption about the original signals should be introduced. The most popular one is based on the assumption of their statistical independence, which is used by a certain class of models that fall within a popular BSS discipline called independent component analysis (ICA) [2].

Since Comon's pioneering paper [3], numerous successful algorithms have been proposed using basic models based either on non-Gaussianity [4–6], nonstationarity [7–9] or spectral diversity (coloration) [10,11] of the original signals. Later, various improvements of the earlier methods were developed [12]. The most recent algorithms provide fast and reliable solutions while attaining the best possible accuracy fundamentally limited by the respective Cramér–Rao lower bound (CRLB) [13,15–17].

* Corresponding author at: Faculty of Mechatronics, Informatics and Interdisciplinary Studies, Technical University of Liberec, Studentská 2, 461 17 Liberec, Czech Republic. Tel.: +420 485 353534; fax: +420 485 353112.
E-mail addresses: zbynek.koldovsky@tul.cz (Z. Koldovský), jiri.malek@tul.cz (J. Málek), tichavsk@utia.cas.cz (P. Tichavský), yannick.deville@ast.obs-mip.fr (Y. Deville), shahram.hosseini@ast.obs-mip.fr (S. Hosseini).

0165-1684/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.sigpro.2009.04.021

While methods assuming non-Gaussianity of signals require computation of higher-order statistics (HOS), the methods using nonstationarity or spectral diversity usually need second-order statistics (SOS) only, which provides faster implementations usually through joint approximate diagonalization of a set of matrices; see e.g., [14,17–19]. On the other hand, each approach cannot separate sources if the respective assumptions are not met, which means certain limitations. For instance, the nonstationarity-based methods cannot separate signals having the same dynamics. In this respect, the non-Gaussianity-based methods are popular thanks to their widest application area, e.g., in telecommunications, biomedical signal processing, or speech and audio processing.

Since real signals may often exhibit both non-Gaussianity, nonstationarity or temporal structure, there are attempts to derive methods that combine two or more models to enhance the application area and to improve the performance [20,21]. However, the theoretical background of combined models is much more complex. Most methods therefore rely on various heuristically chosen criteria [22–24] or decision-driven combinations of basic algorithms [20,25], rather than optimizing the performance in a straightforward way through the theory.

This paper focuses on the model combining the non-Gaussianity and the nonstationarity assumptions through the so-called *piecewise stationary* model. The optimum solution of this model was discussed in [26], but few methods were proposed for finding it. The fully general framework was considered by Pham [27]. He proposed an algorithm, from here on named as NSNG, that performs (quasi)-maximum likelihood estimation (MLE), which yields excellent performance in theory. However, in our experimental tests [28], we have observed cases of instability and misconvergence of NSNG. Specifically, the algorithm seems to work well in simple scenarios, e.g., where "few" signals are separated and their properties perfectly fit the model. By contrast, the method failed with nonnegligible probability in more difficult examples or when separating real-world signals such as EEG data or real audio mixtures.

To provide a reliable algorithm with lower computational burden and comparable performance with that of NSNG, we here introduce a further extension of the very popular FastICA algorithm [5], which was originally developed for non-Gaussian signals. The method is called *Block EFICA*[1] as it is an extension of the EFICA algorithm [16] (a theoretically optimized FastICA variant for non-Gaussian signals) for piecewise stationary signals.

The paper is organized as follows. The following section introduces the piecewise stationary model and basic notations used throughout the paper. Section 3 surveys Cramér–Rao bounds that were derived for several levels of generalizations of the basic non-Gaussianity-based ICA model. The proposal of the Block EFICA algorithm is given in Section 4 after short descriptions of its forgoers: FastICA [5] and EFICA [16]. Section 5 provides performance analysis of behavior of FastICA under the assumption of piecewise stationary signals, and introduces optimized selection of important parameters of Block EFICA to achieve the best performance. Finally, experimental results demonstrating performance of the Block EFICA in comparison with other methods are presented by Section 6.

2. Piecewise stationary model

The basic ICA model exploiting non-Gaussianity of the sources is defined by

$$\mathbf{x} = \mathbf{A}\mathbf{s}, \tag{2}$$

where $\mathbf{s} = [s_1, \ldots, s_d]^T$ is a vector of independent random variables (RVs),[2] and each of them represents one of the unknown original signals. In practice, this means that the data matrices $\mathbf{X}$ and $\mathbf{S}$ consist of N i.i.d. realizations of $\mathbf{x}$ and $\mathbf{s}$, respectively, whose relation is given through the transform $\mathbf{A}$. The very assumption of independence of $s_1, \ldots, s_d$ is used for finding the de-mixing transform $\mathbf{A}^{-1}$, which can be achieved only up to an indeterminable order, scales, and signs of its rows.

Compared to the basic ICA model (2), the piecewise stationary model consists in that the samples of the original signals need *not* be identically distributed. The probability density function (pdf) $f_k(x)$ of s_k thus may be different at each time instant/interval.

However, to allow practical estimation of signal statistics on data blocks, we will assume that there are M blocks of $\mathbf{S}$ of the same integer length N/M, where, within each block, the distribution of the signals is unchanging. Therefore, we will use the superscript (I) to denote quantities, RVs or functions that are related to the I-th block. For instance, this means that for each block of data $\mathbf{X}$, say for the I-th block $\mathbf{X}^{(I)}$, it holds $\mathbf{X}^{(I)} = \mathbf{A}\mathbf{S}^{(I)}$, which corresponds to N/M i.i.d. realizations according to model

$$\mathbf{x}^{(I)} = \mathbf{A}\mathbf{s}^{(I)}, \tag{3}$$

where $\mathbf{x}^{(I)}$ and $\mathbf{s}^{(I)}$ are vectors of corresponding RVs.

A particular case of the piecewise stationary model, which will be called *Block Gaussian* model [14], is when all the distributions of all RVs in (3) are Gaussian. This means that all signals are white and Gaussian within each block, and the piecewise stationarity consists only in that their variances vary block-by-block.

3. Cramér–Rao lower bounds for independent component analysis

We discuss several bounds that limit the accuracy achievable by blind separation. Such limitation may be given by the Cramér–Rao lower bound that is related to the theoretical model of the original signals. In other words, the bound is different for different models requiring various assumptions about the original signals,

[1] The primary version of Block EFICA introduced in [28] was referred to as *Extended EFICA*.

[2] For simplicity, all RVs considered in the paper are assumed to have zero mean and finite variance.

whereby the separation principles are determined. Although the bounds presented here may all be derived as particular cases of the bound given for the most general piecewise stationary model, for convenience, we start the description with the basic ICA bound and then generalize it gradually.

In general, CRLB is defined for an unbiased estimator of some (multivariate) parameter θ, which is being estimated from a data vector x that has the probability density $f_{x|\theta}(x|\theta)$. CRLB is the lower bound for the covariance matrix of any unbiased estimator $\hat{\theta}$ of θ, i.e.,

$$\mathrm{cov}_\theta \hat{\theta} = \mathrm{E}_\theta[(\hat{\theta}-\theta)(\hat{\theta}-\theta)^T]. \tag{4}$$

If the following Fisher information matrix (FIM) and its inverse exist

$$\mathbf{F} = \mathrm{E}_\theta\left[\frac{1}{f^2_{x|\theta}(x|\theta)}\frac{\partial f_{x|\theta}(x|\theta)}{\partial\theta}\left(\frac{\partial f_{x|\theta}(x|\theta)}{\partial\theta}\right)^T\right], \tag{5}$$

under the regularity conditions [29] it holds that

$$\mathrm{cov}_\theta\,\hat{\theta} \geqslant \mathbf{F}^{-1} = \mathrm{CRLB}[\theta],$$

where the matrix inequality means that the matrix $\mathrm{cov}_\theta\,\hat{\theta} - \mathbf{F}^{-1}$ is positive semidefinite.

In case of the instantaneous linear mixture $\mathbf{X}=\mathbf{AS}$, the parameters intended for the estimation are the elements of $\mathbf{A}^{-1}$. Let $\mathbf{W}$ be an unbiased estimator of $\mathbf{A}^{-1}$. Instead of considering CRLB of $\mathbf{W}$, it is useful to derive the bound for the so-called *gain* matrix $\mathbf{G}=\mathbf{WA}$. Without loss of generality, the indeterminacies of order, signs, and scales of the original signals can be assumed to be resolved. $\mathbf{G}$ should then be close to the identity, and the variances of its nondiagonal elements, $\mathrm{var}[\mathbf{G}_{k\ell}]$, $k\neq\ell$, reflect mean value of residual interference between the separated signals $\mathbf{WX}$. Such a criterion, which is commonly used in signal processing, reflects well the accuracy of the estimator $\mathbf{W}$.

The CRLB for the basic ICA model (2) has been well known since the 1990s [30,31]. We will denote the bound by CRLB_1, and it is given by

$$\mathrm{CRLB}_1[\mathbf{G}_{k\ell}] = \frac{1}{N}\frac{\kappa_\ell}{\kappa_k\kappa_\ell - 1}, \quad k\neq\ell, \tag{6}$$

where $\kappa_k = \mathrm{E}[\psi_k^2(x)]$ with $\psi_k(x) = -f_k'(x)/f_k(x)$ being the score function of the probability density function $f_k(x)$ of the k-th RV s_k. Here s_k is assumed to have unit variance, thus, note that κ's are defined for unit-variance score functions.

The bound for the piecewise stationary model (3) with constant (unit) variance signals, denoted by CRLB_2, is given by [28]

$$\mathrm{CRLB}_2[\mathbf{G}_{k\ell}] = \frac{1}{N}\frac{\overline{\kappa}_\ell}{\overline{\kappa}_k\overline{\kappa}_\ell - 1}, \quad k\neq\ell, \tag{7}$$

where $\overline{\kappa}_k \overset{\mathrm{def}}{=} (1/M)\sum_{l=1}^M \kappa_k^{(l)}$.

Now we introduce the most general bound, i.e., for the piecewise stationary model (3) where the variance of the signals is not assumed to be constant. Let $\sigma_k^{2(l)}$ be the variance of $s_k^{(l)}$, $k=1,\ldots,d$, $l=1,\ldots,M$, but $\kappa_k^{(l)}$ is still defined for pdf $f_k^{(l)}(\cdot)$ normalized so as to correspond to RV *normalized to unit variance*. Then, the bound could be written in the form

$$\mathrm{CRLB}_3[\mathbf{G}_{k\ell}] = \frac{1}{N}\frac{B_{k\ell}}{A_{k\ell}B_{k\ell}-1}, \quad k\neq\ell, \tag{8}$$

where

$$A_{k\ell} = \frac{1}{M}\sum_{l=1}^M \frac{\sigma_\ell^{2(l)}}{\sigma_k^{2(l)}}\kappa_k^{(l)}, \tag{9}$$

$$B_{k\ell} = \frac{1}{M}\sum_{l=1}^M \frac{\sigma_k^{2(l)}}{\sigma_\ell^{2(l)}}\kappa_\ell^{(l)}. \tag{10}$$

This result was previously derived, e.g., in [26]. In Appendix A, we provide a simple derivation of the bound using the derivation of FIM from [32].

For the sake of completeness, we introduce the CRLB for the Block Gaussian model, i.e., when all distributions of signals are Gaussian. The bound easily follows from (8) by taking $\kappa_k^{(l)}=1$ in (9) and (10). We will denote this bound by CRLB_4.

4. Block EFICA algorithm

We here describe our novel algorithm that is an extension of its previous variants FastICA and EFICA. First, the underlying methods are reminded in short as they were proposed for solving the model (2). Second, the building block of the proposed algorithm is given, which is a straightforward extension of the one-unit FastICA algorithm to the piecewise stationary signals. Finally, we introduce the proposed algorithm.

4.1. FastICA and EFICA algorithms

The FastICA algorithm [5] was originally derived as a method for solving the basic ICA problem (2). It is based on optimization of a contrast function

$$c(\mathbf{w}_k) = \mathrm{E}[G(\mathbf{w}_k^T\mathbf{z})], \tag{11}$$

subject to the vector $\mathbf{w}_k^T$, whose optimum value is the k-th row of de-mixing transform. The function $G(\cdot)$, which applies elementwise, is a properly chosen nonlinear function whose derivative will be denoted by $g(\cdot)$. The vector $\mathbf{z}$ is derived by transforming signals $\mathbf{x}$ so that the components of $\mathbf{z}$ are not correlated and have unit variance. After this preprocessing, which is commonly referred to as sphering, it holds that $\mathrm{E}[\mathbf{z}\mathbf{z}^T]=\mathbf{I}$.

The optimization of $c(\mathbf{w}_k)$ is based on the iteration

$$\mathbf{w}_k^+ \leftarrow \mathrm{E}[\mathbf{z}g(\mathbf{w}_k^T\mathbf{z})] - \mathbf{w}_k\mathrm{E}[g'(\mathbf{w}_k^T\mathbf{z})]. \tag{12}$$

In practice, i.e., when working with a finite number of signal samples, the theoretical expectations are replaced by respective sample means, thus the resulting de-mixing vectors/matrices are respective estimates thereof.

The original FastICA was developed in two basic versions: the one-unit and the symmetric one. While the one-unit FastICA completes each iteration by normalizing the vector $\mathbf{w}_k^+$, the symmetric FastICA computes d iterations (12) in parallel and does a symmetric

orthogonalization[3] of $[\mathbf{w}_1^+,\dots,\mathbf{w}_d^+]^T$ to yield all rows of the de-mixing matrix, whose practical estimate will be denoted by $\widehat{\mathbf{W}}$.

The theoretical (asymptotic) performance [32] of the one-unit FastICA is characterized by

$$\mathrm{var}[\mathbf{G}_{k\ell}^{1\mathrm{U}}] \approx \frac{1}{N} V_{k\ell}^{1\mathrm{U}}, \quad k \neq \ell, \tag{13}$$

where $\mathbf{G}^{1\mathrm{U}} = \widehat{\mathbf{W}}\mathbf{A}$ is the gain matrix, each of its rows corresponds to the estimation of one de-mixing vector, and $V_{k\ell}^{1\mathrm{U}} = \gamma_k/\tau_k^2$ with

$$\begin{aligned} &\mu_k = \mathrm{E}[s_k g(s_k)], \quad \gamma_k = \beta_k - \mu_k^2, \\ &\nu_k = \mathrm{E}[g'(s_k)], \quad \tau_k = \nu_k - \mu_k, \\ &\beta_k = \mathrm{E}[g^2(s_k)]. \end{aligned} \tag{14}$$

In case that the expectations do not exist it may signify either bad choice of the function $G(\cdot)$ or zero leading term in the asymptotic expansion of the variance (13). It is a well-known feature in ICA that the optimum choice of $G(\cdot)$ comes up to $g(\cdot)$ being the score function of s_k, $\psi_k(\cdot)$ [13,31].

Among other things, this knowledge is taken into account by the recently published EFICA algorithm [16], which is designed to attain the best possible performance limited by (6). The method proceeds in three steps: (1) it preestimates all the original signals by means of the symmetric FastICA with the test of saddle points [32], (2) for each $k = 1,\dots,d$, it adaptively chooses a nonlinearity $g \overset{\mathrm{def.}}{=} g_k$ that approximates the score function $\psi_k(\cdot)$, and (3) it does fine-tunings and a refinement.

The fine-tunings consist in further one-unit FastICA iterations for each signal separately, using the nonlinearities found in step 2. The resulting de-mixing vectors from the fine-tunings $\mathbf{w}_1^+,\dots,\mathbf{w}_d^+$ are then optimally combined by the refinement.

The refinement utilizes optimum weights computed according to

$$c_{k\ell} = \begin{cases} \dfrac{V_{k\ell}^{1\mathrm{U}}}{V_{\ell k}^{1\mathrm{U}}+1}, & k \neq \ell, \\ 1, & k = \ell. \end{cases} \tag{15}$$

We remark that we use slightly different definition of the weights from that in [16] since it is handier for forthcoming description of the Block EFICA. The modification simply consists in normalizing the vectors $\mathbf{w}_1^+,\dots,\mathbf{w}_d^+$, which was not done in [16]; see [33] for details. The weights are used to form matrix

$$\mathbf{W}_k^+ = [c_{k1}\mathbf{w}_1^+/\|\mathbf{w}_1^+\|,\dots,c_{kd}\mathbf{w}_d^+/\|\mathbf{w}_d^+\|]^T. \tag{16}$$

The k-th row of symmetrically orthogonalized version of $\mathbf{W}_k^+$, i.e., of $(\mathbf{W}_k^+\mathbf{W}_k^{+T})^{-1/2}\mathbf{W}_k^+$, yields the final estimate of $\mathbf{w}_k$. This is done for each $k = 1,\dots,d$ separately, which relaxes the orthogonality constraint [31] introduced by the symmetric FastICA.

[3] We use the well-established term "symmetric orthogonalization" although "symmetric orthonormalization" would be more accurate.

The asymptotic performance of EFICA is given by

$$\mathrm{var}[\mathbf{G}_{k\ell}^{\mathrm{EF}}] \approx \frac{1}{N}\frac{V_{k\ell}^{1\mathrm{U}}(V_{\ell k}^{1\mathrm{U}}+1)}{V_{k\ell}^{1\mathrm{U}}+V_{\ell k}^{1\mathrm{U}}+1}, \quad k \neq \ell. \tag{17}$$

The particular case when $g_k = \psi_k$ reveals superior property of EFICA. It holds, then, that $\beta_k = \nu_k = \kappa_k$, $\mu_k = 1$, and $V_{k\ell}^{1\mathrm{U}} = 1/(\kappa_k - 1)$. Substituting this into (17) gives

$$\mathrm{var}[\mathbf{G}_{k\ell}^{\mathrm{EF}}] \approx \frac{1}{N}\frac{\kappa_\ell}{\kappa_k\kappa_\ell - 1}, \quad k \neq \ell.$$

Compared to the CRLB_1 given by (6), the asymptotic efficiency of EFICA in the framework of the basic ICA model (2) follows.

4.2. One-unit FastICA for piecewise stationary signals

To take into account the piecewise stationary model, we introduce a new definition of the contrast function (11), which is

$$c(\mathbf{w}_k) = \lambda_k^{(1)}\mathrm{E}[G_k^{(1)}(\mathbf{w}_k^T\mathbf{z}^{(1)})] + \cdots\cdots + \lambda_k^{(M)}\mathrm{E}[G_k^{(M)}(\mathbf{w}_k^T\mathbf{z}^{(M)})], \tag{18}$$

where $G_k^{(1)},\dots,G_k^{(M)}$ are properly chosen nonlinear functions, and $\lambda_k^{(1)},\dots,\lambda_k^{(M)}$ denote some weights.

It should be noted that this contrast *cannot* be viewed as the contrast (11) with $G(\cdot)$ being a linear combination of $G_k^{(1)},\dots,G_k^{(M)}$, because each expectation in (18) depends on different distributions from corresponding block of the signals. Also, an important fact is that each term in (18) is a valid contrast function itself. Since the mixing matrix is the same in all blocks, (18) is a valid contrast function as well. In other words, all the contrasts represented by the terms in (18) have the same optimum points.

One-unit FastICA using the contrast function (18), from here on referred to as *block one-unit FastICA*, works in the way that it applies a different nonlinearity $g(\cdot)$ on each block of signals. Thus, the iteration (12) changes to

$$\begin{aligned} \mathbf{w}_k^+ \leftarrow{}& \lambda_k^{(1)}(\mathrm{E}[\mathbf{z}^{(1)}g_k^{(1)}(\mathbf{w}_k^T\mathbf{z}^{(1)})] - \mathbf{w}_k\mathrm{E}[g_k^{(1)'}(\mathbf{w}_k^T\mathbf{z}^{(1)})]) \\ &+ \cdots + \lambda_k^{(M)}(\mathrm{E}[\mathbf{z}^{(M)}g_k^{(M)}(\mathbf{w}_k^T\mathbf{z}^{(M)})] \\ &- \mathbf{w}_k\mathrm{E}[g_k^{(M)'}(\mathbf{w}_k^T\mathbf{z}^{(M)})]), \end{aligned} \tag{19}$$

and the expectations are replaced by sample means in practice.

As can be seen, the original one-unit FastICA is, when setting $\lambda_k^{(l)} = 1/M$ and $g_k^{(l)} = g$, for all $l = 1,\dots,M$, a particular case of the block version introduced here. Theoretical conclusions derived later in this paper, therefore, yield an insight into the behavior of the original FastICA (and also of other variants of FastICA) when distributions of signals are different from one block to the other.

4.3. Proposed Block EFICA algorithm

The Block EFICA algorithm takes into account the piecewise stationarity of signals. The approach consists of the following three steps that are similar to those in the original EFICA up to the difference that consists in linking the choice of nonlinearities with the fine-tuning into a

common step due to higher precision. Also a different approach for the choice of nonlinear functions is used, because variance of signals in blocks need not be equal to one as assumed by the approach used in EFICA.

BEF1 Separation by the symmetric FastICA with the test of saddle points in order to obtain a preestimate of the de-mixing matrix $\widehat{\mathbf{W}}$.

BEF2 Fine-tuning of each row of $\widehat{\mathbf{W}}$ by means of the block one-unit FastICA (Section 4.2). The weights and the nonlinearities in (19) are simultaneously updated as described below. The simplified version of the algorithm, called *Uniform Block EFICA*, selects all the weights equal to an arbitrary nonzero value.

BEF3 The refinement to get the most accurate and final estimate of the whole de-mixing matrix.

A simplified illustration of the flow of Block EFICA is shown in Fig. 1. In the following, we provide more details on the steps of the algorithm.

The pre-estimation of the whole de-mixing matrix in BEF1 could be done by any ICA method, which opens up possible variations of the Block EFICA. Nevertheless, our selection, the symmetric FastICA with the test of saddle points, proves being suitable for wide variety of scenarios [16]. The method allows fast and reliable separation of non-Gaussian signals. Moreover, in practice it generally allows significant separation of piecewise stationary signals as well, which follows from the fact that the symmetric FastICA is a special variant of the block one-unit FastICA introduced in the previous subsection. However, it has limited accuracy due to the nonoptimal choice of the nonlinearity that is fixed for all signals and blocks, and also due to the orthogonality constraint [31] introduced by the symmetric orthogonalization. Therefore, it is a suitable initialization for the fine-tuning done by step BEF2.

In the fine-tunings (BEF2), the estimation of the k-th signal, $k=1,\dots,d$, is improved by starting the block one-unit FastICA using appropriately chosen functions $g_k^{(l)}(\cdot)$ and the weights $\lambda_k^{(l)}$, $l=1,\dots,M$. Since the best choice of $g_k^{(l)}(\cdot)$ is the score function $\psi_k^{(l)}(\cdot)$, we use the approximation by Pham's estimator from [13]. The details are given below in the extra subsection.

The choice of the weights $\lambda_k^{(l)}$ has an influence on the performance of fine-tunings as well. Since it should be analyzed first, the choice is given afterwards in Section 5 by the expression (27). In that section, we also justify the introduction of the Uniform Block EFICA algorithm, which sets all the weights to a constant.

Finally, the refinement step is done in the similar way as in the original EFICA [16]. The fine-tuned and normalized rows of the separating matrix resulting from BEF2, $\mathbf{w}_1,\dots,\mathbf{w}_d$, and the weights $c_{k\ell}$ are used to form matrix

$$\mathbf{W}_k^+ = [c_{k1}\mathbf{w}_1,\dots,c_{kd}\mathbf{w}_d]^T.$$

Then, the k-th row of the matrix $(\mathbf{W}_k^+\mathbf{W}_k^{+T})^{-1/2}\mathbf{W}_k^+$ yields the final estimate of $\mathbf{w}_k$. The difference compared to EFICA is that the weights $c_{k\ell}$ should be computed accordingly. Namely, (15) is in fact a function of the performance achieved by the fine-tuning in EFICA, i.e., by that of the one-unit FastICA given by (13). However, the fine-tuning in the Block EFICA is done by means of the block one-unit FastICA algorithm, whose performance is different. The performance is analyzed in Section 5, where the analytical expression (28) for the weights follows.

4.4. Parametric estimation of score functions

Parametric estimation of score functions is a well-established problem in statistical theory [34]. The parametric estimator proposed in [13] is suited for the problems tackled by ICA algorithms. It is defined as the minimizer of the mean square distance between a score function $\psi(\cdot)$ and a linear combination of K *basis* functions $h_1(x),\dots,h_K(x)$, i.e.,

$$\min_{\theta_1,\dots,\theta_K} \mathrm{E}\left[\left(\psi(x)-\sum_{i=1}^{K}\theta_i h_i(x)\right)^2\right]. \tag{20}$$

The merit consists in the fact that $\mathrm{E}[\psi(x)h(x)]=\mathrm{E}[h'(x)]$ for any function $h(x)$. Thanks to this, the minimization is possible without knowledge of $\psi(\cdot)$ and is fast, because it only requires estimation of $\mathrm{E}[h_i(x)h_j(x)]$ and $\mathrm{E}[h_i'(x)]$,

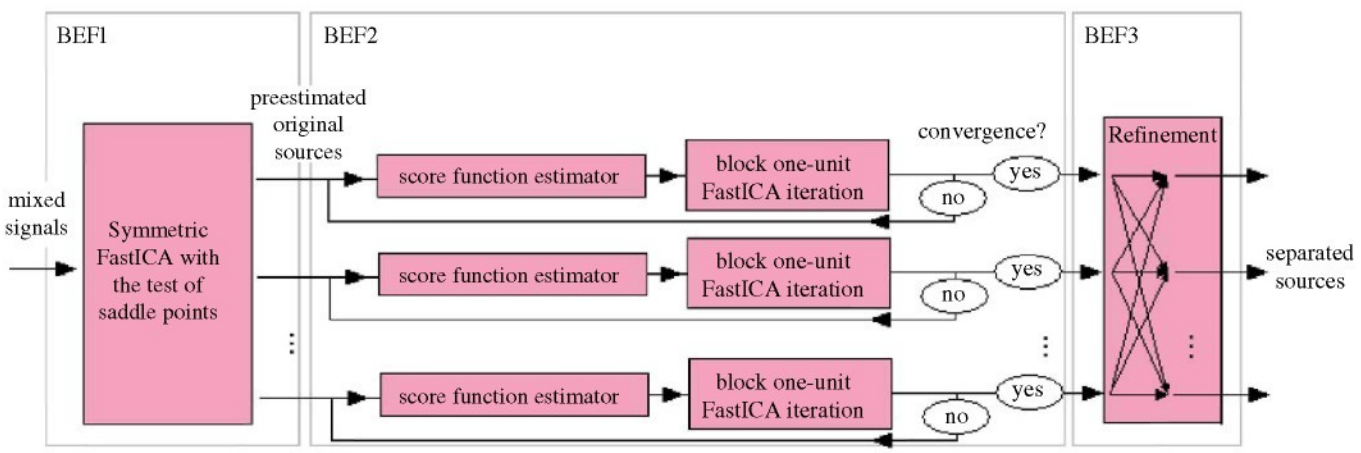


Fig. 1. Flow of the Block EFICA algorithm.

$i,j = 1,\dots,K$. The minimizer is then given by the solution of a set of K linear equations.

In our implementation, we have decided for two ($K = 2$) basis functions: $h_1(x) = x^3$, that is good for sub-Gaussian sources, and $h_2(x) = x/(1+6|x|)^2$ working well with super-Gaussian sources [37]. This choice turns out to be appropriate for a wide class of distributions and offers a good trade-off between accuracy, speed, and flexibility. For instance, when considering signals with generalized Gaussian distributions, the estimator (20) with our settings used within EFICA yields comparable results with the adaptation originally used thereby [35].

Another advantage of this estimator consists in computational savings: once the moments $\mathrm{E}[h_i(x)h_j(x)]$ and $\mathrm{E}[h_i'(x)]$ are estimated, the results can be used where corresponding moments occur, which is, e.g., in the iteration (19). The burden due to the solution of minimizing equations is, for $K = 2$, negligible, thus, the main slowing-down compared to the adaptation used in EFICA consists in that two nonlinear functions h_1 and h_2 must be evaluated.

Here, we should point out that it is relevant to take into account the identity function $h_i(x) = x$ for the third basis function in (20). Unlike in case of the original FastICA/EFICA, this is meaningful in Block EFICA, because each block $\mathbf{z}^{(l)}$ of the sphered data $\mathbf{z}$ may *not* be sphered. Specifically, when considering $g(x) = \alpha x + h(x)$ in (12) with an arbitrary α and a nonlinearity $h(x)$, the effect of the term αx is zeroed no matter how α is chosen since $\mathrm{E}[\mathbf{z}\mathbf{z}^T] = \mathbf{I}$. It is not so in case of the "block-iteration" (19) due to nonsphered blocks $\mathbf{z}^{(l)}$.

Inclusion of the identity function into the score function estimator, in fact, conveys direct utilization of second-order statistics of signals. The consideration is worthwhile especially when separating signals with changing variance. Therefore, we consider this as an option in the Block EFICA, which is slightly more computationally expensive.

4.5. Choice of the number of blocks

The correct number of blocks M is usually not known in practice. The goal is to choose M such that the distribution of $\mathbf{S}$ may be regarded as constant within each block. On the other hand, M should not be overestimated, because overparametrization may cause higher estimation error. Luckily, Block EFICA is not highly sensitive to this parameter, which is demonstrated by results shown in Fig. 4 in Section 6. It is shown that significant overestimations of M as well as its underestimations do not decrease the performance seriously.

Usually, the choice of optimum M is done by taking into account characteristics of signals to be separated. For example, when separating speech signals, it is worth to select M such that the length of blocks corresponds to 20–25 ms where speech is almost stationary.

Blind selection of M may be based on estimation of residual inter-signal interference (signal-to-interference ratio—SIR) using analytical expressions (29) where corresponding statistics are estimated from separated signals. It is thus possible to see the estimated SIR of separated signals as a function of M. At the beginning, SIR usually improves with growing M, but for larger M the growth is slower and slower. We would select M where the increase of SIR becomes slow; see Fig. 4. A similar approach but more computationally demanding would be when Block EFICA was started with different M's taken from a reasonable range, and the optimum M or its effective value was selected subject to the resulting estimate of SIR. Another possible approach for automated choice of M can be found, e.g., in [36].

5. Performance analysis

In this section, we analyze performance of the proposed Block EFICA algorithm to reveal influence of its parameters on accuracy of separation. Optimization of the theoretical performance subject to the parameters gives their final definition, which also completes the description of the algorithm.

The starting point of the analysis is the derivation of the performance of the block one-unit FastICA, which is achieved by the fine-tunings in BEF2. We generalize the analysis of the original one-unit FastICA from [32] that considers the basic ICA model (2) to the piecewise stationary model with M blocks. The analysis yields the result summarized in the following proposition.

Proposition 1. *For $k = 1,\dots,d$ and $l = 1,\dots,M$, assume that*

(i) *the RVs $s_k^{(l)}$ have zero mean and finite variance $\sigma_k^{2(l)}$ such that it holds that (unit scale)*

$$\frac{1}{M}\sum_{l=1}^{M}\sigma_k^{2(l)} = 1,$$

(ii) *the functions $g_k^{(l)}$ are twice continuously differentiable,*

(iii) *the following expectations exist:*

$$\begin{aligned}\mu_k^{(l)} &= \mathrm{E}[s_k^{(l)} g_k^{(l)}(s_k^{(l)})],\\ v_k^{(l)} &= \mathrm{E}[g_k^{(l)\prime}(s_k^{(l)})],\\ \beta_k^{(l)} &= \mathrm{E}[g_k^{2(l)}(s_k^{(l)})],\end{aligned} \tag{21}$$

and

(iv) *the block one-unit FastICA algorithm is started from the correct de-mixing matrix and stops after a single iteration (19).*

Then, the normalized gain matrix elements $N^{1/2}\mathbf{G}_{k\ell}^{\mathrm{B1U}}$ have asymptotically Gaussian distribution $\mathcal{N}(0, V_{k\ell}^{\mathrm{B1U}})$, where

$$V_{k\ell}^{\mathrm{B1U}} = \frac{\overline{\beta}_{k\ell} + \alpha_{k\ell}^2\overline{\sigma}_{k\ell}^2 - 2\alpha_{k\ell}\overline{\mu}_{k\ell}}{\overline{\tau}_k^2} \tag{22}$$

for $k,\ell = 1,\dots,d$, $k \neq \ell$, provided that $\overline{\tau}_k \neq 0$. Here,

$$\overline{\mu}_k = \frac{1}{M}\sum_{l=1}^{M}\lambda_k^{(l)}\mu_k^{(l)},$$

$$\overline{v}_k = \frac{1}{M}\sum_{l=1}^{M}\lambda_k^{(l)}v_k^{(l)},$$

$$\overline{\tau}_k = \overline{v}_k - \overline{\mu}_k,$$
$$\overline{\beta}_{k\ell} = \frac{1}{M}\sum_{l=1}^{M}(\lambda_k^{(l)})^2\beta_k^{(l)}\sigma_\ell^{2(l)},$$
$$\overline{\mu}_{k\ell} = \frac{1}{M}\sum_{l=1}^{M}\lambda_k^{(l)}\mu_k^{(l)}\sigma_\ell^{2(l)},$$
$$\overline{v}_{k\ell} = \frac{1}{M}\sum_{l=1}^{M}\lambda_k^{(l)}v_k^{(l)}\sigma_\ell^{2(l)},$$
$$\overline{\sigma}^2_{k\ell} = \frac{1}{M}\sum_{l=1}^{M}\sigma_k^{2(l)}\sigma_\ell^{2(l)},$$
$$\alpha_{k\ell} = \overline{\mu}_k + (\overline{v}_{k\ell} - \overline{v}_k)/2. \tag{23}$$

Proof. See Appendix B.

The practical conclusion of this proposition is that the variance of the gain matrix elements obtained from the block one-unit FastICA is approximately

$$\mathrm{var}[\mathbf{G}_{k\ell}^{\mathrm{B1U}}] \approx \frac{1}{N}V_{k\ell}^{\mathrm{B1U}}, \quad k\neq\ell. \tag{24}$$

Consequently, the aim is to minimize $V_{k\ell}^{\mathrm{B1U}}$ subject to free parameters (weights) to achieve the best performance in practice. Note that M need not be necessarily equal to a particular value. The proposition is valid if M is such that distributions of signals are constant within each block.

As will be shown later, all the expectations (21), and consequently (23), are important for computing optimum weights (λ's and later the weights for the refinement) needed to achieve the optimum performance. In practice, the expectations are estimated from estimated signals by sample means. Estimation errors are therefore introduced into the weights. Then, the need is that the weights are not much sensitive to the estimation errors so as not to worsen the final performance of the algorithm in practice.

Here we arrive at the problem with the fully general piecewise stationary model. We have found that the resulting formulas for the weights (not shown here to simplify the text) are overparametrized, which causes the higher sensitivity of the weights to the estimation errors of (21) and of the variances $\sigma_k^{2(l)}$. Therefore, to reduce the number of parameters, we introduce an important simplification by assuming the same (unit) variance of signals in all blocks, i.e.,

$$\sigma_k^{2(l)} = 1, \quad k = 1,\dots,d,\ l = 1,\dots,M. \tag{25}$$

Although the assumption restricts our theoretical conclusions to constant-variance signals, we will show by simulations that the performance of the method is not depressed in practice when the variance of signals is changing. The main reason is that the expectations in (21) depend on the distribution of signals and reflect thus the variance sufficiently, and the parameters $\sigma_k^{2(l)}$ become redundant. This is yet more apparent when the identity function is considered in the score function estimator. The variances are then involved in the moments (21), because the functions $g_k^{(l)}$ have the form $g_k^{(l)}(x) = \alpha x + h_k^{(l)}(x)$, where $h_k^{(l)}(x)$ is a combination of nonlinearities.

By using the constant-variance assumption, (22) simplifies to

$$V_{k\ell}^{\mathrm{B1U}} = \frac{\overline{\beta}_k - \overline{\mu}_k^2}{\overline{\tau}_k^2}, \quad k\neq\ell, \tag{26}$$

where $\overline{\beta}_k = (1/M)\sum_{l=1}^{M}(\lambda_k^{(l)})^2\beta_k^{(l)}$. Now, we derive the optimal choice of $\lambda_k^{(1)},\dots,\lambda_k^{(M)}$ by minimizing (26). The result is described by the following proposition.

Proposition 2. *For a fixed $k\in\{1,\dots,d\}$, minimization of $V_{k\ell}^{\mathrm{B1U}}$ given by (26) subject to $\lambda_k^{(1)},\dots,\lambda_k^{(M)}$ is achieved for all $\ell = 1,\dots,d$, $\ell\neq k$, when*

$$\lambda_k^{(J)} = \frac{1}{M}\left(\frac{\tau_k^{(J)}}{\beta_k^{(J)}} + A_k B_k \frac{\mu_k^{(J)}}{\beta_k^{(J)}}\right), \quad J = 1,\dots,M, \tag{27}$$

where

$$A_k = \left(\sum_{l=1}^{M}\frac{\gamma_k^{(l)}}{\beta_k^{(l)}}\right)^{-1}$$

and

$$B_k = \sum_{l=1}^{M}\frac{\mu_k^{(l)}\tau_k^{(l)}}{\beta_k^{(l)}}.$$

Proof. See Appendix C.

After knowing the performance achieved by the fine-tunings stage BEF2, the final performance of the Block EFICA is given after the refinement step BEF3. The refinement, in the original EFICA, utilizes weights $c_{k\ell}$ given by (15), which, in fact, are functions of the performance achieved by the fine-tunings characterized by $V_{k\ell}^{1\mathrm{U}}$. Thanks to this relation, the weights that are optimal for the Block EFICA are simply given when inserting $V_{k\ell}^{\mathrm{B1U}}$ into (15) instead of $V_{k\ell}^{1\mathrm{U}}$.

Namely, the optimum weights $c_{k\ell}$ for the Block EFICA refinement are given by

$$c_{k\ell} = \begin{cases} \dfrac{V_{k\ell}^{\mathrm{B1U}}}{V_{\ell k}^{\mathrm{B1U}}+1}, & k\neq\ell, \\ 1, & k=\ell. \end{cases} \tag{28}$$

Similarly, the performance of the Block EFICA is analogous to (17), i.e., for $\mathbf{G}^{\mathrm{BEF}}$ being the resulting gain matrix,

$$\mathrm{var}[\mathbf{G}_{k\ell}^{\mathrm{BEF}}] \approx \frac{1}{N}\frac{V_{k\ell}^{\mathrm{B1U}}(V_{\ell k}^{\mathrm{B1U}}+1)}{V_{k\ell}^{\mathrm{B1U}}+V_{\ell k}^{\mathrm{B1U}}+1}, \quad k\neq\ell. \tag{29}$$

5.1. Optimal performance

Here, we study the special case when the nonlinearities selected by the score function estimator (20) equal the true score functions, i.e., $g_k^{(l)} = \psi_k^{(l)}$, for $k = 1,\dots,d$, $l = 1,\dots,M$.

Similarly to the equations above after (17), it holds that $\beta_k^{(l)} = v_k^{(l)} = \kappa_k^{(l)}$, $\mu_k^{(l)} = 1$, and $\tau_k^{(l)} = \gamma_k^{(l)} = \kappa_k^{(l)} - 1$. Next, the formula for λ's (27) simplifies to a constant, namely, $\lambda_k^{(l)} = 1/M$, but we may consider all λ's equal to one, because then $\overline{\beta}_k = \overline{v}_k = \overline{\kappa}_k$ and $\overline{\mu}_k = 1$. Now, the

performance (26) becomes equal to

$$V_{k\ell}^{\mathrm{B1U}} = \frac{1}{\overline{\kappa}_k - 1}. \tag{30}$$

Inserting (30) into (29) we get

$$\mathrm{var}[\mathbf{G}_{k\ell}^{\mathrm{BEF}}] \approx \frac{1}{N}\frac{\overline{\kappa}_\ell}{\overline{\kappa}_k\overline{\kappa}_\ell - 1}, \quad k \neq \ell. \tag{31}$$

As compared to the CRLB_2 given by (7), it follows that the Block EFICA is asymptotically efficient within the piecewise stationary model with constant variance signals. Although this does not mean the asymptotic efficiency of Block EFICA for the fully general model, we will show by simulations that its performance is usually very close to the CRLB even when variances of signals are not constant.

The uniformity of the weights (27) for the particular case studied here gives rise to the Uniform Block EFICA, as defined in Section 4.3, because $d \cdot M$ parameters $\lambda_k^{(l)}$ need not be estimated when $g_k^{(l)}(\cdot)$ are assumed to be the score functions. This means further reduction of parameters, which may be useful, for instance, when the number of blocks M is unknown and may be overestimated.

6. Experimental results

We have done several experiments simulating various scenarios to demonstrate good performance and versatility of the proposed Block EFICA algorithm [45]. In comparisons, we select algorithms that are supposed to be the most competitive for a given scenario. Thus, the original symmetric FastICA algorithm [5] with the nonlinearity $g(\cdot) = \tanh(\cdot)$ and the original EFICA algorithm [16,37] are considered as competitive methods within non-Gaussianity-based approaches. In several examples, we also consider the BGL algorithm from [14] that is designed for Gaussian nonstationary signals.

The NSNG algorithm [27] stands for a method belonging to the same class of algorithms as Block EFICA. As stated in Section 1, the method performs well in simple examples with "few" signals, but it is considerably unstable in more realistic scenarios. Therefore, we show its performance only in cases where the method yields meaningful results (Figs. 2 and 6).

A common criterion used in experiments is the interference-to-signal ratio (ISR), for the k-th separated signal defined as

$$\mathrm{ISR}_k = \frac{\sum_{\ell=1,\ell\neq k}^{d} \mathbf{G}_{k\ell}^2}{\mathbf{G}_{kk}^2}, \tag{32}$$

where $\mathbf{G} = \mathbf{WA}$ is the gain matrix computed as the product of the separation matrix $\mathbf{W}$ obtained by an algorithm and the known mixing matrix $\mathbf{A}$. Prior to the computation, the rows of $\mathbf{G}$ are permuted to avoid the indeterminacy of their original order. Such permutation is naturally chosen to yield the best value of the criterion.

For each experiment, we show the average computational loads of methods in legends of the corresponding figures. All simulations were running in Matlab™ on a PC with 3 GHz processor and 2 GB of RAM.

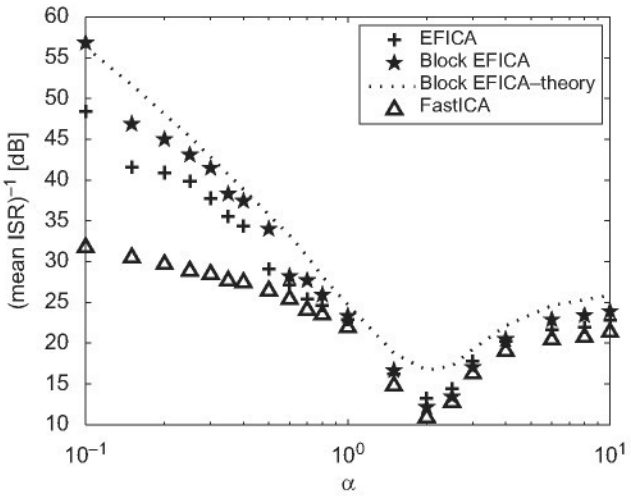


Fig. 2. Mean interference-to-signal ratio of separated signals in the first experiment computed over 100 Monte Carlo trials. Note that CRLB is not defined here, because, in the mixture of 20 signals, some of them are distributed according to RVs with generalized Gaussian distribution with $\alpha \leqslant 0.5$, which have $\kappa = +\infty$; see e.g., Appendix B in [16].

6.1. Validation of the analysis

The examples presented in this subsection aim at validating theoretical conclusions derived in Section 5 and at demonstrating the performance of the Block EFICA in the framework of the piecewise stationary model with constant-variance signals.

To this end, we compare the proposed Block EFICA with the original EFICA algorithm, which performs efficiently when working with signals with generalized Gaussian distributions [16] with parameter α, GGD(α), obeying the basic ICA model (2). However, in the experiments presented here we consider signals with varying distribution from one block to the other. Then, the behavior of EFICA is explained by the analysis of Block EFICA: using the same nonlinear functions in all blocks, the functions cannot match the varying score functions nor the weight for fine-tunings and refinements, thus, the performance of EFICA is *suboptimal*. The same holds for the other FastICA variants.

In the first example, we separate 20 artificial signals of length $N = 10^4$ mixed by a random matrix. Each signal consists of four blocks of the same length $N/4$. The first and the third blocks have Gaussian distribution, which is equivalent with GGD(2), and the second and the fourth blocks have the distribution GGD(α). The parameter α is fixed for each of 20 signals, where its values are uniformly chosen from $[0.1, 10]$. The variance of all the distributions is one, thus, the signals have constant variance.

Theoretical performance, marked in figures by "theory" in the legend, was estimated from separated signals using (26) and (29). Results of this experiment corroborate validity of the analysis due to proximity of the theoretical results with the empirical ones. They also demonstrate the improved performance of the proposed method compared to EFICA thanks to considering different distributions on the four blocks of signals. We do not demonstrate the

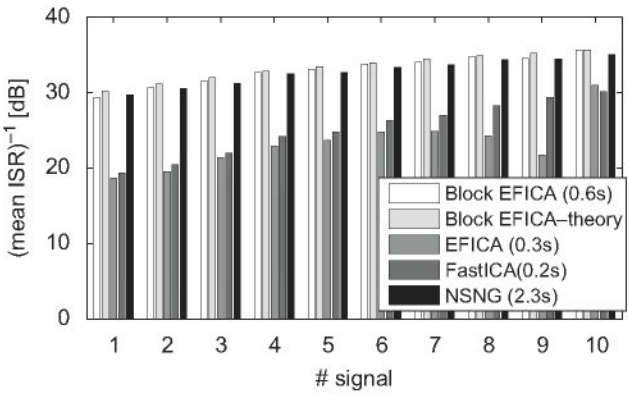


Fig. 3. Mean interference-to-signal ratio of 10 signals of length $N = 10^4$ averaged over 1000 Monte Carlo trials. The first $k \cdot N/10$ samples of the k-th signal are uniformly distributed, and the remainder is Gaussian.

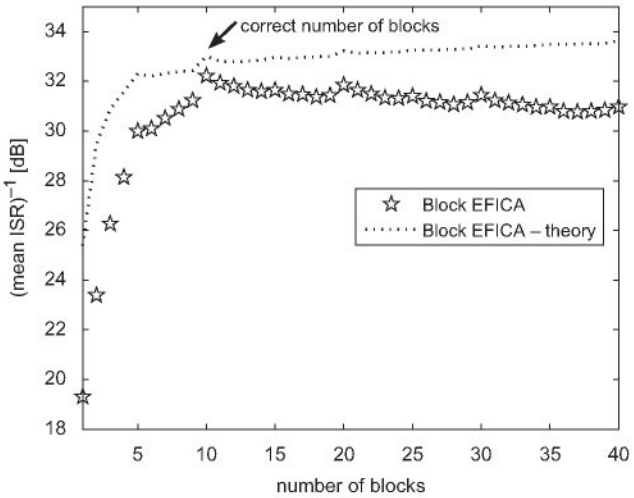


Fig. 4. Average interference-to-signal ratio of 10 sub-Gaussian signals achieved by Block EFICA when changing the number of blocks M considered by the algorithm.

performance of the NSNG algorithm here, because its original implementation[4] is designed for sub-Gaussian signals only, and the method fails to converge in this experiment.

To test a scenario with sub-Gaussian signals, we show in Fig. 3 the performance achieved by separation of 10 signals composed of $M = 10$ blocks. The k-th signal, $k = 1, \ldots, 10$, is uniformly distributed (with variance one) in the first k blocks and Gaussian elsewhere.

Similarly to the previous experiment, this example demonstrates the strongest point of the Block EFICA, which consists in its ability to adapt to varying signal distribution. The same performance was achieved by the NSNG algorithm, and it performed yet better when smaller length of data was considered, which is likely thanks to lower number of parameters compared to Block EFICA. However, also in this scenario, NSNG failed to converge in a few trials. To allow presentation of its performance, the trials where the divergence occurred had to be skipped.

Fig. 4 shows the overall performance averaged over all sources when changing the input parameter M in Block EFICA from 1 to 40. Although performance is optimum for the correct value of $M = 10$, the deterioration of the performance due to overestimation or underestimation of M is not high. For M close to 1 the performance of Block EFICA approaches that of EFICA, which is as expected. Certain local maxima can be observed for M being multiple of 10, which is thanks to fitting the boundaries of blocks exactly to the instants where the distributions of signals are switched. Nevertheless, the negligible improvement demonstrates lower importance of the correct fitting.

The theoretical performance computed using (29) monotonically grows with M. It therefore becomes slightly overoptimistic for higher values of M, because it does not take the practical effect of overparametrization into account. Nevertheless, it may be used in order to choose an effective value of M.

[4] The implementation of the NSNG algorithm was obtained from web-site http://www-lmc.imag.fr/SMS/SASI/bliss.html.

6.2. Signals with changing variance

Since the proposed Block EFICA exploits the piecewise stationary modelling concept, we test its ability to separate nonstationary signals with varying variance. For that purpose, we design a simple experiment where a signal having variable variance is separated from another signal that is stationary. The first (nonstationary) signal has variances, respectively, equal to 1, σ, and σ^2 in the three consecutive blocks of the same length, and the second signal is Gaussian having the constant variance equal to one. An example of the signals for a particular value of the parameter σ, which is considered on interval $(0, 1]$, is shown in Fig. 5.

We consider two situations that differ in selected distribution of the first nonstationary signal. In the first setup, the distribution is Gaussian in all blocks. Then, for σ close to one, where the two signals are almost stationary, the mixture cannot be separated due to Gaussianity of the signals. In the second setup, the distribution is Laplacian, which makes the mixture separable even for σ close to one. The signals can be separated for both cases when σ is close to zero. Then, the first signal is strongly nonstationary and has a different variance-envelope than the second signal, which is the general requirement of the BGL algorithm. Fig. 6 shows results obtained for both settings of the experiment.

The first scenario with Gaussian signals fits the Block Gaussian model. In such a case, the theoretical performance of the BGL algorithm attains corresponding Cramér–Rao bound, here, given by $\mathrm{CRLB}_3 = \mathrm{CRLB}_4$. Therefore, its performance should be optimal, which is confirmed by the results shown by Figs. 6(a) and (b). Similar performance was achieved by the NSNG algorithm without yielding any instability, which reveals its excellent ability to utilize the nonstationarity of signals in simple examples such as the two-dimensional one considered here.

The proposed Block EFICA algorithm achieves comparable results up to $\sigma \in [0.7, 1]$, where the Gaussian signals

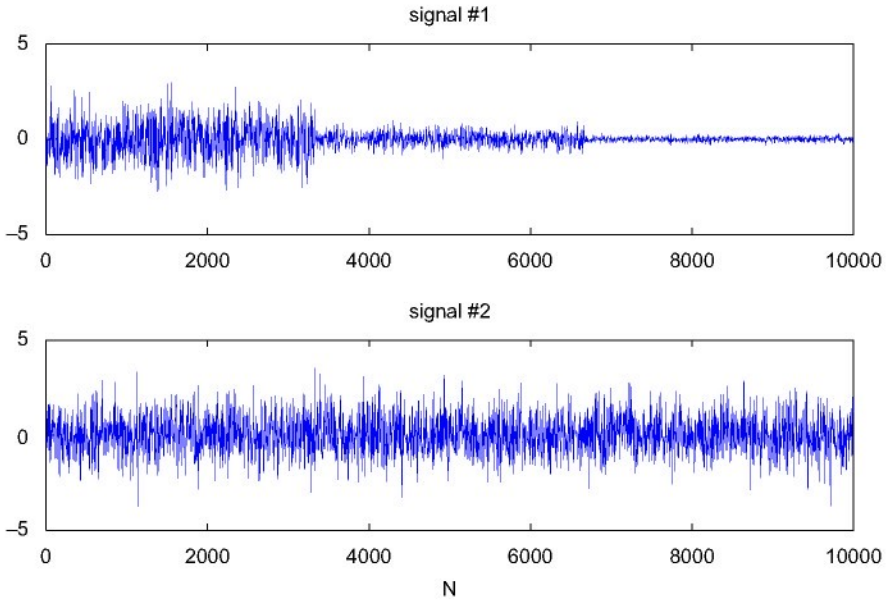


Fig. 5. Illustration of the Gaussian signals of length $N = 10^4$ when the parameter σ that controls the nonstationarity of the first signal equals 0.1.

are almost stationary, which makes them hardly distinguishable for non-Gaussianity-based methods. Hence, the breakdown of the performance is caused by failures of the initialization provided by the symmetric FastICA in the first step BEF1. In our experiments not shown here due to space, we observed that if "good" initialization is guaranteed, the final performance of Block EFICA is comparable with that of the BGL algorithm. Therefore, Block EFICA may be initialized by another method that performs well in this particular case. Nevertheless, our selection, the symmetric FastICA with the test of saddle points, appears to be suitable for most applications as discussed in Section 4.3.

The plots marked by "Block EFICA (identity)" demonstrate further improvement of Block EFICA done via involving the identity function in the score function estimator (see Section 4.4). The better performance shows that the option allows a more effective exploitation of nonstationarity of signals.

The second scenario simulates the case when the original signals exhibit both the non-Gaussianity and the nonstationarity since the distribution of the first signal is Laplacian. Here, the Block EFICA yields performance that is superior to the other methods. The BGL algorithm suffers from stationarity of the signals as σ is approaching one. Conversely, the original EFICA does not utilize effectively their nonstationarity for σ close to zero. The implementation of the NSNG algorithm lacks the ability to accurately estimate the score function of the Laplacian distribution. It has significantly lower performance than EFICA and Block EFICA, nevertheless, its ability to profit both from nonstationarity and non-Gaussianity is confirmed.

6.3. Separation of noisy instantaneous mixtures of speech signals

In this example, we compare performances of algorithms in a noisy scenario. Fig. 7 shows results of separation of 10 speech signals randomly selected from a database, each of length 5000 samples. The signals were mixed by a random matrix, Gaussian noise was added to each mixed channel with the variance corresponding to a given signal-to-noise ratio (input SNR), and the mixture was separated and evaluated in terms of signal-to-interference-plus-noise ratio (SINR). The experiment was designed according to the rules proposed in [39].

Since speech signals often exhibit, beside nonstationarity and non-Gaussianity also spectral diversity, we compare the performance of Block EFICA with the SOBI-RO algorithm from [40] that utilizes the spectral diversity, and ThinICA [11] using also their non-Gaussianity. As can be seen from the results, Block EFICA is not sensitive to the additive noise as inherited from EFICA and FastICA. The achieved SINR decreases smoothly with input SNR. In our example, Block EFICA outperforms the compared algorithms; however, note that the performance strongly depends on properties of the to-be-separated signals.

6.4. Separation of natural convolutive mixture of speech signals

To demonstrate strengths of Block EFICA on real-world data, we present an example where a convolutive mixture of two speech signals recorded by two microphones is separated. The mixture is separated using the procedure from [38] as follows.[5] The first and the most important stage relies on an ICA decomposition of a subspace spanned by delayed signals from microphones, i.e.,

$$x_1(n), x_1(n-1), \ldots, x_1(n-L+1), \\ x_2(n), x_2(n-1), \ldots, x_d(n-L+1), \tag{33}$$

[5] The method from [38] is available at http://itakura.ite.tul.cz/zbynek/tddeconv.htm.

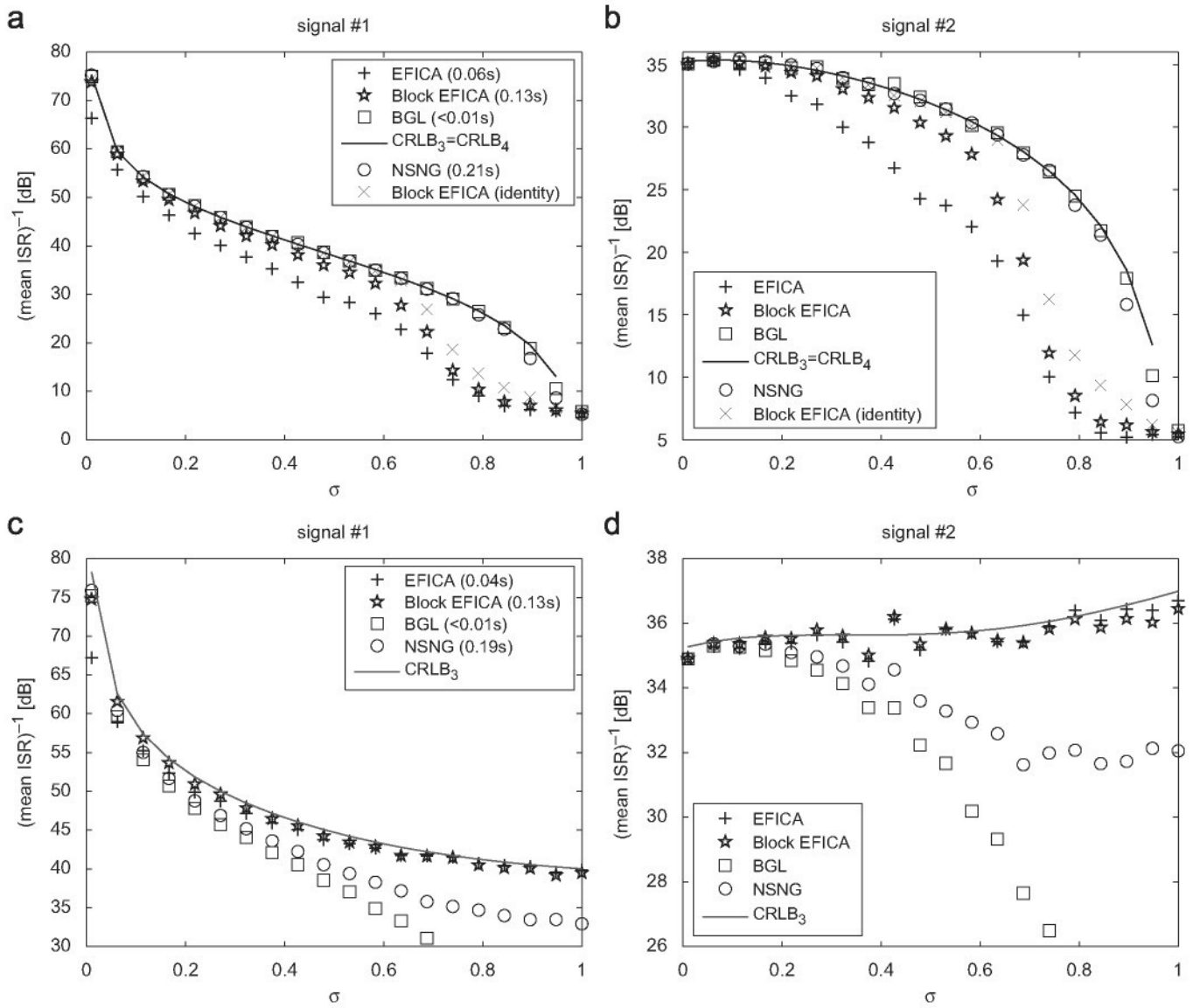


Fig. 6. Results of the experiment with nonstationary signals evaluated by the mean interference-to-signal ratio that was computed from results of 1000 Monte Carlo trials done for each value of the parameter σ. The figures (a) and (b) correspond to the first scenario, where the distribution of signal #1 is Gaussian, while for (c) and (d) the distribution is Laplacian.

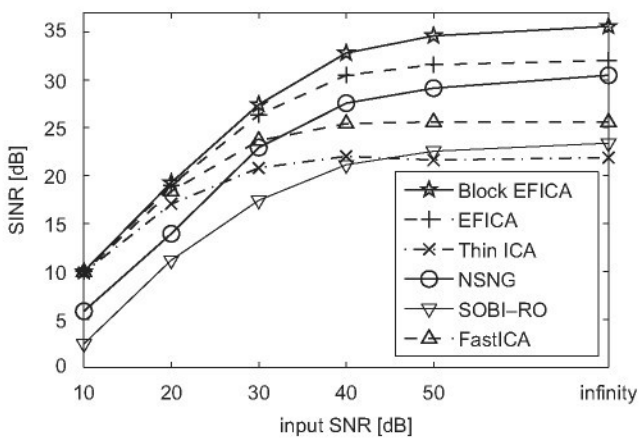


Fig. 7. Results of separation of noisy mixtures of speech signals averaged over 100 independent trials.

where L is the length of separating filters. Note that this way the convolutive mixture problem is transformed into an instantaneous one, thus, we can apply any ICA algorithm that is originally designed for instantaneous mixtures (including Block EFICA). The algorithm thus yields independent components of the subspace (33) that, in fact, correspond to outputs of $d \cdot L$ multiple-input single-output filters of length L. The key objective is that each independent component should contain a contribution of one original source only, which is, in an ideal case, a filtered copy of the source [41–43].

The procedure from [38] continues by grouping the components into clusters that correspond to the same original source. Finally, the clusters (the components in the clusters) are used to reconstruct the original sources; see [38] for further details. Anyway, the idea of this experiment comes from the fact that the final results of

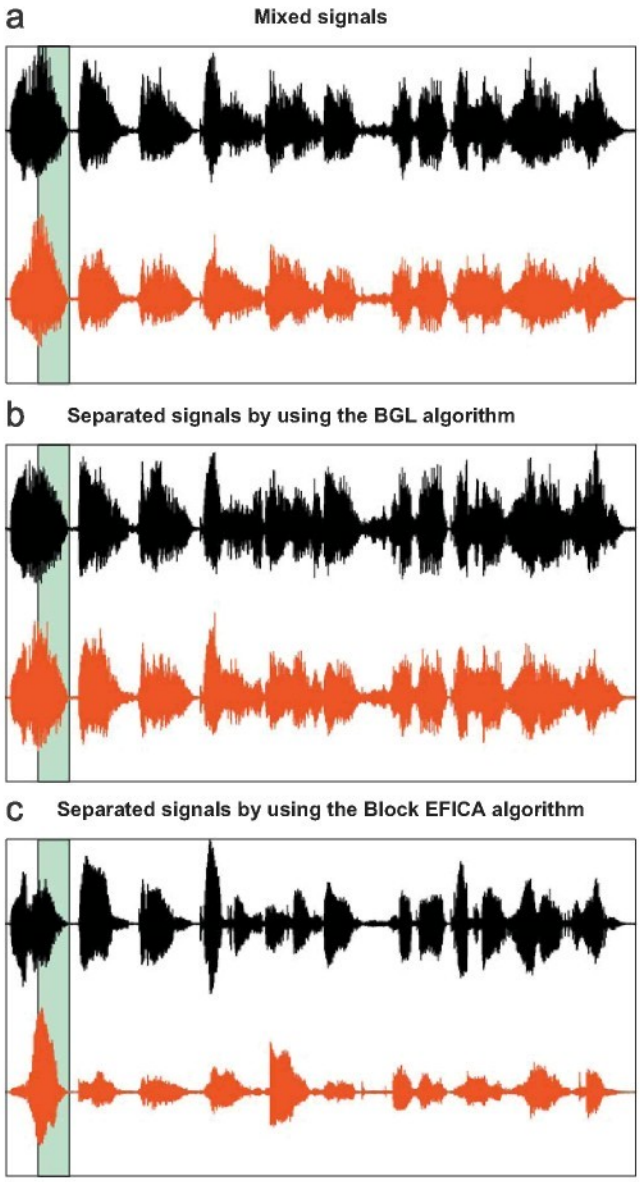


Fig. 8. Results of separation of real-world convolutive mixture of two speech signals recorded by two microphones. Respective ICA methods were applied to the subspace generated by selected data segment of 6000 samples. The segment is delimited by vertical lines in the graphs. (a) Mixed signals. (b) Separated signals by using the BGL algorithm. (c) Separated signals by using Block EFICA algorithm.

the separation provide a benchmark for testing ability of different ICA methods for instantaneous mixtures to separate convolutive audio mixtures, i.e., to yield such independent components that correspond to particular original sources.

Fig. 8(a) shows Lee's data[6] containing real recordings of two speakers (played over loudspeakers) simultaneously saying the digits from one to 10 in English and in Spanish, respectively. The loudspeakers were placed closely to the microphones (60 cm), so direct-path signals and possibly early reflections from the closest objects are much stronger than the other reverberations in the recorded convolutive mixture. Hence, very short separating filters applied through [38] (of the length L) may separate these signals efficiently.

Since the rhythms of the speech signals are similar and synchronized, there occur many short segments (say of length 6000 samples—the sampling frequency is 16 kHz) where the dynamics of the speech signals are very close. Owing to possible changing mixing conditions (e.g., moving sources), the aim is to separate as short segments of signals as possible. However, the similar dynamics of sources in short segments cause malfunctioning of nonstationarity-based methods. From this point of view, the methods that use not only the nonstationarity but also the non-Gaussianity of speech are more flexible, because they do not fail in such situations.

To demonstrate this, Figs. 8(b) and (c) show results of separation with $L = 20$ via BGL[7] and Block EFICA, respectively, when only using a short segment of data for the mixture identification (learning data). Then, the resulting separating filters are applied to the whole signals. Since the mixture is here stationary (the loudspeakers and microphones remain in their positions during the whole recording), the separated signals reveal ability of the ICA methods to separate them using data from the given data segment only.

Since the dynamics of signals are too similar in the chosen segment, the nonstationarity-based BGL algorithm yields poorly separated components of (33), so that average SIR of the finally separated sources is 3.3 dB,[8] while the original SIR of the mixed signals is 3.4 dB. By contrast, the Block EFICA algorithm succeeded to separate the signals yielding average SIR of 12.2 dB, which means "good" result in this convolutive audio source separation task.

7. Conclusions

We have proposed the Block EFICA algorithm that effectively exploits both the non-Gaussianity and the nonstationarity of original signals to separate them. The method efficiently solves the ICA task defined by the piecewise stationary model. It yields comparable performance as methods only intended for marginal cases: the non-Gaussianity-based model or the Block Gaussian model. Namely, it has about the same performance as the EFICA algorithm if the separated signals are stationary and non-Gaussian. In case of Gaussian piecewise stationary signals, Block EFICA is not claimed to be optimum in theory, but in our simulations we have shown that its performance may be close to that of the BGL algorithm that performs optimally in this case.

In so doing, Block EFICA performs best in case of compound scenarios involving non-Gaussian and

[6] Lee's data are available online at http://www.cnl.salk.edu/~tewon/Blind/blind_audio.html.

[7] In fact, the method from [38] utilizes a fast variant of BGL named BG_WEDGE. The algorithm is based on a fast joint diagonalization algorithm with adaptive weights proposed in [17].

[8] The signal-to-interference ratio was evaluated by means of the BSS_EVAL toolbox from [46] that uses projections of signals to avoid indeterminacies due to arbitrary filtering of separated signals. Lee's separated signals were used as the reference "correct" signals.

nonstationary signals. The considered number of blocks M need not be precisely determined, as the method is not highly sensitive to it. Moreover, it yields equivalent performance with that of EFICA when M is equal to one. Finally, Block EFICA provides an appealing alternative to the theoretically optimum NSNG algorithm in terms of better stability and lower computational complexity, especially, when applied to high-dimensional data and, therefore, may be successfully applied to real-world BSS problems.

Acknowledgments

This work was supported by Ministry of Education, Youth and Sports of the Czech Republic through the project 1M0572 and by Grant Agency of the Czech Republic through the projects 102/09/1278 and 102/07/P384.

Appendix A. Derivation of CRLB$_3$

In this appendix, we provide a simple derivation of CRLB$_3$ based on results from [32] and the corrections [44].

We start from Eq. (36) of [32] that, for $N=1$, gives the mn-th element of the Fisher information matrix[9] of an independent observation of (2)

$$(\mathbf{F_I})_{mn} = \delta_{ju}\delta_{vi} + \delta_{ji}\delta_{vu}\delta_{vi}(\eta_i - \kappa_i - 2) + \delta_{iu}\delta_{vj}\kappa_i, \tag{34}$$

where $m=(i-1)d+j$ and $n=(u-1)d+v$ with $i,j,u,v=1,\dots,d$, $\eta_i = \mathrm{E}[s_i^2\psi_i^2(s_i)]$, and δ_{ji} is Kronecker's delta. This result can be easily extended for signals with general variance $\sigma_j^2 = \mathrm{E}[s_j^2]$ ([32, p. 1201], the first column, the fourth line of the second item in the enumeration), which gives

$$(\mathbf{F_I})_{mn} = \delta_{ju}\delta_{vi} + \delta_{ji}\delta_{vu}\delta_{vi}\left(\eta_i - \frac{\sigma_v^2}{\sigma_i^2}\kappa_i - 2\right) + \delta_{iu}\delta_{vj}\frac{\sigma_v^2}{\sigma_i^2}\kappa_i, \tag{35}$$

where κ_i and η_i are defined for *normalized* pdfs of the sources in order to be scale-invariant.

Now it follows that the FIM of an observation from the I-th block of the piecewise stationary model (3) should have the block-dependent quantities labelled by the superscript (I), and the FIM of all N independent observations has the mn-th element equal to

$$(\mathbf{F_I})_{mn} = N\left[\delta_{ju}\delta_{vi} + \delta_{ji}\delta_{vu}\delta_{vi}\frac{1}{M}\sum_{I=1}^{M}\left(\eta_i^{(I)} - \frac{\sigma_v^{2(I)}}{\sigma_i^{2(I)}}\kappa_i^{(I)} - 2\right) + \delta_{iu}\delta_{vj}\frac{1}{M}\sum_{I=1}^{M}\frac{\sigma_v^{2(I)}}{\sigma_i^{2(I)}}\kappa_i^{(I)}\right]. \tag{36}$$

The structure of the FIM (36) is the same as in case of the basic ICA model (2), i.e., it can be written in a form $\mathbf{F_I} = \mathbf{P} + \mathbf{\Sigma}$ with $\mathbf{P}$ being a special permutation matrix and $\mathbf{\Sigma}$ being diagonal

$$\mathbf{\Sigma} = \frac{1}{M}\sum_{I=1}^{M}\mathrm{diag}\left[\eta_1^{(I)} - 2, \frac{\sigma_2^{2(I)}}{\sigma_1^{2(I)}}\kappa_1^{(I)}, \dots, \frac{\sigma_d^{2(I)}}{\sigma_1^{2(I)}}\kappa_1^{(I)}, \frac{\sigma_1^{2(I)}}{\sigma_2^{2(I)}}\kappa_2^{(I)}, \eta_2^{(I)} - 2, \frac{\sigma_3^{2(I)}}{\sigma_2^{2(I)}}\kappa_2^{(I)}, \dots, \frac{\sigma_d^{2(I)}}{\sigma_2^{2(I)}}\kappa_2^{(I)}, \dots, \frac{\sigma_1^{2(I)}}{\sigma_d^{2(I)}}\kappa_d^{(I)}, \dots, \frac{\sigma_{d-1}^{2(I)}}{\sigma_d^{2(I)}}\kappa_d^{(I)}, \eta_d^{(I)} - 2\right]. \tag{37}$$

Therefore, the inversion of $\mathbf{F_I}$ can be derived using Appendix D of [32]; see the simplification due to the corrections. Using appropriate substitutions according to (90) in [32], the resulting CRLB$_3$ given by (8) readily follows.

Appendix B. Proof of Proposition 1

We will follow the easiest way by generalizing proof of analogous proposition in [32] (see the Appendix A therefrom). Similar notations will be used, namely, $\mathbf{s}_k$ will be $N\times 1$ vector of samples of the k-th original signal, i.e., the k-th row of $\mathbf{S}$, with the difference that the I-th block of N/M samples is distributed according to RV $s_k^{(I)}$. Owing to the indeterminacy of scale of original signals, the variances of $s_k^{(I)}$ can be assumed to be such that $\mathbf{s}_k$ has unit scale (assumption (i) of the proposition).

Next, the vector $\mathbf{u}_k$ contains normalized elements of $\mathbf{s}_k$ so that $\mathbf{u}_k$ has the second-order sample-moment exactly equal to one. The vectors $\mathbf{z}_k$ and $\mathbf{x}_k$ denote samples of the respective signals. The nonlinearity $g(\cdot)$ used for the k-th signal will be distinguished by the subscript k, i.e., $g_k(\cdot)$. It applies to the vectors elementwise, so that function $\lambda_k^{(I)}g_k^{(I)}(\cdot)$ applies to the I-th block of N/M elements.

Now, using the third assumption of Proposition 1 given by (21), Eqs. (40) and (41) from [32] change, respectively, to

$$N^{-1}\mathbf{s}_k^T g_k(\mathbf{s}_k) \overset{N\to+\infty}{\longrightarrow} \overline{\mu}_k, \tag{38}$$

$$N^{-1}g_k'^T(\mathbf{s}_k)\mathbf{1}_N \overset{N\to+\infty}{\longrightarrow} \overline{\nu}_k. \tag{39}$$

Note that ν denotes the same expectations that are in [32] denoted by ρ. $\mathbf{1}_N$ stands for $N\times 1$ vector of ones.

Using this, all Eqs. (42)–(64) in [32] change according to the substitutions

$$\mu_k \leftarrow \overline{\mu}_k, \tag{40}$$

$$\rho_k \leftarrow \overline{\nu}_k. \tag{41}$$

The only exceptions are Eqs. (42), (45), and (62), which should be revised due to different variance in blocks, and it gives, respectively,

$$N^{-1}g_k'^T(\mathbf{s}_k)(\mathbf{s}_\ell \odot \mathbf{s}_\ell) \overset{N\to+\infty}{\longrightarrow} \overline{\nu}_{k\ell}, \tag{42}$$

$$\mathbf{g}_k'^T(\mathbf{u}_\ell \odot \mathbf{u}_\ell) = N\overline{\nu}_{k\ell} + o_p(N), \tag{43}$$

$$\mathrm{E}[(\mathbf{g}_k^T\mathbf{u}_\ell)^2] = N\overline{\beta}_{k\ell}, \tag{44}$$

where $\mathbf{g}_k$ is the simplified notation of $g_k(\mathbf{u}_k)$. Recomputation of (65), (71), and (75) in [32] using the above substitutions readily yields the result of the proposition given by (22).

[9] In the corrections [44], it is shown that the first term in (36) of [32] should be removed. This means that, for $N=1$, the relation is correct.

Appendix C. Proof of Proposition 2

The criterion (26) can be written in the form

$$V_{k\ell}^{\mathrm{B1U}} = \frac{\mathbb{I}_k^T \boldsymbol{\Gamma}_k \mathbb{I}_k}{\mathbb{I}_k^T \mathbb{t}_k \mathbb{t}_k^T \mathbb{I}_k} \tag{45}$$

with

$$\mathbb{I}_k = [\lambda_k^{(1)}, \ldots, \lambda_k^{(M)}]^T, \tag{46}$$

$$\boldsymbol{\Gamma}_k = \mathrm{diag}[\beta_k^{(1)}, \ldots, \beta_k^{(M)}] - \frac{1}{M} \mathbb{m}_k \mathbb{m}_k^T, \tag{47}$$

$$\mathbb{m}_k = [\mu_k^{(1)}, \ldots, \mu_k^{(M)}]^T, \tag{48}$$

$$\mathbb{t}_k = [\tau_k^{(1)}, \ldots, \tau_k^{(M)}]^T, \tag{49}$$

$$\tau_k^{(l)} = v_k^{(l)} - \mu_k^{(l)}. \tag{50}$$

The goal is to minimize (45) subject to elements of $\mathbb{I}_k$, which is equivalent with maximizing

$$\max_{\mathbb{I}_k} \frac{\mathbb{I}_k^T \mathbb{t}_k \mathbb{t}_k^T \mathbb{I}_k}{\mathbb{I}_k^T \boldsymbol{\Gamma}_k \mathbb{I}_k}. \tag{51}$$

Let $\mathbb{y}_k = \boldsymbol{\Gamma}_k^{1/2} \mathbb{I}_k$, where the matrix $\boldsymbol{\Gamma}_k^{1/2}$ obeying $\boldsymbol{\Gamma}_k^{1/2} \boldsymbol{\Gamma}_k^{1/2} = \boldsymbol{\Gamma}_k$ exists thanks to positive semidefiniteness of $\boldsymbol{\Gamma}_k$ ($V_{k\ell}^{\mathrm{B1U}}$ denotes variance, which must be always nonnegative). Since (45) is invariant subject to nonzero multiple of $\mathbb{I}_k$, we can introduce a constraint $\|\mathbb{I}_k\| = \text{const.}$, and (51) can be written in the form of classical eigenvalue problem

$$\max_{\|\mathbb{y}_k\|=1} \frac{\mathbb{y}_k^T \boldsymbol{\Gamma}_k^{-1/2} \mathbb{t}_k \mathbb{t}_k^T \boldsymbol{\Gamma}_k^{-1/2} \mathbb{y}_k}{\mathbb{y}_k^T \mathbb{y}_k}. \tag{52}$$

The rank of the matrix $\boldsymbol{\Gamma}_k^{-1/2} \mathbb{t}_k \mathbb{t}_k^T \boldsymbol{\Gamma}_k^{-1/2}$ is one, thus, the eigenvector corresponding to the only nonzero eigenvalue, i.e., the solution of (52), is $\mathbb{y}_k = \boldsymbol{\Gamma}_k^{-1/2} \mathbb{t}_k$. Hence, $\mathbb{I}_k$ that minimizes (45) is

$$\mathbb{I}_k = \boldsymbol{\Gamma}_k^{-1} \mathbb{t}_k. \tag{53}$$

Using the matrix inversion lemma for computation of $\boldsymbol{\Gamma}_k^{-1}$, (27) follows.

References

[1] A. Cichocki, S.-I. Amari, Adaptive Signal and Image Processing: Learning Algorithms and Applications, Wiley, New York, 2002.

[2] A. Hyvärinen, J. Karhunen, E. Oja, Independent Component Analysis, Wiley-Interscience, New York, 2001.

[3] P. Comon, Independent Component Analysis, a new concept?, Signal Processing 36 (3) (April 1994) 287–314.

[4] J.-F. Cardoso, A. Souloumiac, Blind beamforming from non-Gaussian Signals, IEE Proceedings—F 140 (6) (December 1993) 362–370.

[5] A. Hyvärinen, Fast and robust fixed-point algorithms for independent component analysis, IEEE Transactions on Neural Networks 10 (3) (1999) 626–634.

[6] A.J. Bell, T.J. Sejnowski, An information-maximization approach to blind separation and blind deconvolution, Neural Computation 7 (1995) 1129–1159.

[7] K. Matsuoka, M. Ohya, M. Kawamoto, A neural net for blind separation of nonstationary signals, Neural Networks 8 (3) (May 1995) 411–419.

[8] S. Choi, A. Cichocki, S. Amari, Equivariant nonstationary source separation, Neural Networks 15 (1) (January 2002) 121–130.

[9] J.-I. Hirayama, S.-I. Maeda, S. Ishii, Markov and semi-Markov switching of source appearances for nonstationary independent component analysis, IEEE Transactions on Neural Networks 18 (5) (September 2007).

[10] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, E. Moulines, A blind source separation technique using second-order statistics, IEEE Transactions on Signal Processing 45 (February 1997) 434–444.

[11] S. Cruces, A. Cichocki, Combining blind source extraction with joint approximate diagonalization: thin algorithms for ICA, in: Proceedings of the 4th International Symposium on Independent Component Analysis and Blind Signal Separation (ICA2003), Nara, Japan, April 2003, pp. 463–468.

[12] T.-W. Lee, M. Girolami, T.J. Sejnowski, Independent component analysis using an extended infomax algorithm for mixed sub-Gaussian and super-Gaussian sources, Neural Computation 11 (2) (1999) 417–441.

[13] D.T. Pham, P. Garat, Blind separation of mixture of independent sources through a quasi-maximum likelihood approach, IEEE Transactions on Signal Processing 45 (7) (July 1997) 1712–1725.

[14] D.-T. Pham, J.-F. Cardoso, Blind separation of instantaneous mixtures of non stationary sources, IEEE Transactions on Signal Processing 49 (9) (2001) 1837–1848.

[15] R. Boscolo, H. Pan, V.P. Roychowdhury, Independent component analysis based on nonparametric density estimation, IEEE Transactions on Neural Networks 15 (1) (2004) 55–65.

[16] Z. Koldovský, P. Tichavský, E. Oja, Efficient variant of algorithm FastICA for independent component analysis attaining the Cramér–Rao lower bound, IEEE Transactions on Neural Networks 17 (5) (September 2006).

[17] P. Tichavský, A. Yeredor, Fast approximate joint diagonalization incorporating weight matrices, IEEE Transactions on Signal Processing 57 (3) (March 2009) 878–891 codes available at ⟨http://si.utia.cas.cz/downloadPT.htm⟩.

[18] L. Parra, P. Sajda, Blind source separation via generalized eigenvalue decomposition, Journal of Machine Learning Research 4 (2003) 1261–1269.

[19] J.-F. Cardoso, A. Souloumiac, Jacobi angles for simultaneous diagonalization, SIAM Journal on Matrix Analysis and Applications 17 (1996) 161–164.

[20] F. Meinecke, A. Ziehe, M. Kawanabe, K.R. Müller, A resampling approach to estimate the stability of one-dimensional or multi-dimensional independent components, IEEE Transactions on Biomedical Engineering 49 (12) (December 2002) 1514–1525.

[21] H. Buchner, R. Aichner, W. Kellermann, TRINICON: a versatile framework for multichannel blind signal processing, Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Montreal, Canada 3 (May 2004) 889–892.

[22] K.R. Müller, P. Philips, A. Ziehe, JADETD: combining higher-order statistics and temporal information for blind source separation (with noise), in: Proceedings of ICA '99, Aussois, January 1999.

[23] S. Choi, A. Cichocki, A. Belouchrani, Second order nonstationary source separation, Journal of VLSI Signal Processing 32 (1–2) (August 2002) 93–104.

[24] A. Hyvärinen, A unifying model for blind separation of independent sources, Signal Processing 85 (7) (2005) 1419–1427.

[25] P. Tichavský, Z. Koldovský, A. Yeredor, G.G. Herrero, E. Doron, A hybrid technique for blind non-Gaussian and time-correlated sources using a multicomponent approach, IEEE Transactions on Neural Networks 19 (3) (March 2008) 421–430.

[26] J.-F. Cardoso, D.T. Pham, Separation of non stationary sources. Algorithms and performance, in: S.J. Roberts, R.M. Everson (Eds.), Independent Components Analysis: Principles and Practice, Cambridge University Press, Cambridge, 2001, pp. 158–180.

[27] D.-T. Pham, Blind separation of non stationary non Gaussian sources, in: Proceeding of the EUSIPCO 2002 Conference, Toulouse, France, September 2002.

[28] Z. Koldovský, J. Málek, P. Tichavský, Y. Deville, S. Hosseini, Extension of EFICA algorithm for blind separation of piecewise stationary non Gaussian sources, in: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Las Vegas, Nevada, USA, April 2008, pp. 1913–1916.

[29] R.C. Rao, Linear Statistical Inference and its Applications, second ed., Wiley, New York, 1973.

[30] O. Shalvi, E. Weinstein, Maximum likelihood and lower bounds in system identification with non-Gaussian inputs, IEEE Transactions on Information Theory 40 (2) (March 1994) 328–339.

[31] J.-F. Cardoso, Blind signal separation: statistical principles, Proceedings of the IEEE 90 (8) (October 1998) 2009–2026.

[32] P. Tichavský, Z. Koldovský, E. Oja, Performance analysis of the FastICA algorithm and Cramér–Rao bounds for linear independent

component analysis, IEEE Transactions on Signal Processing 54 (4) (April 2006).
[33] Z. Koldovský, J. Málek, P. Tichavský, Y. Deville, S. Hosseini, Performance Analysis of Extended EFICA Algorithm, Technical Report No. 2199, ÚTIA, AV ČR, 2007. Available at ⟨http://itakura.ite.tul.cz/zbynek/publications.htm⟩.
[34] B.W. Silvermann, Density Estimation for Statistics and Data Analysis, Chapman & Hall, London, 1986.
[35] J. Málek, Z. Koldovský, S. Hosseini, Y. Deville, A variant of EFICA algorithm with adaptive parametric density estimator, in: 8th International Workshop on Electronics, Control, Modelling, Measurement, and Signals (ECMS 2007), Liberec, Czech Republic, May 2007, pp. 79–84.
[36] Y. Deville, M. Benali, F. Abrard, Differential source separation for underdetermined instantaneous or convolutive mixtures: concept and algorithms, Signal Processing 84 (10) (October 2004) 1759–1776.
[37] P. Tichavský, Z. Koldovský, E. Oja, Speed and accuracy enhancement of linear ICA techniques using rational nonlinear functions, in: Proceedings of ICA2007, September 2007, pp. 285–292.
[38] Z. Koldovský, P. Tichavský, Time-domain blind audio source separation using advanced component clustering and reconstruction, in: Proceedings of Joint Workshop on Hands-free Speech Communication and Microphone Arrays (HSCMA 2008), Trento, Italy, May 2008, pp. 216–219.
[39] Z. Koldovský, P. Tichavský, Methods of fair comparison of performance of linear ICA techniques in presence of additive noise, in: Proceedings of ICASSP 2006, Toulouse, no. V, May 2006, pp. 873–876.
[40] A. Belouchrani, A. Cichocki, Robust whitening procedure in blind source separation context, Electronics Letters 36 (24) (2000) 2050–2053.
[41] H. Buchner, R. Aichner, W. Kellermann, A generalization of blind source separation algorithms for convolutive mixtures based on second-order statistics, IEEE Transactions on Speech and Audio Processing 13 (1) (January 2005) 120–134.
[42] S. Douglas, M. Gupta, H. Sawada, S. Makino, Spatio-temporal FastICA algorithms for the blind separation of convolutive mixtures, IEEE Transactions on Audio, Speech and Language Processing 15 (5) (July 2007) 1511–1520.
[43] J. Thomas, Y. Deville, S. Hosseini, Time-domain fast fixed-point algorithms for convolutive ICA, IEEE Signal Processing Letters 13 (4) (2006) 228–231.
[44] P. Tichavský, Z. Koldovský, E. Oja, Corrections of the 'performance analysis of the FastICA algorithm and Cramér–Rao bounds for linear independent component analysis, IEEE Transactions on Signal Processing 56 (4) (April 2008) 1715–1716.
[45] [online] The Matlab™ implementation of the Block EFICA algorithm: ⟨http://itakura.ite.tul.cz/zbynek/eefica.htm⟩.
[46] E. Vincent, R. Gribonval, C. Févotte, Performance measurement in blind audio source separation, IEEE Transactions on Speech and Audio Processing 14 (4) (July 2006) 1462–1469.

A Hybrid Technique for Blind Separation of Non-Gaussian and Time-Correlated Sources Using a Multicomponent Approach

Petr Tichavský, *Senior Member, IEEE*, Zbyněk Koldovský, *Member, IEEE*, Arie Yeredor, *Senior Member, IEEE*, Germán Gómez-Herrero, and Eran Doron

***Abstract*—Blind inversion of a linear and instantaneous mixture of source signals is a problem often encountered in many signal processing applications. Efficient fastICA (EFICA) offers an asymptotically optimal solution to this problem when all of the sources obey a generalized Gaussian distribution, at most one of them is Gaussian, and each is independent and identically distributed (i.i.d.) in time. Likewise, weights-adjusted second-order blind identification (WASOBI) is asymptotically optimal when all the sources are Gaussian and can be modeled as autoregressive (AR) processes with distinct spectra. Nevertheless, real-life mixtures are likely to contain both Gaussian AR and non-Gaussian i.i.d. sources, rendering WASOBI and EFICA severely suboptimal. In this paper, we propose a novel scheme for combining the strengths of EFICA and WASOBI in order to deal with such hybrid mixtures. Simulations show that our approach outperforms competing algorithms designed for separating similar mixtures.**

***Index Terms*—Blind source separation, independent component analysis (ICA).**

I. Introduction

In this paper, we address the classical real-valued square (invertible) instantaneous linear independent components analysis (ICA) model $\mathbf{x} = \mathbf{A}\mathbf{s}$, where $\mathbf{s}, \mathbf{x} \in \mathbb{R}^{d \times N}$ contain the d unknown independent source signals and their observed mixtures (respectively), each of length N, and $\mathbf{A} \in \mathbf{R}^{d \times d}$ is the unknown mixing matrix.

The goal is to estimate the mixing matrix $\mathbf{A}$ or, equivalently, the demixing matrix $\mathbf{W} \triangleq \mathbf{A}^{-1}$ or, equivalently, the original source signals $\mathbf{s}$. We employ an assumption of zero-mean unit variance sources, and we assume for simplicity of the exposition that the remaining permutation ambiguity can be arbitrated (e.g., using the reordering method proposed in [25], which is also used in our simulations).

At least three classes of source models have been considered in the literature (see, e.g., [5]) with associated separation approaches based on either "non-Gaussianity," "nonwhiteness," or "nonstationarity" of the source signals.[1] For each of these models, there exist algorithms which are asymptotically optimal (in some sense, to be discussed in Section III) under the following conditions: 1) efficient fastICA (EFICA, [18]) for independent white generalized-Gaussian-distributed sources, 2) weights-adjusted second-order blind identification (WASOBI, [33], [9], [28]) for wide sense stationary (WSS) parametric Gaussian sources with *spectral diversity*, and 3) block Gaussian likelihood (BGL, [22]) for Gaussian sources with *time-varying variances*. Note that EFICA is a recently developed modification of the popular fastICA [13]. A speed enhancement of fastICA/EFICA using rational nonlinear functions (used in this paper) was proposed in [30]. The WASOBI is an enhanced version of the popular algorithm second-order blind identification (SOBI) [2].

Often in cases of real-data processing, no single model of these three classes offers a correct representation of all sources. For example, in biomedicine, both non-Gaussianity-based and spectral diversity-based blind separation methods are currently studied; see [16] and [27]. Merits in combining these two kinds of methods were already demonstrated on an example with an electroencephalography (EEG) data in [12].

The aim of this paper is to develop a method that can account for a combination of the first two model classes, by combining the strengths of EFICA and WASOBI. There is no claim of inherited asymptotic optimality of the resulting algorithm. However, simulations show that our approach outperforms previous attempts to address combinations of those two source classes, namely, the algorithms joint approximate diagonalization of eigenmatrices (JADE_{TD}) [23], joint cumulant and correlation-based separation (JCC) [12], and thinICA (TICA, [6]). Another, *ad hoc* algorithm, addressing combinations of *all three* classes, was proposed by Hyvärinen in [15], an extension of a complexity pursuit algorithm [14]. Unfortunately, however, this algorithm was not developed in sufficient generality. In particular, the implementation that is available so far is only suitable

Manuscript received January 9, 2007; revised June 12, 2007; accepted July 18, 2007. Parts of this work were presented at the 2006 European Signal Processing Conference (EUSIPCO). This work was supported by the Ministry of Education, Youth and Sports of the Czech Republic under Project 1M0572, the Grant Agency of the Czech Republic under Projects 201/08/0646 and 102/07/P384, and the Academy of Finland under Project 213462 (Finnish Centre of Excellence Program 2006-2011).

P. Tichavský and Z. Koldovský are with the Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, 182 08 Prague 8, Czech Republic (e-mail: tichavsk@utia.cas.cz; zbynek.koldovsky@tul.cz).

A. Yeredor is with the School of Electrical Engineering, Tel-Aviv University, Tel-Aviv 69978, Israel and also with the NICE Systems Inc., Ra'anana 43107, Israel (e-mail: arie@eng.tau.ac.il).

G. Gómez-Herrero is with the Institute of Signal Processing, Tampere University of Technology, Tampere, 33101 TUT, Finland (e-mail: german.gomezherrero@tut.fi).

E. Doron was with the School of Electrical Engineering, Tel-Aviv University, Tel-Aviv 69978, Israel. He is now with the Amimon Ltd., Herzlia 46733, Israel (e-mail: Eran@amimon.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TNN.2007.908648

[1] This terminology is quite "loose": for instance, the essence of the "nonwhiteness" property should better be termed "spectral diversity." Note that colored sources cannot be separated using second-order statistics blindly, unless their spectra are distinct.

to separate first-order autoregressive (AR) sources. With certain parametrization of such first-order AR sources, Hyvärinen's algorithm has been observed (in our simulations) to outperform our proposed algorithm.

A previous, more basic method for combining EFICA and WASOBI was recently presented (by us) in [29]. As explained in the sequel, the algorithm presented in here considerably enhances that method by properly accounting for multidimensional independent components within the observed mixtures [1], [3], [7]. Note that unlike [1], [3], and [7], we do not consider multicomponents associated with dependent sources, but only linear mixtures of independent sources which either EFICA or WASOBI fails to separate properly.

The key to successful combination of the two methods lies with the ability to predict (estimate) their resulting performance from their outputs. This information can in turn be used for successive data-adaptive "matching" of each algorithm to the subset(s) of sources for which it outperforms the other. To elaborate, we briefly address the issue of performance assessment in Section II. In Section III, we provide a brief overview of the "building blocks" of the algorithm, which is outlined in Section IV. Extensive simulation results are presented in Section V and some conclusions are drawn in Section VI.

II. Output-Based Performance Assessment

A common measure for evaluating the separation accuracy is the interference-to-signal ratio (ISR). For a given estimate of the demixing matrix $\widehat{\mathbf{W}}$, the "realization-ISR" matrix $\mathbf{rISR}$ is given (elementwise) by $\mathbf{rISR}_{k\ell} = \mathbf{G}_{k\ell}^2/\mathbf{G}_{kk}^2$, where $\mathbf{G} \triangleq \widehat{\mathbf{W}}\mathbf{A}$. The total "realization-ISR" of the kth estimated signal can also be defined as $\mathbf{risr}_k \triangleq \sum_{\ell=1,\ell\neq k}^{d} \mathbf{rISR}_{k\ell}$. Naturally, evaluation of both requires knowledge of the true mixing matrix $\mathbf{A}$, which is normally unavailable (except in simulations).

If the signal separation experiment is repeated in a Monte Carlo fashion, a general key property of any separation algorithm is its "mean-ISR" (or simply its "ISR" for short), given by the expected value of its "realization-ISR," $\mathbf{ISR} \triangleq E[\mathbf{rISR}]$ (with a similar definition for the isr vector). This $\mathbf{ISR}$ depends, in general, on the statistical model of the data generating process.

For some algorithms, the ISR can be determined by analysis, and thanks to the well-known equivariance property (e.g., [4]), this ISR usually does not depend on the unknown $\mathbf{A}$, but only on statistical properties of the sources, which, although unknown as well, may sometimes be estimated empirically from the separated (estimated) sources.

The ability to assess the ISR of an algorithm from simple empirical estimates of statistical properties of its outputs is a desirable but rare feature, shared by very few ICA algorithms. Fortunately, both EFICA and WASOBI do share that attractive feature, which will prove instrumental in the sequel.

Moreover, as we will show in simulation, the validity of the mean ISR estimates for both EFICA and WASOBI is maintained even when the data generating process is somewhat modified. In particular, the following will be shown.

- The EFICA ISR expression, derived assuming temporally white sources, remains approximately valid when the sources are mildly colored.
- The WASOBI ISR expression, derived assuming Gaussian AR sources, remains approximately valid when the Gaussian driving noise is replaced with non-Gaussian noise, as long as the AR coefficients (namely, the spectral shapes of the sources) are maintained.

A partial intuitive explanation may be that EFICA is based only on the marginal distributions of the sources, ignoring any time structures, whereas WASOBI is based only on second-order statistics, ignoring any higher order statistical information. We elaborate on this issue in Section III.

In addition, note the following arguments supporting the idea of the output performance assessment, even for poorly separated sources.

When EFICA fails to separate some of the sources, they remain mixed together and the mixtures' probability distributions would usually be close to Gaussian, thanks to the central limit theorem (because each unseparated observation would still be a linear combination of several independent inputs). As a result, the estimate of the EFICA mean-isr would be relatively high, as the true mean-isr of EFICA is well known to be high for sources with nearly Gaussian distributions.

Similarly, when WASOBI fails to separate some of the sources, if the remaining mixtures are poorly separated, they are prone to have fairly similar spectra (some kind of slightly differently weighted "average" spectra of the sources involved). As a result, the estimate of the WASOBI mean-isr would be high, as the true mean-isr of WASOBI is well known to be high for sources with nearly similar spectra.

Admittedly, these arguments cannot be regarded as rigorous justification of our claim. However, they indicate that the general trend of the estimated mean-isrs can usually be expected to conform with the true situation, even when the separation is poor.

III. Building Blocks

In this section, we briefly describe the essential building blocks of the proposed algorithm. These building blocks are the EFICA and WASOBI separation algorithms, as well as a previously proposed, more basic combination scheme.

The Cramér–Rao lower bound (CRLB) on the (unbiased) estimation of $\mathbf{W}$ induces a different type of lower bound (see, e.g., [11]) on the attainable ISR, in the form of an ISR-like matrix with elementwise bounds. We would refer to that bound as the Cramér–Rao-induced bound (CRIB). A separation algorithm is said to be "optimal" (for a specified mixing model) when its ISR matrix equals the respective CRIB. Both EFICA and WASOBI have been shown to be asymptotically optimal (under some mild conditions) for their respective model classes [18], [9].

A. EFICA

EFICA is essentially a modification of the popular fastICA algorithm [13], belonging to a wide family of ICA algorithms which exploit non-Gaussianity of the sources' distributions (ignoring any time structure). In its general form, fastICA requires a user-defined choice of a set of nonlinear functions $g_k(\cdot)$ $(k = 1, 2, \ldots, d)$ for extracting each of the d sources. EFICA enhances fastICA by offering an elaborate data-adaptive choice of these nonlinearities, followed by a refinement step.

Under the assumption that each row $\mathbf{s}_k$ $(k = 1, \ldots, d)$ of $\mathbf{s}$ contains N independent realizations of non-Gaussian[2] random variables ξ_k, it is shown in [18] that the asymptotic ISR matrix has as elements

$$\mathbf{ISR}_{k\ell} = \frac{1}{N} \frac{\gamma_k(\gamma_\ell + \tau_\ell^2)}{\tau_\ell^2\gamma_k + \tau_k^2(\gamma_\ell + \tau_\ell^2)} \tag{1}$$

where

$$\begin{array}{rclrcl} \gamma_k & = & \beta_k - \mu_k^2 & \mu_k & = & \mathrm{E}[\xi_k g_k(\xi_k)] \\ \tau_k & = & |\mu_k - \rho_k| & \rho_k & = & \mathrm{E}[g_k'(\xi_k)] \\ & & & \beta_k & = & \mathrm{E}[g_k^2(\xi_k)] \end{array}$$

and where $E[\cdot]$ denotes the expectation operator and $g_k'(\cdot)$ denotes the derivative of $g_k(\cdot)$. In the best possible case, obtained by EFICA for sources with generalized Gaussian distributions, (1) equals the respective CRIB [26].

B. WASOBI

WASOBI [33], [9], [28] is a weighted version of the well-known SOBI [2] algorithm, belonging to a wide family of second-order-statistics-based ICA algorithms, which rely on time structures in the sources' correlations. Both SOBI and WASOBI are based on approximate joint diagonalization (AJD) of several (say M) time-lagged estimated correlation matrices

$$\widehat{\mathbf{R}}_{\mathbf{x}}[\tau] = \frac{1}{N-\tau} \sum_{n=1}^{N-\tau} \mathbf{x}[n]\mathbf{x}^T[n+\tau] \tau = 0, \ldots, M-1 \tag{2}$$

where $\mathbf{x}[n]$ denotes the nth column of $\mathbf{x}$.

Unlike SOBI, WASOBI incorporates proper weighting (inversely proportional to the covariance in the correlation estimates) into the AJD process. The weighting is asymptotically optimal for the case of Gaussian sources.

In particular, if all sources are Gaussian AR of order $M-1$, then under asymptotic conditions the ISR matrix attained by WASOBI can be shown to equal the respective CRIB [11]

$$\mathbf{ISR}_{k\ell} = \frac{1}{N} \frac{\phi_{k\ell}}{\phi_{k\ell}\phi_{\ell k} - 1} \frac{\sigma_k^2 R_\ell[0]}{\sigma_\ell^2 R_k[0]} \tag{3}$$

where σ_k^2 is the variance of the innovation sequence of the kth source and $\phi_{k\ell}$ are given by

$$\phi_{k\ell} = \frac{1}{\sigma_k^2} \sum_{i,j=0}^{M-1} a_{i\ell} a_{j\ell} R_k[i-j]$$

where $\{a_{i\ell}\}_{i=0}^{M-1}$ are the AR coefficients of the ℓth source with $a_{0\ell} = 1$ for $k, \ell = \mathbf{1}, \ldots, d$, and $R_k[m]$ is the autocorrelation of the kth source at time lag m (we use a unit-variance scaling assumption $R_k[0] = R_\ell[0] = 1$ in our model).

C. Combined WASOBI and EFICA (COMBI)

An intuitively appealing selection approach would be to apply both EFICA and WASOBI to $\mathbf{x}$ and select for each source the reconstructed version that has the best total realization-ISR of the two. This basic selection approach can then be turned into a successive scheme, such that in each iteration only the "best" separated sources are "accepted," and the remaining signals (which are still weakly separated mixtures of the remaining sources) are subjected to an additional iteration of separation and selection.

The "realization-ISR" matrices are obviously unknown (nor can they be consistently estimated from the data). However, it is possible to substitute these with the "mean-ISRs," thereby attaining a selection strategy which implies proper selection "on the average." Consistent estimates of the mean-ISR matrices $\widehat{\mathbf{ISR}}^{EF}$ and $\widehat{\mathbf{ISR}}^{WA}$ for both EFICA and WASOBI can indeed be obtained from (1) and (3), respectively, by substituting the true sources with the estimated sources and the true expectations with the empirical means. Then, all individual **isr**s estimates $\widehat{\mathrm{isr}}_k^{EF}$ and $\widehat{\mathrm{isr}}_k^{WA}$ (for all k) can be extracted from these matrices. The COMBI algorithm [29] employs these estimates in the following procedure.

1) Let $\mathbf{z} = \mathbf{x}$

2) Apply both EFICA and WASOBI to $\mathbf{z}$; denote the estimated sources as $\mathbf{s}^{EF}$ and $\mathbf{s}^{WA}$, respectively, and the respective estimated isrs as $\widehat{\mathrm{isr}}^{EF}$ and $\widehat{\mathrm{isr}}^{WA}$.

3) Let $E = \min_k \widehat{\mathrm{isr}}_k^{EF}$ and $W = \min_k \widehat{\mathrm{isr}}_k^{WA}$

4) **If** $E < W$

a) accept those signals $\mathbf{s}^{EF}$ for which $\widehat{\mathrm{isr}}_k^{EF} < W$ and redefine $\mathbf{z}$ as the rejected signals of $\mathbf{s}^{EF}$ **else**

b) accept those signals $\mathbf{s}^{WA}$ for which $\widehat{\mathrm{isr}}_k^{WA} < E$ and redefine $\mathbf{z}$ as the rejected signals of $\mathbf{s}^{WA}$

5) If there are more than one rejected signal remaining, go to (2). Otherwise, if any, accept the rejected signal.

Each of the two ISR expressions (1) and (3) was derived under the assumption that all of the sources comply with their respective model assumption. However, when the mixture consists of both non-Gaussian i.i.d. and Gaussian time-structured sources, neither of the model assumptions can be satisfied by all sources. Strictly speaking, this mismatch may undermine the theoretical reliability of the output-based ISR estimates. However, as already mentioned, it has been empirically verified (and will be demonstrated in simulation) that the ISR estimators usually remain reasonably accurate even when the respective model assumptions are mildly violated and when the separation is not perfect.

Moreover, it has to be emphasized that exact ISR values are of little or no interest here, since only their comparative relations are used in the selection process. We note, in addition, that other empirical methods for assessing the resulting ISRs could be considered, such as bootstrap resampling [20]. However, these approaches usually involve a computationally extensive repeated resampling and separation scheme, and may be more suited for i.i.d. sources than for time-structured sources. Thus, the possibility to exploit the analytical expressions (1) and (3) for EFICA and WASOBI is rather appealing and serves as one of the cornerstones of the proposed approach.

Still, a remaining major drawback of the COMBI algorithm described previously is the following. Suppose that one of the two algorithms (EFICA or WASOBI) can attain a nearly block-diagonal ISR matrix, namely, can well separate the mixture into *groups* of sources, but still with poor separation within each

[2]To be precise, at most one of the random variables is allowed to be Gaussian.

group. Then, subsequent application of the other algorithm to each group (separately) may be able to eventually attain good separation of all of the sources. Unfortunately, COMBI would not be able to exploit such potential "two-stage cooperation" between the two algorithms. This is because COMBI is essentially unaware of the group-separation ability of the first algorithm (because only the individual sources' isrs are accounted for).

We, therefore, propose (in Section IV) an enhanced version of COMBI, aimed at applying a more "systematic" approach, capable of accounting for such cases. A simple demonstration of the sources constellation in question, presenting both the drawback and its solution, would appear in Example 4 in Section V.

IV. Proposed Method: Multi-COMBI

A "multidimensional component" is a cluster of signal components that can together be well separated from the other components in the mixture, yet are difficult to separate from one another [1], [3]. For EFICA, only components that have (nearly) Gaussian distributions might form such a cluster, hence at most one such cluster may exist. For WASOBI, any components sharing similar correlation structures (i.e., power spectra) are hardly separable from one another, but may be easily separated as a cluster, hence several such clusters might coexist.

Each cluster is characterized by the set of indices of the sources it contains, denoted $\mathcal{I}$, $\mathcal{I} \subset \{1,\ldots,d\}$. Using an estimate of the ISR matrix in (1) or in (3), the $\widehat{\text{isr}}$ of a cluster $\mathcal{I}$ (with respect to all the other sources) can be defined as

$$\widehat{\text{isr}}(\mathcal{I}) = \alpha_{\mathcal{I}} \sum_{k\in\mathcal{I},\ell\notin\mathcal{I}} \widehat{\mathbf{ISR}}_{k\ell} \tag{4}$$

where $\alpha_{\mathcal{I}}$ is some normalization coefficient depending on the cluster's cardinality (dimension) $|\mathcal{I}|$ and on d. We propose to take

$$\alpha_{\mathcal{I}} = \frac{d-1}{|\mathcal{I}|(d-|\mathcal{I}|)}$$

so that $\text{isr}(\mathcal{I})$ has the meaning of $d-1$ times the average of the entries in the sum in (4). This choice is compatible with the basic definition of isr_k for $\mathcal{I}=\{k\}$.

The proposed "multi-COMBI" algorithm works recursively with a stack of clusters $\mathcal{S}$. In each step, one of the clusters in the stack, that is not a *singleton*, i.e., does not have dimension 1, is decomposed into two or more smaller clusters, until all clusters are singletons. The algorithm can be summarized as follows.

To initialize, let the stack of clusters $\mathcal{S}$ be comprised of a single cluster containing the entire set $\mathcal{S} := \{\mathbf{x}\}$.

1) Pick any cluster in $\mathcal{S}$ that is not a singleton, and denote this cluster as $\mathbf{z}$ (obviously, $\mathbf{z}=\mathbf{x}$ in the first step).

2) Apply both EFICA and WASOBI to $\mathbf{z}$; obtain the separated signals $\mathbf{s}^{EF}$ and $\mathbf{s}^{WA}$ and the corresponding estimated ISR matrices $\widehat{\mathbf{ISR}}^{EF}$ and $\widehat{\mathbf{ISR}}^{WA}$, estimated from the separated data using (1) and (3).

3) Construct a set $\mathcal{C}$ of possible clusters $\mathcal{I} \subset \{1,\ldots,\dim(\mathbf{z})\}$. For example, if $\mathbf{z}$ contains three signals, then $\mathcal{C} = \{\{1\},\{2\},\{3\},\{1,2\},\{1,3\},\{2,3\}\}$. Note that $\mathcal{C}$ does not have to include *all* possible clusters—see Section IV-A for a further discussion.

4) Based on the estimated ISR matrices, compute (using (4)) $\widehat{\text{isr}}^{EF}(\mathcal{I})$ and $\widehat{\text{isr}}^{WA}(\mathcal{I})$ for each $\mathcal{I}\in\mathcal{C}$. Namely, in the same example, compute $\widehat{\text{isr}}^{EF}(\{1\}), \widehat{\text{isr}}^{EF}(\{2\}), \widehat{\text{isr}}^{EF}(\{3\}), \widehat{\text{isr}}^{EF}(\{1,2\})$, etc.

5) Let $E := \min_{\mathcal{I}} \widehat{\text{isr}}^{EF}(\mathcal{I})$ and $W := \min_{\mathcal{I}} \widehat{\text{isr}}^{WA}(\mathcal{I})$

6) **If** $E < W$, pick up the set of "best" EFICA-separated clusters as follows:

$$\mathcal{I}_1 := \text{argmin}_{\mathcal{I}\in\mathcal{C}} \widehat{\text{isr}}^{EF}(\mathcal{I})$$

and then, for $k=1,2,\ldots$, repeat the following:

$$\mathcal{I}_{k+1} := \text{argmin}_{\mathcal{I}\in\mathcal{C}-\{\mathcal{I}:\exists\ell\in[1,k]:\mathcal{I}_\ell\cap\mathcal{I}\neq\emptyset\}} \widehat{\text{isr}}^{EF}(\mathcal{I})$$

until either

$$\widehat{\text{isr}}^{EF}(\mathcal{I}_{k+1}) > W$$

or $\mathcal{C} - \{\mathcal{I};\exists\ell\in[1,k];\mathcal{I}_\ell\cap\mathcal{I}\neq\emptyset\}$ is empty. This procedure picks up the "best" (lowest **isr**) EFICA-separated clusters one by one: At each step, the best remaining cluster in $\mathcal{C}$ (among those disjoint with the clusters picked up so far) is picked up (such a scheme is sometimes called a *greedy* algorithm). The procedure stops either when all clusters have been picked up, or when the best remaining cluster is already worse than the best WASOBI-separated cluster. The value of k upon exit is denoted M.

Let $\mathcal{J} := \{\mathbf{1},\ldots,\dim(\mathbf{z})\} - \cup\mathcal{I}_k$. If $\mathcal{J}$ is not empty, let $M := M+1$ and $\mathcal{I}_M := \mathcal{J}$.

The new clusters $\mathbf{s}_1,\ldots,\mathbf{s}_M$ are extracted from $\mathbf{s}^{EF}$ according to the partitioning $\mathcal{I}_1,\ldots\mathcal{I}_M$.

else (for $E > W$) extract $\mathbf{s}_1,\ldots,\mathbf{s}_M$ similarly from $\mathbf{s}^{WA}$ using $\widehat{\text{isr}}^{WA}$.

7) update $\mathcal{S}$ by substituting $\mathbf{z}$ with $\mathbf{s}_1,\ldots,\mathbf{s}_M$

$$\mathcal{S} := (\mathcal{S} - \{\mathbf{z}\}) \cup \{\mathbf{s}_1,\ldots,\mathbf{s}_M\}$$

8) If all clusters in $\mathcal{S}$ are already singletons, stop. Otherwise return to 1).

A simplified demonstration of the progress of the algorithm can be found in the context of Example 4 in Section V.

A. Alternative (Proposed) Construction of the Set C

When d is not large, then the set $\mathcal{C}$ of the cluster candidates in Step 3) can contain all $2^{\dim(\mathbf{z})}-2$ nontrivial subsets of $\{\mathbf{1},\ldots,\dim(\mathbf{z})\}$. However, when d is large, say $d \geq 20$, computing isrs of all of these subsets can be prohibitively slow. We, therefore, propose, in high-dimensional cases, considering a smaller set of relevant cluster candidates $\mathcal{C}$. The set $\mathcal{C}$ can be constructed using any well-established clustering method such as K-means, hierarchical clustering and many others (see, e.g., [24]).

For the EFICA-separated signals $\mathbf{s}^{EF}$, no clustering is actually required, as $\mathcal{C}$ can be simply determined as the set of all singletons. This is because, for EFICA, we know *a priori* that at

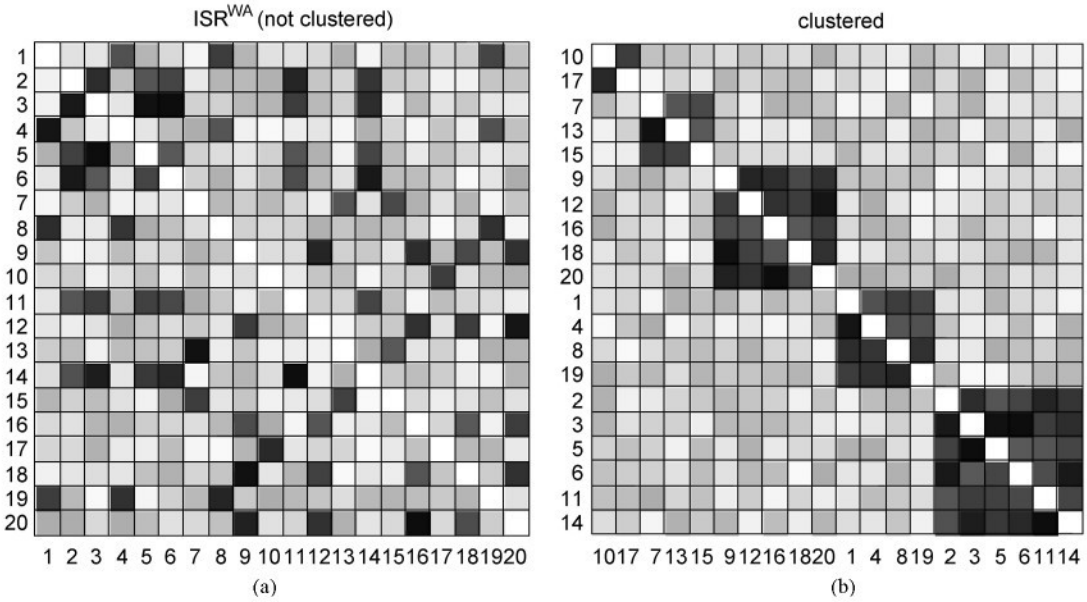


Fig. 1. Example of hierarchical clustering of a WASOBI ISR matrix (in grayscale). (a) Before. (b) After.

most one cluster of nondistinguishable (nearly Gaussian) components can exist. This cluster would be found as the remainder set $\mathcal{J}$ in Step 5). Therefore, a clustering method is needed only to process the WASOBI-separated signals $\mathbf{s}^{WA}$.

To apply such clustering, note that we may regard any estimated ISR matrix ($\widehat{\mathbf{ISR}}^{WA}$ in our case) as describing inverse distances between nodes on a graph, where the nodes are the source signals. A high value in $\widehat{\mathbf{ISR}}_{k\ell}$ means that sources k and ℓ are "close," namely, not well separated and should, therefore, belong to the same cluster. Conversely, a low $\widehat{\mathbf{ISR}}_{k\ell}$ implies that sources k and ℓ are well separated and should, therefore, belong to different clusters. However, since we are not interested in clustering a directed graph (namely, we do not distinguish between $\widehat{\mathbf{ISR}}_{k\ell}$ and $\widehat{\mathbf{ISR}}_{\ell k}$ for the clustering), we can base the clustering on a symmetrized version of the $\mathbf{ISR}$ matrices, $\mathbf{D} \triangleq \widehat{\mathbf{ISR}} + \widehat{\mathbf{ISR}}^T$.

In this paper, we suggest to construct the set of cluster candidates $\mathcal{C}$ using a *hierarchical clustering with a single linking strategy* [24]. Here, the set $\mathcal{C}$ is built recursively, so that in the beginning it contains all singletons. At each step, we look for the couple (k, ℓ) for which $\mathbf{D}_{k\ell}$ obtains its maximum value, and then create and add a new cluster to $\mathcal{C}$, formed by the union of the most recently created cluster containing signal k and the most recently created cluster containing signal ℓ. In addition, we zero-out the (k, ℓ) and (ℓ, k) entries in $\mathbf{D}$, so as not to reuse the same couple in subsequent steps. The update of $\mathcal{C}$ terminates after $\dim(\mathbf{z}) - 1$ steps and contains $2\dim(\mathbf{z}) - 2$ entries at the end.[3] Note that the cardinality of $\mathcal{C}$ would usually be significantly smaller than the number of all possible clusters $2^{\dim(\mathbf{z})} - 2$.

Once the set $\mathcal{C}$ of candidate clusters is obtained, the "leading clusters" can be selected, e.g., using a greedy algorithm based on each cluster's $\widehat{\text{isr}}(\mathcal{I})$ [calculated using (4)]. This selection is required in Step 5).

[3]It is because, in each update, the number of the clusters available for further fusion decreases by one.

The clustering scheme described previously is an *ad hoc* algorithm, which can be replaced by a more sophisticated method in the future. However, in our simulations, this scheme works well and seems more accurate than the spectral clustering method advocated in [20] in a similar context.

We illustrate a typical clustering result of this clustering algorithm in Fig. 1. On the left-hand side, we show the ISR matrix for 20 sources in grayscale colors, where lighter colors denote low ISR (good separation) and darker colors denote high ISR (poor separation). The resulting reordering and partition into clusters is clearly observed on the right-hand side.

B. Cluster Issues

We note in passing that under poor separation conditions (e.g., short data length N), situations containing poorly distinguishable (overlapping) clustering might also occur. Indeed, theoretically (and asymptotically), for the EFICA model, there can only be one cluster of inseparable sources, namely, a cluster of Gaussian sources. For WASOBI, there can be several clusters that group sources with identical spectra (different between clusters). Therefore, strictly speaking, the residual clusters produced by each method separately should not overlap. However, in reality (especially under nonasymptotic conditions), this might not hold true in some situations, e.g., if there are some similarities in spectra between sources in different clusters of WASOBI or if there are sources which are "roughly" Gaussian (for EFICA). In such cases, the clusters might not be strictly disjoint. However, the algorithm relies on some thresholding of the ISR, which would eventually yield some (possibly inaccurate) disjoint clustering, hopefully (but not necessarily) a "good" one. Nevertheless, under the specified model assumptions, as the observation length N increases, the clusters are guaranteed to become well distinguishable.

V. Simulation Results

We conducted a series of simulation experiments aimed at comparative evaluation of the proposed multi-COMBI

approach, as well as at verifying the validity of the intermediate ISR estimates. As discussed earlier, the analytic ISR expressions were obtained under their respective "nominal" homogeneous model assumptions, which are deliberately breached in our experiments' setup. Moreover, when using these expressions, some true (unknown) quantities are replaced by their empirical estimates from the output signals, which might not be well separated. It is, therefore, essential to verify (at least empirically) that the output-based ISR estimates, on which the entire multi-COMBI approach is based, are indeed valid.

Thus, the first three simulation examples in this section demonstrate the remarkable agreement (under mild deviations from the model assumptions) between the empirical performance of EFICA and WASOBI and their theoretical predictions obtained using (1) and (3) (with empirical quantities). In addition, we compare the resulting multi-COMBI and COMBI performance to some competing algorithms.

The fourth example illustrates the advantages of multi-COMBI with respect to the less sophisticated COMBI in the presence of clusters. The last three examples challenge the robustness of all algorithms, demonstrating the maintained superiority of multi-COMBI in larger scale problems (containing several large clusters) and in the presence of additive noise.

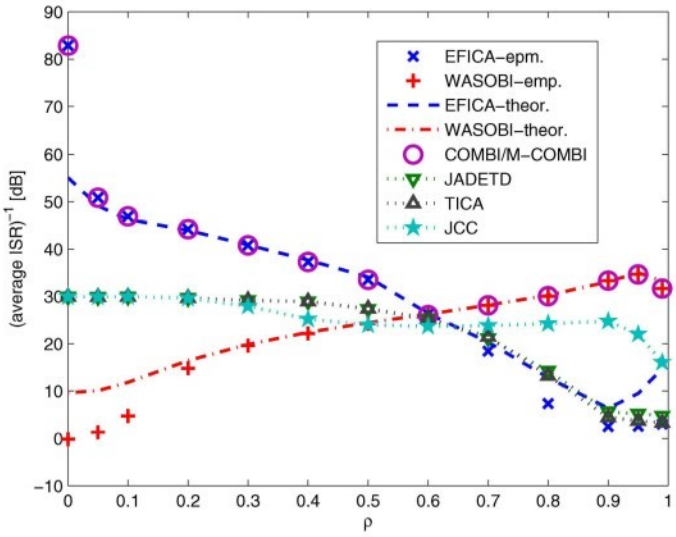


Fig. 2. Inverted average ISR achieved in separation of five AR signals obtained by passing BPSK i.i.d. sequences of length $N = 1000$ through all-pole filters with AR coefficients $[1, \rho]$, $[1, 0, \rho]$, $[1, 0, 0, \rho]$, $[1, 0, 0, 0, \rho]$, and $[1, 0, 0, 0, 0, \rho]$, respectively, and the theoretically predicted ISR (1) and (3), estimated using the separated signals, versus varying ρ. Each simulation point is an average of 100 trials.

Example 1—Fig. 2

In the first experiment, we consider the separation of five colored non-Gaussian sources versus a parameterized variation of their spectral diversity. $N = 1000$ samples of each source were generated by filtering statistically independent random binary phase-shift keying (BPSK) sequences using all-pole filters. For each $k = 1, 2, \ldots, 5$, the kth filter was constructed of k poles, located at all k roots of the real-valued parameter ρ. In other words, the filters' AR coefficients were $[1, \rho]$, $[1, 0, \rho]$, $[1, 0, 0, \rho]$, $[1, 0, 0, 0, \rho]$, and $[1, 0, 0, 0, 0, \rho]$, for $0 \leq \rho < 1$.

For small values of ρ, the sources are strongly non-Gaussian, having a weak (and rather similar) temporal correlation structure, so EFICA should be superior to WASOBI. Conversely, as ρ approaches 1, the sources can be equivalently reproduced with effectively very long finite impulse response (FIR) filters, and, therefore (by the central limit theorem), have nearly Gaussian marginal distributions, yet with strong different temporal correlation structures, so WASOBI should clearly outperform EFICA.

Since the obtained ISR values in each experiment were roughly similar for all of the sources, we merely display the performance in terms of a single, average ISR (inverted, for convenience), averaged over all sources and over all trials. In each trial, all elements of the mixing matrix were redrawn independently from a standard Gaussian distribution.

The theoretically predicted ISRs were obtained empirically in each trial, by substituting the unknown statistical properties in (1) and (3) with their empirically obtained values from the separated sources. These ISR values were also averaged over all sources and over all trials and their inverted values were displayed versus the spectral-shape parameter ρ.

We note the remarkable agreement of the performance of both EFICA and WASOBI with their theoretical prediction over the entire range of ρ, except for the extreme cases $\rho \approx 0$ and $\rho \approx 1$, where the deviation is more significant. In the higher region of ρ, the predicted ISR of EFICA is slightly overoptimistic, i.e., the inverted mean predicted ISR is slightly higher than the actual inverted ISR (yet the relative order is evidently maintained).

The performance of COMBI and multi-COMBI is compared in this and in subsequent experiments with the other following algorithms: JADE$_{\mathrm{TD}}$ [23] with parameters 0:5, JCC [12] with parameters $p = 5$ and $\tau = [2, 3, 4, 5, 6]$, and TICA [6] with parameters $(d1, d2, d3) = (5, 0, 5)$.

Example 2—Fig. 3

In this experiment, we fed the all-pole filters described in the previous example with (super-Gaussian) i.i.d. samples taken from a generalized Gaussian distribution with parameter α.[4] For easy reference, the distribution is denoted GG(α). Fig. 3(a) shows the result for $\alpha = 0.5$ as a function of parameter ρ and Fig. 3(b) shows the result for $\rho = 0.5$ and varying α. Each simulation point is an average of 100 trials.

The general behavior in Fig. 3(a) is similar to that observed in the previous example, with the difference that the performance of all algorithms is statistically less stable than in the first experiment, probably due to the long tail distribution of the data.

Fig. 3(b) shows that for α below 0.5, the non-Gaussian character of the data is the dominant key property for separation and, therefore, EFICA is more accurate than WASOBI. For α higher than 0.5, WASOBI is more accurate since the temporal correlation structure becomes the dominant key property for separation. As in the previous examples, both COMBI and multi-COMBI are able to effectively combine the advantages of EFICA and WASOBI and, at the same time, outperform the other competing algorithms.

[4]This distribution has a density proportional to $\exp(-\beta|x|^{\alpha})$, where $\alpha > 0$ controls the shape of the distribution and $\beta > 0$ controls the variance. See, e.g., [18] for more details.

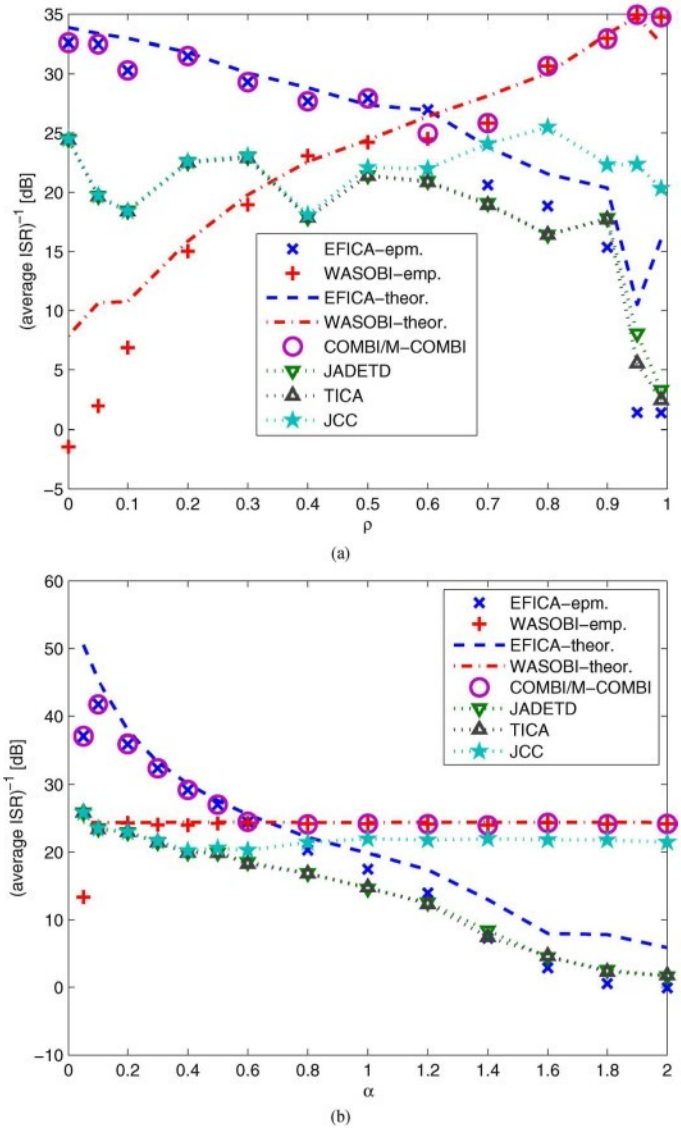


Fig. 3. Inverted average ISR achieved in separation of five AR signals obtained by passing white i.i.d. GG(α) distributed sequences of length $N = 1000$ through all-pole filters with AR coefficients $[1, \rho]$, $[1, 0, \rho]$, $[1, 0, 0, \rho]$, $[1, 0, 0, 0, \rho]$, and $[1, 0, 0, 0, 0, \rho]$, respectively, and the theoretically predicted ISR (1) and (3), estimated using the separated signals. (a) Result for a fixed $\alpha = 0.5$ versus varying ρ. (b) Result for $\rho = 0.5$ versus varying α.

Example 3—Fig. 4

This experiment demonstrates the advantage of COMBI and multi-COMBI in scenarios where neither EFICA nor WASOBI are able to separate all signals, yet COMBI, multi-COMBI, JADE$_{\mathrm{TD}}$[23], TICA [6], JCC [12], and Hyvärinen's algorithm "unified" [15] can. Moreover, we demonstrate how, due to its ability to account for clusters, multi-COMBI can outperform COMBI and almost all of the other competitors. We considered four AR sources. The first source was generated by filtering an independent BPSK sequence using an all-pole filter with coefficients $[1, \rho]$. The second source was generated by feeding the same filter with Gaussian i.i.d. samples. The third and fourth sources were generated in the same way but the coefficients of the filter were $[1, -\rho]$. Obviously, the first and second pairs are each nonseparable by WASOBI, while the second and fourth sources are nonseparable by EFICA.

COMBI first separates the two non-Gaussian sources using EFICA and subsequently separates the two remaining sources with WASOBI. Note, however, that for large values of ρ the initial separation of the non-Gaussian sources may be rather poor, since the increased effective length of the filters renders the marginal distributions of their outputs nearly Gaussian. Significantly better separation is then achieved by multi-COMBI, which is able to exploit the excellent ability of WASOBI to first separate the pair (cluster) of components with one spectral density from the pair (cluster) with the other spectrum, leaving for EFICA the remaining task of separating each pair internally. For example, in one trial with $\rho = 0.6$, we got the following ISR matrices (in natural ratio numbers, not in decibels):

$$\widehat{\mathbf{ISR}}^{WA} = \begin{bmatrix} - & 0.5234 & 0.0004 & 0.0003 \\ 0.5716 & - & 0.0004 & 0.0004 \\ 0.0004 & 0.0004 & - & 0.0389 \\ 0.0003 & 0.0003 & 0.0269 & - \end{bmatrix}$$

$$\widehat{\mathbf{ISR}}^{EF} = \begin{bmatrix} - & 0.0013 & 0.0013 & 0.0008 \\ 0.0022 & - & 0.1272 & 0.0019 \\ 0.0023 & 0.1281 & - & 0.0019 \\ 0.0007 & 0.0009 & 0.0009 & - \end{bmatrix}.$$

From these ISR matrices, we can see that the clusters of components $\{1, 2\}$ and $\{3, 4\}$ in WASOBI are better separated from one another (having lower residual presence of each in other) than clusters $\{1\}$, $\{2, 3\}$, and $\{4\}$ in the EFICA result. More specifically

$$\widehat{\mathrm{isr}}^{WA}(\{1, 2\}) \approx \widehat{\mathrm{isr}}^{WA}(\{3, 4\}) \approx 0.0004$$

and

$$\widehat{\mathrm{isr}}^{EF}(\{2, 3\}) \approx 0.0021\,.$$

In this situation, WASOBI cannot accurately resolve individual components but it separates the two 2-D clusters better than EFICA. Contrary to COMBI, multi-COMBI detects this fact and correctly chooses WASOBI for the initial separation, yielding improved performance.

Note that multi-COMBI outperforms almost all of the competitors for almost all values of ρ, with one significant exception: Hyvärinen's algorithm outperforms multi-COMBI (in all four experiments) for ρ above 0.6–0.7. This means that there is still room from improvement, as multi-COMBI does not (and is not claimed to) inherit the optimality of its building blocks EFICA and WASOBI. We note in passing, that the implementation of Hyvärinen's algorithm that is available so far is inapplicable to separation of AR processes of higher orders and to separation of sources of an unknown type (super-Gaussian/sub-Gaussian), because each type requires a different built-in nonlinear function. (In this example, we have used "pow3" to achieve a good performance.)

Example 4–Fig. 5

In this experiment, we mixed (and separated) 20 AR sources comprised of four groups of five sources each. Each of the five groups was generated with the same set of filters used in the first experiment with $\rho = 0.6$. The only difference between the groups was the distribution of the i.i.d. "driving noise," which

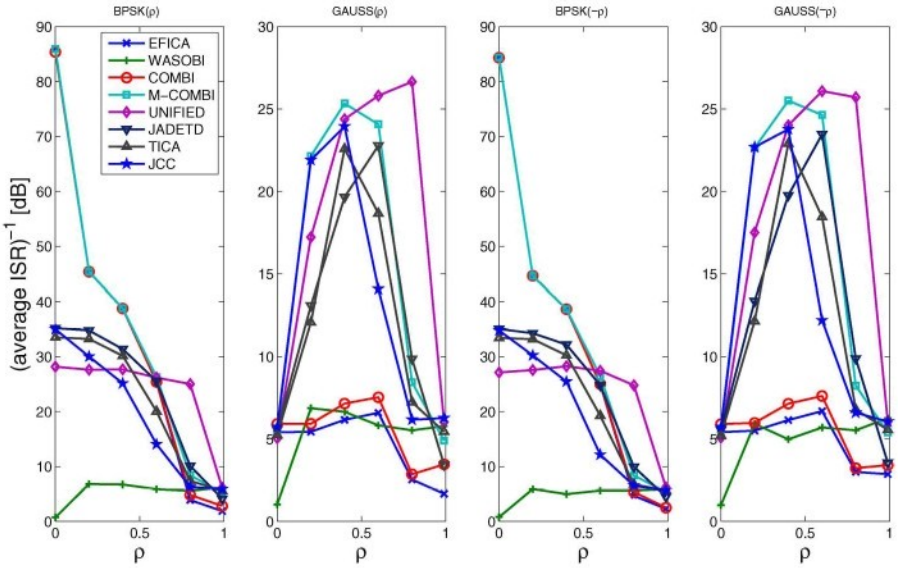


Fig. 4. Inverted average ISR achieved in separation of four AR signals obtained by passing BPSK, Gaussian, BPSK, and Gaussian i.i.d. sequences of length $N = 1000$ through all-pole filters whose AR coefficients were $[1, \rho]$, $[1, \rho]$, $[1, -\rho]$, and $[1, -\rho]$, respectively. Each simulation point is an average of 100 trials.

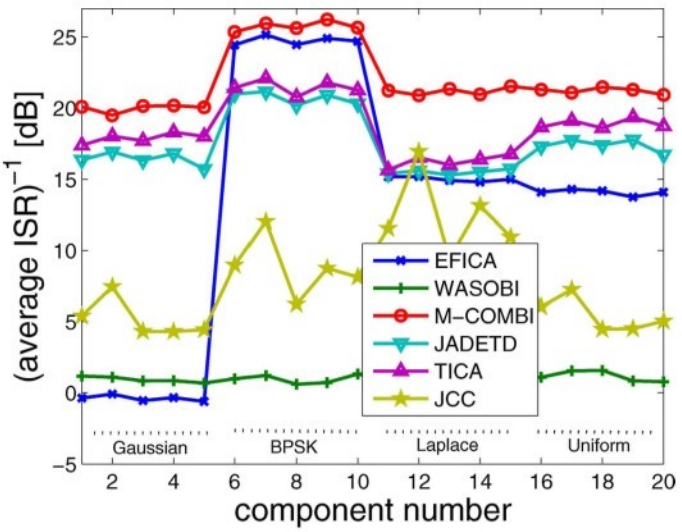


Fig. 5. Inverted average isr (for each component separately) achieved in separation of 20 AR sources of length $N = 5000$, driven by i.i.d. sequences of Gaussian, BPSK, Laplace, and Uniform distributions passing through all-pole filters with AR coefficients $[1, \rho]$, $[1, 0, \rho]$, $[1, 0, 0, \rho]$, $[1, 0, 0, 0, \rho]$, and $[1, 0, 0, 0, 0, \rho]$ for $\rho = 0.6$. Each simulation point is an average of 100 trials.

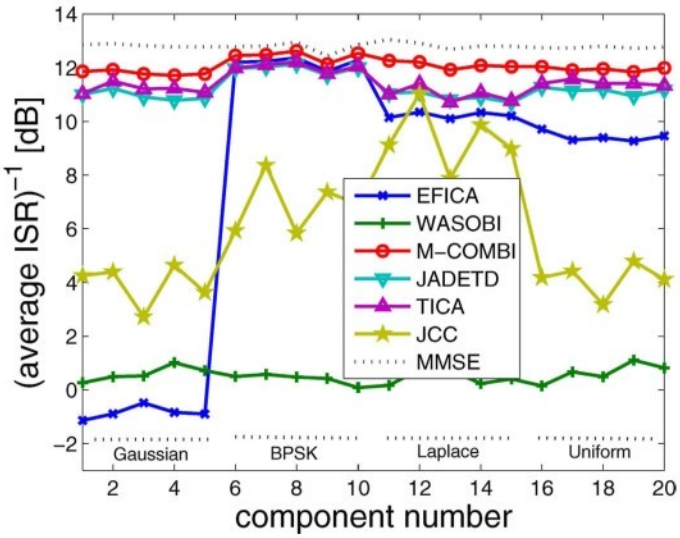


Fig. 6. Inverted average INSR achieved in separation of the same mixtures of 20 AR sources as in Fig. 5, contaminated by AWGN at 0-dB SNR. MMSE denotes the performance of a hypothetical "oracle" separator which uses the known mixing matrix and noises' variances.

was Gaussian for the first group, BPSK for the second, Laplace ($=$ GG(1)) for the third, and Uniform ($=$ $\lim \mathrm{GG}(\alpha)$ for $\alpha \to \infty$) for the fourth. Thus, for EFICA, the first group of five Gaussian signals comprises a nonseparable cluster, whereas for WASOBI there are five different clusters, each comprised of four signals with similar spectra (and different marginal distributions).

The results are shown in terms of the inverted average ISR for all 20 sources. Across-the-board superiority of multi-COMBI is clearly evident.

Example 5–Fig. 6

In this experiment, the scenario of the previous experiment is repeated with the exception that now the observations are contaminated by additive white Gaussian noise (AWGN). The noises' variances were set so as to maintain input signal-to-noise ratio (SNR) of 0 dB for all sources. The mixing matrices were taken at random with independent Gaussian distributed elements, normalized such that each row of $\mathbf{A}^{-1}$ had unit norm [17], and censored so that their condition numbers lie in the interval $[10, 100]$. The results are shown in terms of the inverted averaged interference-plus-noise-to-signal ratio (INSR) and are also compared to the empirical performance of an "oracle" minimum mean square error (MMSE) separator, which uses the known mixing matrix and noise variance. It is evident that the superior performances of multi-COMBI with these sources is maintained also in the presence of AWGN.

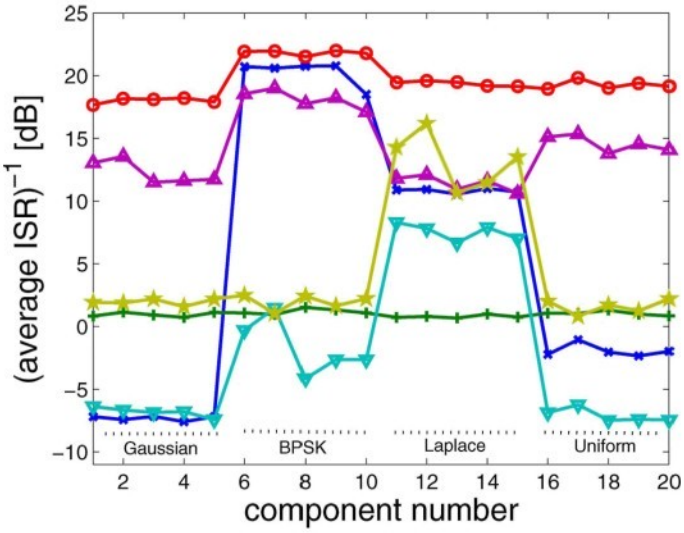


Fig. 7. Inverted average isr (for each component separately) achieved in separation of 40 sources; 20 AR sources of Fig. 5 plus 20 white Gaussian sources. Since the Gaussian sources are inseparable, they are excluded from the Figure. Each simulation point is an average of 100 trials.

TABLE I
RUNNING TIMES (IN SECONDS) OF DIFFERENT SEPARATION ALGORITHMS

Algorithm	20 sources (ex. 5)	40 sources (ex. 7)
EFICA	2.5	6.7
WASOBI	0.5	2.1
MULTI-COMBI	5.2	22.6
JADE_{TD}	2.6	434.7
TICA	22.6	378.2
JCC	22.4	712.7

Example 6–Fig. 7

In this experiment, 20 white Gaussian (unresolvable) sources were added to the scenario considered in Example 5, yielding 40 mixtures of 40 sources, with only 20 being separable from each other (as well as from the other 20). We display the results for the separable sources only. Again, multi-COMBI is clearly shown to outperform the other algorithms.

Computational Aspects

The computational load of each algorithm was compared when operating on the large-scale mixtures of 20 sources (Example 5) and of 40 sources (Example 7). Our hierarchical clustering algorithm was used in multi-COMBI, as described in Section IV. The average running times of each algorithm with the parameters specified in Example 1 and running on the same personal computer (PC; P4 3-GHz, 2-GB RAM, Windows XP) in Matlab® version 7.0/R14 are summarized in Table I.

VI. CONCLUSION

We have proposed a novel ICA algorithm[5] that effectively combines the two powerful ICA methods, EFICA and WASOBI, thereby allowing separation of mixtures of sources that would be otherwise poorly separated by either one. Computer simulations show good performance of the algorithm compared to competing algorithms, such as JADE_{TD}, JCC, and TICA, both in terms of separation quality and in terms of computational efficiency.

[5]The Matlab codes for EFICA, WASOBI, COMBI, and multi-COMBI can be downloaded from http://si.utia.cas.cz/Tichavsky.html. A Matlab implementation of JADE_{TD} is available at http://www.cs.tut.fi/~omezher/software.htm

ACKNOWLEDGMENT

The authors would like to thank I. Gorodnitsky [12] and Prof. S. C. Álvarez, the author of TICA [6], for sending us their Matlab codes, and Dr. A. Hyvärinen, the author of the unified algorithm [15], for his assistance with his Matlab code.

REFERENCES

[1] F. R. Bach and M. I. Jordan, "Beyond independent components: Trees and clusters," *J. Mach. Learn. Res.*, vol. 4, pp. 1205–1233, 2003.
[2] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, and E. Moulines, "A blind source separation technique using second-order statistics," *IEEE Trans. Signal Process.*, vol. 45, no. 2, pp. 434–444, Feb. 1997.
[3] J.-F. Cardoso, "Multidimensional independent component analysis," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Seattle, WA, 1998, pp. 1941–1944.
[4] J.-F. Cardoso, "Blind signal separation: Statistical principles," *Proc. IEEE*, vol. 86, pp. 2009–2025, 1998.
[5] J.-F. Cardoso, "The three easy routes to independent component analysis; contrasts and geometry," in *Proc. Int. Conf. Independent Component Anal. Blind Signal Separat. (ICA)*, San Diego, CA, Dec. 2001.
[6] S. A. Cruces-Alvarez and A. Cichocki, "Combining blind source extraction with joint approximate diagonalization: Thin algorithms for ICA," in *Proc. Int. Conf. Independent Component Anal. Blind Signal Separat. (ICA)*, San Diego, CA, Dec. 2001, pp. 740–745.
[7] L. De Lathauwer, D. Callaerts, B. De Moor, and J. Vandewalle, "Fetal electrocardiogram extraction by source subspace separation," in *Proc. IEEE Signal Process./ATHOS Workshop Higher-Order Statist.*, Girona, Spain, Jun. 12–14, 1995, pp. 134–138.
[8] A. Delorme *et al.*, "EEGLAB: A Software Toolbox for Electrophysiological Data Analysis," The Salk Inst., La Jolla/Swartz Cntr. Comput. Neurosci., Inst. Neural Comput., Univ. California San Diego, La Jolla, CA, 2002 [Online]. Available: http://sccn.ucsd.edu/eeglab/
[9] E. Doron and A. Yeredor, "Asymptotically optimal blind separation of parametric Gaussian sources," in *Proc. Int. Conf. Independent Component Anal. Blind Signal Separat. (ICA)*, Granada, Spain, Sep. 2004, pp. 390–397.
[10] E. Doron, "Asymptotically optimal blind separation of parametric Gaussian sources," M.Sc. thesis, Dept. Electr. Eng. Syst., Tel-Aviv Univ., Tel-Aviv, Israel, 2003.
[11] E. Doron, A. Yeredor, and P. Tichavský, "Cramér-Rao-induced bound for blind separation of stationary parametric Gaussian sources," *IEEE Signal Process. Lett.*, vol. 14, no. 6, pp. 417–420, Jun. 2007.
[12] I. F. Gorodnitsky and A. Belouchrani, "Joint cumulant and correlation based signal separation with application to EEG data," in *Proc. Int. Conf. Independent Component Anal. Blind Signal Separat. (ICA)*, San Diego, CA, Dec. 2001, pp. 475–480.
[13] A. Hyvärinen, "Fast and robust fixed-point algorithms for independent component analysis," *IEEE Trans. Neural Netw.*, vol. 10, no. 3, pp. 626–634, May 1999.
[14] A. Hyvärinen, "Complexity pursuit: Separating interesting components from time-series," *Neural Comput.*, vol. 13, no. 4, pp. 883–898, 2001.
[15] A. Hyvärinen, "A unifying model for blind separation of independent sources," *Signal Process.*, vol. 85, no. 7, pp. 1419–1427, 2005.
[16] C. J. James and C. V. Hesse, "Independent component analysis for biomedical signals," *Physiol. Meas.*, vol. 26, pp. R15–R39, 2004.
[17] Z. Koldovský and P. Tichavský, "Methods of fair comparison of performance of linear ICA techniques in presence of additive noise," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Toulouse, France, May 14–19, 2006, vol. 5, pp. 873–876.
[18] Z. Koldovský, P. Tichavský, and E. Oja, "Efficient variant of algorithm FastICA for independent component analysis attaining the Cramer-Rao lower bound," *IEEE Trans. Neural Netw.*, vol. 17, no. 5, pp. 1265–1277, Sep. 2006.
[19] M. Ye, "Global convergence analysis of a discrete time nonnegative ICA algorithm," *IEEE Trans. Neural Netw.*, vol. 17, no. 1, pp. 253–256, Jan. 2006.

[20] F. Meinecke, A. Ziehe, M. Kawanabe, and K. R. Müller, "A resampling approach to estimate the stability of one-dimensional or multidimensional independent components," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 12, pp. 1514–1525, Dec. 2002.
[21] E. Oja and Z. Yuan, "The FastICA algorithm revisited: Convergence analysis," *IEEE Trans. Neural Netw.*, vol. 17, no. 6, pp. 1370–1381, Nov. 2006.
[22] D.-T. Pham and J.-F. Cardoso, "Blind separation of instantaneous mixtures of non stationary sources," *IEEE Trans. Signal Process.*, vol. 49, no. 9, pp. 1837–1848, Sep. 2001.
[23] K. R. Müller, P. Philips, and A. Ziehe, "$\mathrm{JADE}_{\mathrm{TD}}$: Combining higher-order statistics and temporal information for blind source separation (with noise)," in *Proc. Int. Conf. Independent Component Anal. Blind Signal Separat. (ICA)*, Aussois, Jan. 1999, pp. 87–92.
[24] H. Späth, *Cluster Analysis Algorithms for Data Reduction and Classification of Objects*. Chichester, U.K.: Ellis Horwood, 1980.
[25] P. Tichavský and Z. Koldovský, "Optimal pairing of signal components separated by blind techniques," *IEEE Signal Process. Lett.*, vol. 11, no. 2, pp. 119–122, Feb. 2004.
[26] P. Tichavský, Z. Koldovský, and E. Oja, "Performance analysis of the FastICA algorithm and Cramér-Rao bounds for linear independent component analysis," *IEEE Trans. Signal Process.*, vol. 54, no. 4, pp. 1189–1203, Apr. 2006.
[27] P. Tichavský, J. Nielsen, and V. Krajča, "Identification of epileptic activity of brain in EEG recordings using four ICA techniques," in *Proc. Biosignal*, Brno, Czech Republic, Jun. 28–30, 2006, pp. 166–168.
[28] P. Tichavský, E. Doron, A. Yeredor, and J. Nielsen, "A computationally affordable implementation of an asymptotically optimal BSS algorithm for AR sources," in *Proc. Eur. Signal Process. Conf. (EUSIPCO)*, Florence, Italy, Sep. 2006.
[29] P. Tichavský, Z. Koldovský, E. Doron, A. Yeredor, and G. Gomez-Herrero, "Blind signal separation by combining two ICA algorithms: HOS-based EFICA and time structure-based WASOBI," presented at the Eur. Signal Process. Conf. (EUSIPCO), Florence, Sep. 2006.
[30] P. Tichavsky, Z. Koldovsky, and E. Oja, "Speed and accuracy enhancement of linear ICA techniques using rational nonlinear functions," in *Proc. Int. Conf. Independent Component Anal. Blind Signal Separat. (ICA)*, London, U.K., Sep. 9–12, 2007, pp. 285–292.
[31] F. Vrins, J. A. Lee, and M. Verleysen, "A minimum-range approach to blind extraction of bounded sources," *IEEE Trans. Neural Netw.*, vol. 18, no. 3, pp. 809–822, May 2007.
[32] L. Wei and J. C. Rajapakse, "Approach and applications of constrained ICA," *IEEE Trans. Neural Netw.*, vol. 18, no. 3, pp. 809–822, May 2006.
[33] A. Yeredor, "Blind separation of Gaussian sources via second-order statistics with asymptotically optimal weighting," *IEEE Signal Process. Lett.*, vol. 7, no. 7, pp. 197–200, Jul. 2000.

Petr Tichavský (M'98–SM'04) received the M.S. degree in mathematics from the Czech Technical University, Prague, Czechoslovakia, in 1987 and the Ph.D. degree in theoretical cybernetics from the Czechoslovak Academy of Sciences, Prague, Czech Republic, in 1992.

Currently, he is with the Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Prague, Czech Republic. In 1994, he received the *Fulbright grant* for a ten-month Fellowship at the Department of Electrical Engineering, Yale University, New Haven, CT. He is the author and coauthor of research papers in the area of sinusoidal frequency/frequency-rate estimation, adaptive filtering and tracking of time varying signal parameters, algorithm-independent bounds on achievable performance, sensor array processing, independent component analysis, and blind signal separation.

Dr. Tichavský served as an Associate Editor of the IEEE SIGNAL PROCESSING LETTERS from 2002 to 2004, and since 2005, he has served as an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING. In 2002, he received the *Otto Wichterle Award* from Academy of Sciences of the Czech Republic.

Zbyněk Koldovský (S'03–M'04) was born in Jablonec nad Nisou, Czech Republic, in 1979. He received the M.S. and Ph.D. degrees in mathematical modeling from Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University, Prague, Czech Republic, in 2002 and 2006, respectively.

Currently, he is with the Institute of Information Technology and Electronics, Technical University of Liberec, Liberec, Czech Republic. He has also been with the Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Prague, Czech Republic, since 2002. His main research interests are in audioprocessing, independent component analysis, and blind deconvolution.

Arie Yeredor (M'99–SM'02) received the B.Sc. (*summa cum laude*) and Ph.D. degrees in electrical engineering from Tel-Aviv University (TAU), Tel Aviv, Israel, in 1984 and 1997, respectively.

Currently, he is with the Department of Electrical Engineering—Systems, at TAU's School of Electrical Engineering, where he teaches courses in statistical and digital signal processing. He also holds a consulting position with NICE Systems Inc., Ra'anana, Israel, in the fields of speech and audio processing, video processing, and emitter location algorithms.

Dr. Yeredor is an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS and the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEM II, and is a member of the IEEE Signal Processing Society's Signal Processing Theory and Methods (SPTM) Technical Committee. He has been awarded the yearly Best Lecturer of the Faculty of Engineering Award at TAU five times.

Germán Gómez-Herrero was born in Zaragoza, Spain, in 1979. He received the M.Sc. degree in telecommunications engineering from the University of Zaragoza, Zaragoza, Spain and the M.Sc. degree in information technology from Tampere University of Technology, Tampere, Finland, in 2003. Currently, he is working towards the Ph.D. degree in biomedical signal processing at the Institute of Signal Processing, Tampere University of Technology, Tampere, Finland.

His research interests focus on machine learning, statistical pattern recognition, and classification and applications related with the analysis and processing of physiological signals.

Eran Doron was born in Israel in 1975. He received the B.Sc. and M.Sc. degrees (*magna cum laude*) in electrical engineering from Tel-Aviv University, Tel-Aviv, Israel, in 2002 and 2004, respectively.

Currently, he is an Algorithm Engineer in the field of wireless communication at the Amimon Ltd. His research interests include estimation theory and blind source separation.

A COMPARISON OF INDEPENDENT COMPONENT AND INDEPENDENT SUBSPACE ANALYSIS ALGORITHMS

Zbyněk Koldovský[1] *and Petr Tichavský*[2]

[1]Faculty of Mechatronic, Informatics and Interdisciplinary Studies, Technical University of Liberec
Studentská 2, 461 17 Liberec, Czech Republic, phone: + (420) 485 353 534, fax: + (420) 485 353 112
email: zbynek.koldovsky@tul.cz, web: http://itakura.ite.tul.cz/zbynek
[2]Institute of Information Theory and Automation, P.O.Box 18, 182 08 Prague 8, Czech Republic
phone: + (420) 266 052 292, fax: + (420) 286 890 300, email: tichavsk@utia.cas.cz
web: http://si.utia.cas.cz/Tichavsky.html

ABSTRACT

Recent advances in separation of convolutive mixtures of audio signals have shown that the problem can be successfully solved in time-domain in a multistep procedure including an application of some method of instantaneous independent component analysis (ICA) or independent subspace analysis (ISA), as one of the steps. In this paper we propose a test that allows a comparison of different ICA and ISA algorithms from this perspective. The test consists in evaluating separation of a pseudo-convolutive mixture of given independent signals. The mixture has features of real-world convolutive mixtures and of instantaneous mixtures simultaneously. We apply the proposed test to compare performance of several ICA and ISA algorithms in four different scenarios, taking in mind that suitability of the algorithms depends on properties of the separated signals.

1. INTRODUCTION

In this paper, we aim at comparing different ICA/ISA methods when applied to blind audio source separation (BASS), which is a popular discipline in recent decade due to emerging applications in multi-microphone systems. The goal of BASS is to separate simultaneously sounding audio sources that are mixed in a natural acoustical environment through the convolutive model

$$x_i(n) = \sum_{j=1}^{d} \sum_{\tau=0}^{M_{ij}-1} h_{ij}(\tau) s_j(n-\tau), \quad i = 1,\ldots,m, \qquad (1)$$

where $x_1(n),\ldots,x_m(n)$ are the observed signals on microphones, $s_1(n),\ldots,s_d(n)$ are the unknown original sources, and h_{ij}'s are source-microphone impulse responses each of length M_{ij}. The original sources can be estimated by passing the mixture through a separating (de-mixing) filter

$$\widehat{s}_i(n) = \sum_{j=1}^{m} \sum_{\tau=0}^{L-1} w_{ij}(\tau) x_j(n-\tau), \quad i = 1,\ldots,d. \qquad (2)$$

of a finite length L.

A popular way is to ground the separation on the assumption that the original sources are statistically independent. The solution of the problem is then based on methods related to the ICA [1]. However, original ICA methods assume instantaneous mixture, i.e., when $M_{ij} = 1$ for all i, j. The problem given by (1), therefore, needs to be transformed. This is usually done either in the frequency-domain [2] or in the time-domain [3]. In general, performance of the BASS algorithms can be evaluated e.g. by aid of the BSS_eval toolbox [4]. In this paper, we propose a special method of comparison of different ICA and ISA algorithms, with respect to their performance *inside* a *time-domain* BSS method.

In the time-domain methods, the convolution operation is written in terms of a vector/matrix product. In particular, the output of the separating filter in (2) corresponds to a direction in the subspace spanned by rows of an $mL \times (N_2 - N_1 + 1)$ matrix

$$\mathbf{X} = \begin{bmatrix} x_1(N_1) & \ldots & \ldots & x_1(N_2) \\ x_1(N_1-1) & \ldots & \ldots & x_1(N_2-1) \\ \vdots & & & \vdots \\ x_1(N_1-L+1) & \ldots & \ldots & x_1(N_2-L+1) \\ x_2(N_1) & \ldots & \ldots & x_2(N_2) \\ x_2(N_1-1) & \ldots & \ldots & x_2(N_2-1) \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ x_m(N_1-L+1) & \ldots & \ldots & x_m(N_2-L+1) \end{bmatrix}, \qquad (3)$$

where N_1 and N_2, $N_2 > N_1$, determine part of recorded signals used to define $\mathbf{X}$.

Time-domain BASS methods seek for such a linear transform that splits the row-space of $\mathbf{X}$ to independent subspaces so that each of them corresponds to a separated audio signal. To separate the subspaces, it is possible to use some algorithm for Independent Subspace Analysis (ISA) [5, 6, 7]. Another way is to apply one of large number of known ICA algorithms to estimate several one-dimensional components of each original source [8, 3, 9], and the subspaces are obtained by a suitable grouping (clustering) of the components [6, 9]. It was shown in [9] that under some condition, even quite short filters ($L = 10\ldots40$) can produce effective separation results.

Since the applied ICA/ISA algorithm is the hearth of the time-domain separation, a natural question is which one is suited best for that purpose. The objective evaluation of the decomposition of $\mathbf{X}$ is however an intricate problem due to

1. unpredictable performance limitations caused by the finite length of separating filters (the finite number of rows of $\mathbf{X}$), and

[0]This work was supported by Ministry of Education, Youth and Sports of the Czech Republic through the project 1M0572 and by Grant Agency of the Czech Republic through the projects 102/07/P384 and 102/09/1278.

2. an ambiguity of optimal solution, because the original signals can be retrieved up to unknown filterings [10].

Note that both the features affect the achieved signal-to-interference ratio [4].

In the following section, we propose a test that is designed so that the above problems are avoided, and the evaluation is done through a criterion tailored to the test. In Section 3, we discuss usefulness of several ICA models and select representative methods for experimental comparison by the proposed test that is described in Section 4. Sections 5-6 provide ranking of the methods and suggest conclusions.

2. TEST PROPOSAL

The main idea of the test is to define a source matrix of the original sources $s_1(n),\ldots,s_d(n)$ as

$$\mathbf{S} = \begin{bmatrix} s_1(N_1) & \ldots & \ldots & s_1(N_2) \\ s_1(N_1-1) & \ldots & \ldots & s_1(N_2-1) \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ s_d(N_1-L+1) & \ldots & \ldots & s_d(N_2-L+1) \end{bmatrix}. \quad (4)$$

The mixture is then simply given by

$$\mathbf{X} = \mathbf{AS}, \quad (5)$$

where $\mathbf{A}$ is a regular $dL \times dL$ matrix. The matrix can have the block Sylvester structure as it exists in the true convolutive mixtures [8, 3].

Unlike the true convolutive mixture in (3), the mixture in (5) can, in theory, be separated perfectly by $\mathbf{W} = \mathbf{A}^{-1}$. By contrast, the common feature is that ICA or ISA methods applied to (5) tend to produce arbitrarily filtered counterparts of $s_1(n),\ldots,s_d(n)$, because rows of $\mathbf{S}$ corresponding to delayed versions of the same source are *not* independent due to temporal structures of original (audio) sources.

2.1 Choice of the mixing matrix A

Most of ICA and ISA algorithms (all that were included in our comparative study) are equivariant. This means that outcome of the separation is essentially the same (up to the order of components or subspaces) if the input data are mixed by an arbitrary regular mixing matrix. It follows that it makes no difference if the mixing matrix in simulations has a certain structure or not. However, if someone wants to study a separation algorithm that relies on the special structure of the mixing matrix, a fair comparison would be obtained only if the mixing matrix has the same structure.

2.2 Grouping of components

Let ISA/ICA algorithms under the test be applied to the mixture $\mathbf{X}$. ISA algorithms produce d independent L-dimensional subspaces, which only have to be properly reordered to fit the original signal order. ICA algorithms yield one-dimensional components that have yet to be grouped. In our test, we do not want to let the choice of the grouping procedure interfere with the estimated quality of the separation. Therefore we resort to the optimum grouping of the components subject to the signal-to-interference ratio (SIR) both for ICA and ISA algorithms.

Consider the SIR of the jth separated component, denoted by $c_j(n)$, with respect to the ith source. Since $c_j(n)$ was obtained as the jth row of

$$\mathbf{C} = \widehat{\mathbf{W}}\mathbf{X} = \widehat{\mathbf{W}}\mathbf{AS} \stackrel{\text{def.}}{=} \mathbf{GS}, \quad (6)$$

it can be written as a linear combination of $s_i(n)$ and its time delays plus the remainder, which represents the interference. Thus, the SIR can be defined as

$$\mathrm{SIR}_j^i = \frac{\hat{\mathrm{E}}[\sum_{\ell=1}^{L} \mathbf{G}_{j,(i-1)L+\ell} s_i(n-\ell+1)]^2}{\hat{\mathrm{E}}[c_j(n) - \sum_{\ell=1}^{L} \mathbf{G}_{j,(i-1)L+\ell} s_i(n-\ell+1)]^2}, \quad (7)$$

where $\hat{\mathrm{E}}$ stands for the sample mean operator, and $\mathbf{G}_{j,k}$ are elements of the so-called gain matrix $\mathbf{G}$.

Now, for each source s_i we assign those L separated components c_j that have the largest SIR_j^i.

2.3 Criteria

Once we have the components assigned to the sources, we can judge quality of the separation. We propose two ways. First, we measure the distance of the true and estimated subspaces in terms of the angle of these subspaces in the vector space spanned by all rows of the matrix $\mathbf{S}$ in (5). In Matlab it is realized by the command `subspace`.

Second, we propose an alternative way which goes one step further towards the estimation of the source signals, using the time-shift structure of the matrix $\mathbf{S}$.

Let J_i denote a set of the indices of components that were assigned to the ith source. Then, an estimate of the ith source delayed by ℓ samples, i.e. of $s_i(n-\ell)$, can be obtained, avoiding unknown permutations in $\mathbf{G}$, through the *inverse* of $\mathbf{G}$ as

$$\widehat{s}_i^{\ell}(n) = \sum_{j \in J_i} (\mathbf{G}^{-1})_{(i-1)L+\ell,j} c_j(n), \quad (8)$$

for $\ell = 0,\ldots,L-1$, and these estimates of $s_i(n-\ell)$ can be combined together by simple time-shifting and averaging,

$$\widehat{s}_i(n) = \frac{1}{L} \sum_{\ell=1}^{L} \widehat{s}_i^{\ell}(n+\ell)\,. \quad (9)$$

The resultant reconstructed signal $\widehat{s}_i(n)$ is then written in the form signal-plus-interference, and the corresponding SIR yields the final criterion for the overall estimation of the ith source. Note that $\widehat{s}_i(n) = s_i(n)$ if and only if $\mathbf{G}$ is exactly block-diagonal (up to the order of its rows). Therefore, the SIR of $\widehat{s}_i(n)$ reflects the error of the block blind separation in a comprehensive way!

Another important point to note here is that the values SIR_j^i, $j \in J_i$, in (7) do *not* provide objective measures for evaluating the overall separation of $s_i(n)$ because of an unknown filtering of the respective components.

A third alternative of computing the SIR of the subspaces was advocated in [8]. It consists in applying a SIMO blind identification method to each subspace. This approach has the disadvantage that it introduces another source of error in the SIR computation.

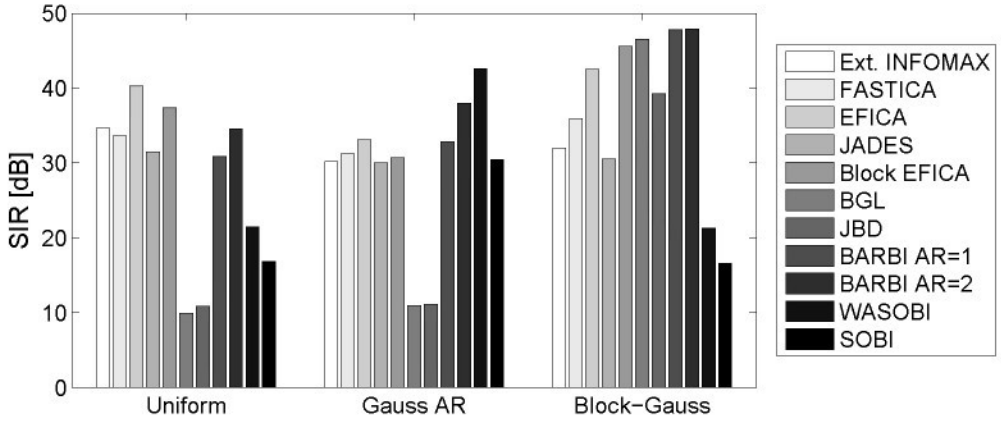


Figure 1: SIR of three separated artificial signals mixed via (4-5) with $L = 3$ averaged over 100 independent trials. Note that both settings of BARBI (first and second-order block-AR model) yield good results in this example.

3. REPRESENTATIVE METHODS

The main ICA algorithms for separation of instantaneous mixtures are based either on non-Gaussianity, distinct coloration (spectral diversity), or non-stationarity. While the first class uses higher-order statistics (nonlinear transforms) of the data, the other two classes are based on second-order statistics. Recently, combinations of the models have been considered, also [3, 13, 14, 21].

We note that the non-Gaussianity based methods tend to produce temporally whitened versions of $s_1(n), \ldots, s_d(n)$. The whitened signals are sometimes called the partial innovations. The reason is that the innovations of each sources and their mutual time-shifted copies are usually the most non-Gaussian signals that can be obtained by linear transformations of the data.

In our experiments, we consider Extended Infomax [11], FastICA [1] (the symmetric approach with "tanh" nonlinearity), EFICA [15], and SJADE [20] as the representatives for this class. Unlike the other methods, SJADE is an ISA algorithm.

Methods relying on nonstationarity divide mixed signals in non-overlapping segments of a given length, compute signal covariance matrices on each segment, and do an approximate joint diagonalization (AJD) of these matrices. These methods cannot separate sources having the same variance profiles. Hence, they cannot distinguish delayed copies of the same source as the delays are negligible compared to the length of segments. Separated components of (5) thus form clusters of arbitrarily filtered original sources, which is required for the separation. The class is represented by BGL [12] and JBD [7] algorithms. While BGL searches one-dimensional components, JBD is an ISA algorithm.

Methods relying on spectral diversity of the signals are based on (block-)AJD of cross-covariance matrices of mixed signals. In simulations, we shall consider the earlier and popular SOBI algorithm [16] and its weight-adjusted version WASOBI [17].

We will also consider methods that combine the basic ICA models, namely, Block EFICA [18] combining the non-Gaussianity with the nonstationarity, and the recently proposed BARBI algorithm [19, 21] combining the nonstationarity and the spectral diversity principles via block AR modeling of signals.

4. EXPERIMENTS

First, we present a simple example with three artificial signals obeying the basic ICA models: a non-Gaussian i.i.d. signal that is uniformly distributed, a stationary Gaussian process with AR coefficients $(1, 0.7)$, and a nonstationary block-Gaussian white signal whose each of four blocks has, respectively, the variance 1, 0.09, 0.01, and 1.21. These signals were used to form (4) with $L = 3$, which was mixed by a randomly generated mixing matrix via (5). Then, ICA methods were applied to separate the mixture and the resulting signals were evaluated by the proposed SIR.

Results of this example shown in Fig. 1 confirm characteristic features of the selected methods. Extended INFOMAX, FastICA, SJADE, EFICA, and Block EFICA succeeded to roughly separate all signals, because the 1^{st} signal is non-Gaussian, the 3^{rd} nonstationary signal behaves like being non-Gaussian, and one signal is allowed to be Gaussian, which is the 2^{nd} one. BGL and JBD failed to separate the 1^{st} and 2^{nd} signals since they have the same dynamic profiles. SOBI and WASOBI separated the 1^{st} and 3^{rd} signals poorly due to their similar spectra. Finally, BARBI succeeded to separate all signals since the 1^{st} and 2^{nd} signals have different spectra and different dynamics from that of the 3^{rd} signal.

In our main experiment, we did extensive testing of algorithms by separating the convolutive-like mixtures of two audio sources. Four different combinations of acoustical signals each of length 6.5s (10^5 samples) sampled at 16kHz were considered. In two scenarios, we mixed a male and a female speech and two speeches of the same male speaker, respectively, which stands for the situation where speakers' voices have different and similar spectra. In the fourth and third scenario, the male speech was mixed with a musical signal: First, with a long synthesizer tone having almost static variance, and, second, with a piece of a rhythmic music.

To simulate Monte-Carlo trials, we used the method of sliding time-window gradually shifted throughout the whole recordings. In each trial, the time-window of length 8000

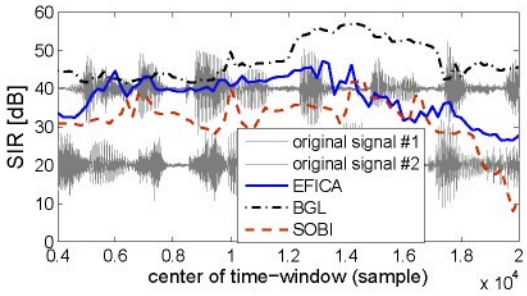


Figure 2: Example of course of the SIR averaged over two separated signals.

samples (0.5s) was shifted by 200 samples (12.5ms, i.e., there are 461 trials in each scenario), and the matrix (4) was constructed from the corresponding segment of signals with $L = 10$ and multiplied by a random mixing matrix. Then, the mixture was separated by the ICA and ISA methods mentioned in the previous section, with the following parameters: Block EFICA, BGL, JBD and BARBI had the number of blocks set to 40 (so that each block had the length 200). The methods based on spectral diversity, computed the separation by AJD of 11 covariance matrices with time lags $0, 1, \ldots, 10$.

Due to lack of space we present results in terms of the SIR only; results obtained by angles between subspaces were similar. An example of the resulting course of SIR is shown in Fig. 2.

The evolutions of resulting SIRs are indicative of the behavior of respective algorithms when signals are changing in time. Therefore, we use the three following characteristics of the SIR for evaluation: (*A*) the mean value, (*B*) the standard deviation, and (*C*) the mean of absolute value of variation, which is the difference between SIRs achieved in two successive time-windows. These characteristics of SIR are shown in Table 1 in the form $A \pm B(C)$.

Note that "good" performance must be characterized by continuous behavior of the resulting SIR in time. The range of SIR corresponds with the standard deviation *B*, and the speed of changes is reflected by the mean variation *C*. Higher value of the latter criterion signifies unstable performance. Conversely, small *C* and *B* means stable performance that is less dependent on signal characteristics.

5. DISCUSSION

5.1 Methods based on non-Gaussianity

Performances of algorithms using non-Gaussianity (INFOMAX, FastICA, EFICA, SJADE) appear to be not the best of all algorithms, but are quite good in all scenarios. EFICA slightly outperforms the other algorithms (INFOMAX, FastICA, and SJADE) thanks to being more advanced. All these four algorithms were outperformed by Block EFICA, which, in addition to the non-Gaussianity, utilizes non-stationarity of the signals as well.

5.2 Methods based on non-stationarity

These methods gave the best separation results in our study. Among them, BGL and JBD are based on non-stationarity only, BARBI combines it with the spectral diversity. Here, BGL and BARBI with AR order 1 appear to be the most successful algorithms. It is interesting to compare the results in Table 1 with results in [19] that deals with a separation of an *instantaneous* mixture of speech signals. In the latter study, BARBI was a clear winner, outperforming the other algorithms (including the BGL) by several dB. In this comparative study, both methods give similar results. Hence we can see that there is a qualitative difference between the instantaneous mixtures and the pseudo-convolutive mixtures. BARBI with AR order 2 was less successful both in our study and in separating the instantaneous mixtures of speech signals [19].

5.3 Methods based on spectral diversity

We observe that WASOBI fails in many trials in all scenarios, the results of SOBI are stable, moreover, SOBI yields surprisingly good results in the third scenario. This is another example showing the difference between instantaneous and pseudo-convolutive mixtures, because WASOBI is normally known to outperform SOBI in separating instantaneous mixtures [17]. In order to explain the failure of WASOBI, we note that blocks of cross-covariance matrices of (4) are *not* diagonally dominant except for the zero-lag cross-covariance. Therefore the AJD procedure in WASOBI might terminate at transformed matrices that are "more diagonal" in a sense but "less block-diagonal" than they should be.

5.4 Comparison of ICA and ISA algorithms

Our comparative study does not show any clear advantage of subspace (ISA) algorithms (JBD, SJADE) over the ICA algorithms. In order to make sure that the difference between the algorithms is not only in the separation criteria, we also compared performance of the ISA algorithms with their ICA versions. The ICA versions were obtained by setting the subspace dimensions equal to one. The separation results of the ISA algorithms and their ICA variants were approximately the same; the truly subspace algorithms were only faster.

6. CONCLUSION

We have proposed a method of comparing performance of different ICA and ISA methods in time-domain separation of convolutive mixtures of audio sources. In our test, the best separation results were obtained by the BGL and BARBI algorithms. Note, however, that the results depend, in general, on properties of the to-be separated signals.

REFERENCES

[1] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, Wiley-Interscience, New York, 2001.

[2] H. Sawada, R. Mukai, S. Araki, and S. Makino, "A robust and precise method for solving the permutation problem of frequency-domain blind source separation," *IEEE Trans. Speech Audio Processing*, vol. 12, no. 5, pp. 530-538, 2004.

[3] H. Buchner, R. Aichner, and W. Kellermann, "A Generalization of Blind Source Separation Algorithms for Convolutive Mixtures Based on Second-Order Statistics," *IEEE Trans. on Speech and Audio Processing*, Vol. 13, No. 1, pp. 120-134, Jan. 2005.

[4] E. Vincent, R. Gribonval, and C. Févotte, "Performance Measurement in Blind Audio Source Separation", *IEEE Trans. on Speech and Audio Processing*, 14(4):1462-1469, July 2006.

METHOD	SIR [dB] female speech male speech #1	male speech #2 male speech #1	music #1 male speech #1	music #2 male speech #1	average comp. time per trial (secs)
Ext. INFOMAX	29.6±7.8(1.9) 30.4±7.8(1.8)	31.3±6.0(1.8) 30.3±6.1(1.5)	30.1±7.6(2.4) 28.2±6.3(1.5)	31.5±8.1(1.7) 31.4±8.1(1.8)	13.5
FastICA	30.4±6.6(1.3) 31.8±6.4(1.5)	31.8±5.8(1.5) 30.6±5.5(1.4)	29.4±10.6(*6.5*) 31.4±5.0(1.7)	33.6±6.5(1.7) 33.9±4.7(1.4)	0.65
EFICA	32.7±7.0(1.8) 36.9±6.5(2.2)	34.1±6.3(1.9) 34.4±6.1(2.0)	35.9±6.2(2.8) 33.4±6.1(1.5)	34.6±5.5(1.8) 36.7±6.1(1.7)	1.06
SJADE	25.3±6.3(1.5) 26.4±6.5(1.7)	27.9±6.2(1.7) 26.4±7.0(1.7)	33.4±6.1(2.2) 26.8±4.5(1.1)	28.4±5.0(1.3) 28.3±4.6(1.3)	6.0
Block EFICA	35.0±8.4(2.2) 36.2±6.8(2.2)	36.4±7.9(2.0) 33.8±7.1(1.8)	34.2±6.9(3.1) **35.1**±8.2(2.1)	35.2±6.1(1.8) 39.0±8.3(1.8)	2.98
BGL	**41.2**±7.6(1.2) **43.8**±5.2(1.1)	**41.8**±8.5(1.3) **40.1**±6.3(1.1)	36.8±7.3(2.2) **37.1**±10.0(1.2)	**40.8**±5.8(1.1) **42.5**±8.1(1.1)	0.046
JBD	27.8±8.0(1.1) 30.7±6.6(1.4)	30.6±7.6(1.4) 29.1±7.4(1.2)	**38.5**±7.0(2.0) 29.4±7.9(1.1)	33.8±8.0(1.1) 31.5±7.1(1.3)	10.68
BARBI AR=1	**42.3**±8.2(1.4) **42.5**±7.1(1.6)	**42.4** ±7.3(2.2) **41.4**±7.5(2.2)	34.9±7.6(4.8) 16.8±13.8(*6.9*)	**38.7**±6.5(2.1) **43.9**±9.6(1.8)	0.059
BARBI AR=2	15.0±15.8(*15.0*) 12.8±14.7(*14.4*)	17.4±14.6(*15.0*) 13.1±13.4(*14.8*)	7.8±10.8(*11.4*) 3.6±10.1(*10.7*)	20.6±16.0(*15.1*) 20.1±13.3(*12.5*)	0.092
WASOBI	29.2±10.2(*9.3*) 27.0±11.7(*11.0*)	25.1±9.4(*9.5*) 23.3±10.8(*10.1*)	28.4±14.9(*16.5*) 9.4±12.0(*13.3*)	33.2±11.8(*10.1*) 32.0±10.2(*7.6*)	0.17
SOBI	25.9±7.3(2.0) 25.2±7.1(2.1)	21.9±8.0(2.3) 20.7±7.6(2.2)	**46.0**±5.9(2.1) 32.4±4.0(1.0)	37.5±6.0(1.6) 33.1±4.7(1.3)	3.24

Table 1: Separation SIR obtained by the ICA and ISA algorithms in four pseudoconvolutive mixtures. The best two mean values are written in bold, and the mean variations above 5dB are written in italic. The test was running in Matlab^{TM} v7.2 on a PC with i7-920 processor, 2.66GHz, 3GB RAM.

[5] L. De Lathauwer, B. De Moor, J. Vandewalle, "Fetal Electrocardiogram Extraction by Blind Source Subspace Separation," *IEEE Trans. Biomedical Engineering*, Vol. 47, No. 5, pp. 567-572, May 2000.

[6] K. Abed-Meraim and A. Belouchrani, "Algorithms for Joint Block Diagonalization," *Proc. of EUSIPCO 2004*, pp. 209212, Vienna, Austria, 2004.

[7] C. Févotte and F. J. Theis, "Orthonormal approximate joint block-diagonalization," *Technical Report GET/Tlcom Paris 2007D007*, 2007.

[8] C. Févotte, A. Debiolles, and C. Doncarli. "Blind separation of FIR convolutive mixtures: application to speech signals," *Proc. 1st ISCA Workshop on Non-Linear Speech Processing*, 2003.

[9] Z. Koldovský, P. Tichavský, "Time-domain blind audio source separation using advanced component clustering and reconstruction," *Proc. of HSCMA 2008*, Trento, Italy, pp. 216-219, May 2008.

[10] K. Matsuoka and S. Nakashima, "Minimal distortion principle for blind source separation", *Proc of ICA-2001*, pp. 722-727, 2001.

[11] T-W. Lee, M. Girolami and T. J. Sejnowski. "Independent Component Analysis using an Extended Infomax Algorithm for Mixed Sub-Gaussian and Super-Gaussian Sources", *Neural Computation*, Vol.11(2): 417-441, 1999.

[12] D-T. Pham and J-F. Cardoso. "Blind separation of instantaneous mixtures of non stationary sources", *IEEE Trans. Signal Processing*, pp. 1837-1848, vol. 49, no. 9, 2001.

[13] U. Manmontri and P. A. Naylor, "A Class of Frobenius Norm-Based Algorithms Using Penalty Term and Natural Gradient for Blind Signal Separation," *IEEE Trans. on Audio, Speech and Language Processing*, 16(6):1181-1193, Aug. 2008.

[14] S. A. Cruces-Alvarez and A. Cichocki, "Combining blind source extraction with joint approximate diagonalization: thin algorithms for ICA," *in Proc. of ICA 2001*, San Diego, pp. 740-745, Dec. 2001.

[15] Z. Koldovský, P. Tichavský and E. Oja, "Efficient Variant of Algorithm FastICA for Independent Component Analysis Attaining the Cramér-Rao Lower Bound", *IEEE Trans. on Neural Networks*, Vol. 17, No. 5, Sept 2006.

[16] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, and E. Moulines, "A blind source separation technique using second-order statistics," *IEEE Trans. Signal Processing*, vol. 45, pp. 434-444, Feb. 1997.

[17] P. Tichavský, E. Doron, A. Yeredor and J. Nielsen, "A Computationally Affordable Implementation of An Asymptotically Optimal BSS Algorithm for AR Sources", *Proc. EUSIPCO-2006*, Florence, Italy, Sep. 2006.

[18] Z. Koldovský, J. Málek, P. Tichavský, Y. Deville, and S. Hosseini, "Extension of EFICA Algorithm for Blind Separation of Piecewise Stationary non Gaussian Sources", in *Proc. ICASSP 2008*, Las Vegas, Nevada, USA, pp. 1913-1916, April 2008.

[19] P. Tichavský, A. Yeredor, and Z. Koldovský, "A Fast Asymptotically Efficient Algorithm for Blind Separation of a Linear Mixture of Block-Wise Stationary Autoregressive Processes", *Proc. of ICASSP 2009*, Taipei, April 2009.

[20] F. J. Theis, "Towards a general independent subspace analysis," *Advances in Neural Information Processing Systems 19*, pp. 1361-1368, 2007.

[21] P. Tichavský and A. Yeredor, "Fast Approximate Joint Diagonalization Incorporating Weight Matrices, " *IEEE Transactions of Signal Processing*, vol. 57, no.3, pp. 878-891, March 2009.

Part II

Convolutive Mixtures of Audio Signals

Chapter 5

Introduction

Blind separation of audio signals that are simultaneously sounding in a natural environment has many potential applications. It allows to suppress noise, jamming signals or other negative effects such as the reverberation from a targeted signal. Blind processing of signals is popular since no prior knowledge about them or about mixing conditions is required. There are also many emerging applications, especially in speech processing [44], such as automatic speech recognition, hands-free and distant-talking speech communication, human/machine interactions, music, etc.

However, the problem is difficult, since acoustical signals are reflected by walls and other obstacles, and their speed of propagation must be taken into account even when the sampling frequency of the measuring device is very small ($<$ 8 kHz). Therefore, the mixed signals measured on microphones contain delayed and colored versions of the original signals due to the reverberation. The mixing system is therefore the convolutive one [45], which was already mentioned in Section 1.2.2.

5.1 Problem Statement

5.1.1 Convolutive Model

The goal is to blindly retrieve d original audio signals (sources) from their convolutive mixtures recorded by m microphones. The ith mixed signal is equal to

$$x_i(n) = \sum_{j=1}^{d} \sum_{\tau=0}^{M_{ij}-1} h_{ij}(\tau) s_j(n-\tau), \qquad i = 1, \ldots, m, \tag{5.1}$$

where $s_1(n), \ldots, s_d(n)$ are the unknown original (audio) signals, h_{ij} denotes the impulse responses between the jth source and the ith microphone whose length is

M_{ij}. The responses characterize propagation of sound in the recording room and are unknown. It will be assumed that the system is time-invariant. This requires that positions of all objects in the mixing environment, including sources and microphones, do not change within recording. The length of recorded data will be denoted by N.

The mixing model (5.1) can be written in the form

$$x_i(n) = \sum_{j=1}^{d} s_j^i(n) \tag{5.2}$$

where

$$s_j^i(n) = \sum_{\tau=0}^{M_{ij}-1} h_{ij}(\tau) s_j(n-\tau) \tag{5.3}$$

is called the *response* of the jth source on the ith microphone. It corresponds to the jth original signal how it is heard in the position of the ith microphone when all the other sources are silent. It also corresponds to a filtered version of $s_j(n)$ by the filter represented by h_{ij}.

To separate the signals in a linear way, one has to find a MIMO filter whose inputs are the mixed signals from microphones, and its outputs are the separated (estimated) signals. For instance, an estimate of the jth original signal is

$$\widehat{s}_j(n) = \sum_{i=1}^{m} \sum_{\tau=0}^{L_i-1} w_{ji}(\tau) x_i(n-\tau), \qquad j = 1, \ldots, d, \tag{5.4}$$

where the filters w_{ji}, each of the length L_i, parametrize the separating system. The number of parameters to-be estimated is thus equal to $d(\sum_{i=1}^{m} L_i)$. Since filters that achieve (almost) exact inversion of (5.1) are usually long (L_is are large), the task to find them is usually very difficult.

5.2 Separation through ICA

There are several approaches for finding the separating filters in a blind way. This part of the thesis focuses on the use of ICA. The only assumption about the original signals, therefore, is that they are independent. The goal is to restore the independence of signals by transforming the mixed signals. Here, it will be assumed that the number of microphones is the same as the number of sources, $m = d$, so the mixing system is determined.

5.2.1 Indeterminacies

As follows from the ambiguity of ICA, which was discussed in Section 2.1.1, the original order, signs and scales of signals cannot be retrieved. Nevertheless, the solution is yet more ambiguous in case of the convolutive mixing model. The reason is that the original signals can each be arbitrarily filtered while being still mutually independent. For instance, the responses of signals on a given microphone defined by (5.3) are independent whenever $s_1(n), \ldots, s_d(n)$ are independent. It follows that, based on the independence assumption, signals can be retrieved up to their original order and spectra.

Minimal Distortion Principle

The fact that the original spectra of signals cannot be determined entails an important problem when separating audio signals. Fortunately, properly defined spectra are (only) those of the responses (5.3), because they are observed through microphones. It is therefore meaningful to aim at estimating the responses, because the original spectra of sources cannot be retrieved without their prior knowledge.

Besides the advantage that the only ambiguity of this solution is the unknown order of signals, there are also further appealing features that were described in [46]. The approach is called Minimal Distortion Principle (MDP).

5.2.2 From Convolutive to Instantaneous Mixing Model

The original ICA algorithms assume the instantaneous mixing model, which is equivalent to (5.1) only if $M_{ij} = 0$ for all i and j. To enable the utilization of ICA algorithms designed for instantaneous mixtures in processing of convolutive mixtures, the model needs to be transformed first.

There are two basic approaches to transform the convolutive model. The *frequency-domain* (FD) approach comes from the fact that the Fourier transform converts the convolution operation in (5.1) between the original signals and room impulse responses to the ordinary multiplication of their Fourier images. Then, (5.1) is represented by instantaneous mixtures

$$\mathbf{X}(\theta) = \mathbf{H}(\theta)\mathbf{S}(\theta), \tag{5.5}$$

one for each frequency θ. Here, the elements of $\mathbf{X}(\theta)$ and $\mathbf{S}(\theta)$ are equal to the Fourier transforms of the mixed and original signals, respectively, and $\mathbf{H}(\theta)$ is the $d \times d$ matrix whose ijth element is the Fourier transform of $h_{ij}(n)$ at θ. Consequently, complex-domain ICA methods could be applied separately to each of the instantaneous mixture.

The second basic approach works with signals directly in *time-domain* (TD). Generally, it can be described by defining a data matrix constructed of the samples of mixed signals. It has a special structure different from that of $\mathbf{X}$ defined in (1.1). The de-mixing transform is searched by an ICA algorithm as a separating matrix, because the ordinary matrix multiplication conveys the convolution operation thanks to the special structure of the data matrix.

Both approaches will be described in more details in the next chapter.

5.3 Performance Measurement

It is difficult to evaluate the quality of separated audio signals in a definite way using a single-valued criterion. The problem arises due to the ambiguity of the spectra of separated signals. Standard measures like the Signal-to-Interference ratio are not invariant to the filtering of signals [48].

On one hand, the Minimum Distortion principle defines the target spectra of separated signals properly. On the other hand, it does not imply that these signals provide the best solution in terms of other aspects like intelligibility, for example. (The responses (5.3) are reverberated due to room acoustic.) Therefore, further criteria reflecting other aspects of the quality should be introduced.

5.3.1 Signal Acquisition

To evaluate the separation, the mixing system or the original signals must be known. However, it is difficult or even impossible to measure the room impulse responses of acoustical environments with the required precision. On the other hand, the original signals are easily provided in a testing scenario. The problem is that the room impulse responses are usually very long, so it is difficult to decompose any measured signal into the sum of target signal and interference.

A much easier way is to record each source separately, that is, when the other sources are silent, that is, to record the responses (5.3) separately. Then, the mixed signals can be obtained as the sum of responses according to (5.2).

The alternative and more realistic way is to record the mixed signals, that is, when all sources are active. The latter approach requires perfect time synchronization of recordings so that the recorded responses have no delay in comparison with how they occur in the mixed signals. In other words, the relation (5.2) between recorded signals must be valid up to some noise. Then, it follows that the difference between the left and right side of (5.2) provide information about the additive noise, which will be specified below.

5.3.2 Signal-to-Interference Ratio

Let the order of separated sources be the same as the original one. Next, let $\widehat{s}_j^i$ be the estimated response of the jth source on the ith microphone, and w_k, $k = 1, \ldots, d$, be the separating filters. The estimated response can be written as

$$\widehat{s}_j^i(n) = \sum_{k=1}^{d} \{w_k \star x_k\}(n) \tag{5.6}$$

where $\star$ denotes the convolution operation. Using (5.2), the signal can be decomposed as

$$\widehat{s}_j^i(n) = \sum_{k=1}^{d} \{w_k \star s_j^k\}(n) + \sum_{\substack{\ell=1 \\ \ell \neq j}}^{d} \sum_{k=1}^{d} \{w_k \star s_\ell^k\}(n), \tag{5.7}$$

where the first term corresponds to the contribution of the target signal, and the second term is the residual interference. Signal-to-Interference Ratio (SIR) of $\widehat{s}_j^i(n)$ is thus defined as

$$\mathrm{SIR}_j^i = \frac{\widehat{\mathrm{E}}\left[\sum_{k=1}^{d} \{w_k \star s_j^k\}(n)\right]^2}{\widehat{\mathrm{E}}\left[\sum_{\substack{\ell=1 \\ \ell \neq j}}^{d} \sum_{k=1}^{d} \{w_k \star s_\ell^k\}(n)\right]^2}, \tag{5.8}$$

where $\widehat{\mathrm{E}}$ stands for the sample mean operator. For further use, the average SIR of the jth source will be denoted by SIR_j, i.e.

$$\mathrm{SIR}_j = \frac{1}{d} \sum_{i=1}^{d} \mathrm{SIR}_j^i. \tag{5.9}$$

As pointed above, SIR is *not* invariant to a filtering of the evaluated signal. Since the filtering can focus an arbitrary frequency band where the signal is or is not active (and similarly the other signals), the SIR can be made arbitrarily high or low by the filtering. Consequently, it should be noted that the criterion does not reflect the audience quality of the signal.

5.3.3 Signal-to-Distortion Ratio

A criterion that reveals whether the spectrum of a separated response corresponds with that of the original response is the Signal-to-Distortion ratio (SDR). For $\widehat{s}_j^i(n)$, it is defined in [49] as

$$\mathrm{SDR}_j^i = \max_{\alpha \in \mathbb{R}, \tau \in \mathbb{Z}} \frac{\widehat{\mathrm{E}}\left[s_j^i(n)\right]^2}{\widehat{\mathrm{E}}\left[s_j^i(n) - \alpha \sum_{k=1}^{d} \{w_k \star s_j^k\}(n - \tau)\right]^2}. \tag{5.10}$$

The maximizations over α and τ guarantee that the criterion is invariant to the scale and delay of separated signals. The denominator, called the *distortion*, is equal to zero if the first term in (5.7) is equal to $s_j^i(n)$. It is highly sensitive to the filtering of the estimated response $\widehat{s}_j^i(n)$, as required. On the other hand, the criterion neglects any the residual interference in the evaluated signal. The alternative definition is therefore

$$\mathrm{SDR}_j^i = \max_{\alpha\in\mathbb{R},\tau\in\mathbb{Z}} \frac{\widehat{\mathrm{E}}\left[s_j^i(n)\right]^2}{\widehat{\mathrm{E}}\left[s_j^i(n) - \alpha\widehat{s}_j^i(n-\tau)\right]^2}. \tag{5.11}$$

This criterion is easier to implement, but it might be too sensitive to the filtering and interference in $\widehat{s}_j^i(n)$.

5.3.4 Signal-to-Noise Ratio

Assume that the second approach described in Section 5.3.1 for acquiring the recordings was done. It means that the mixed signals and the responses were obtained independently. In practice, each recording contains certain amount of additive noise. Using the difference between the recordings, the noise level can be evaluated.

It is natural to assume that each recording is disturbed by a Gaussian white noise with zero mean and variance σ_i^2 where i is the index of microphone. Let $\mathrm{n}_j^i(n)$ be the noise that was added to the recording of $s_j^i(n)$, and $\mathrm{n}_i(n)$ be the noise added to $x_i(n)$. The signals $\mathrm{n}_j^i(n)$ and $\mathrm{n}_i(n)$ can be seen as independent realizations of the same Gaussian noise with the variance σ_i^2. The difference between the mixed signals and sum of responses on the ith microphone thus gives

$$x_i(n) - \sum_{k=1}^{d} s_k^i(n) = \mathrm{n}_i(n) - \sum_{k=1}^{d} \mathrm{n}_k^i(n). \tag{5.12}$$

Using the independence of noise realizations, the variance of (5.12) is $(d+1)\sigma_i^2$. By computing the sample variance of (5.12), σ_i^2 can be estimated as

$$\widehat{\sigma}_i^2 = \frac{1}{d+1}\widehat{\mathrm{E}}\left[x_i(n) - \sum_{k=1}^{d} s_k^i(n)\right]^2. \tag{5.13}$$

Consequently, the Signal-to-Noise Ratio (SNR) of the recorded signals $s_j^i(n)$ and $x_i(n)$ can, respectively, be defined as

$$\mathrm{SNR}_j^i = \frac{\widehat{\mathrm{E}}[s_j^i(n)]^2 - \widehat{\sigma}_i^2}{\widehat{\sigma}_i^2} \tag{5.14}$$

$$\mathrm{SNR}_i = \frac{\widehat{\mathrm{E}}[x_i(n)]^2 - \widehat{\sigma}_i^2}{\widehat{\sigma}_i^2}. \tag{5.15}$$

Chapter 6

Methods

6.1 Frequency-Domain Approaches

Separation in the frequency-domain [47] works with the signals obtained from microphones transformed by the Fourier transform. By applying the Discrete-Time Fourier Transform (DTFT) to the convolutive model (5.1), the problem is transformed to complex-valued instantaneous mixtures (5.5), one for each frequency.

In practice, where finite data are processed, the DTFT is replaced by Discrete Fourier Transform (DFT) that is computed in a sliding time-windows. This transform is commonly known as the Short-Time Fourier Transform (STFT). In this way, K instantaneous mixtures

$$\mathbf{X}(\theta_k) = \mathbf{H}(\theta_k)\mathbf{S}(\theta_k), \qquad k = 1, \ldots, K, \tag{6.1}$$

are obtained, where K corresponds to the half of the length of DFT. The number of available samples of $\mathbf{X}(\theta_k)$ is equal to the number of time-windows where DFT is computed.

Now, an ICA algorithm designed to work with complex-valued signals could be applied to find the right sides of (6.1). Once $\mathbf{H}(\theta_k)$ and $\mathbf{S}(\theta_k)$ are estimated for all k, the separated signals are obtained by the inverse STFT. The minimal distortion principle is respected when each separated signal, that is, a row of $\mathbf{S}(\theta_k)$, is multiplied by the corresponding column of $\mathbf{H}(\theta_k)$. Specifically, for the jth signal, the rows of

$$\mathbf{H}(\theta_k)_{:,j}\,\mathbf{S}(\theta_k)_{j,:},$$

which is a matrix of the same size as $\mathbf{X}(\theta_k)$, contain Fourier coefficients of the responses of the jth source for θ_k. Here, the subscripts $\cdot_{:,j}$ and $\cdot_{j,:}$ denote, respectively, the jth column and row of a given matrix.

The advantage of the FD approach is that the dimension of (6.1) is d, which is reasonably low, and the ICA decomposition is fast. On the other hand, ICA

requires sufficient length of data K due to the accuracy of separation. Therefore, sufficiently long recordings should be provided.

The main difficulties of the FD approach consist in the indeterminacy of order of original signals, which cause that the order of columns of $\mathbf{H}(\theta_k)$ and of rows of $\mathbf{S}(\theta_k)$ is random (thus different) for each θ_k. For the sake of successful separation, it is necessary that the separated frequency components corresponds to the same source. Otherwise, the sources would be mixed again after the inverse STFT. The goal is thus to re-order the separated frequency components. The task is named the *permutation problem*.

6.1.1 Approaches to Solve the Permutation Problem

The permutation problem has been addressed by many researches, and various methods have been proposed to solve it. Some earlier methods utilize the fact that spectra of audio signals should be smooth. The frequency components are permuted so that a criterion of smoothness is optimized. For instance, in [50], the criterion requires that the length of separating filters in time-domain is limited. The reason is that FIR filters have polynomial spectral characteristics, so they are smoother when they are "short", that is, they have as small number of free parameters as possible.

Another way is to exploit the dependency between separated components across frequencies. Many audio signals such as speech or tones of musical instruments embody the dependencies thanks to being abundant in harmonics. The dependencies can be measured by correlations [51] or even with the aid of higher-order statistics.

Robust approaches utilize the directivity of sources through various propagation models [52]. In general, the models try to extract some hidden information following from the structure of $\mathbf{H}(\theta_k)$ or of its inverse matrix. For instance, the time difference of arrival (TDOA) of sources can be utilized this way [51].

A recursive computation of the ICA is utilized in [58]. Also this approach relies on the smoothness of frequency characteristics of separating filters. ICA in a given frequency bin is computed by doing one optimization step of an iterative ICA algorithm initialized by the resulting transform computed in the previous frequency bin. The recursion can be repeated several times to achieve the convergence. A smooth behavior of the ICA algorithm is needed.

6.2 Time-Domain Approaches

TD approaches consists in that the convolution operation is done through operations with data vectors or matrices of a special structure. Typically, a data matrix

of mixed signals $\mathbf{X}$ is defined as

$$\mathbf{X} = \begin{bmatrix} x_1(N_1) & \dots & \dots & x_1(N_2) \\ x_1(N_1-1) & \dots & \dots & x_1(N_2-1) \\ \vdots & \vdots & \vdots & \vdots \\ x_1(N_1-L+1) & \dots & \dots & x_1(N_2-L+1) \\ x_2(N_1) & \dots & \dots & x_2(N_2) \\ x_2(N_1-1) & \dots & \dots & x_2(N_2-1) \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ x_d(N_1-L+1) & \dots & \dots & x_d(N_2-L+1) \end{bmatrix}, \tag{6.2}$$

where N_1 and N_2, $1 \leq N_1 < N_2 \leq N$, determine a segment of recordings that is used for computations, and L is a free integer parameter.

Consider a $Ld \times 1$ vector $\mathbf{w}$ whose elements are equal to the entries of a MISO filter of the length L

$$\mathbf{w} = [w_1(0) \dots w_1(L-1)\, w_2(0) \dots \dots w_d(L-1)]^T.$$

The resulting signal from a multichannel convolutive operation

$$\sum_{k=1}^{d} \{w_k \star x_k\}(n), \qquad n = N_1, \dots, N_2,$$

can be represented (up to starting and ending effects) as the resulting row-vector of the vector/matrix product

$$\mathbf{w}^T\mathbf{X}, \tag{6.3}$$

which holds thanks to the structure of $\mathbf{X}$. Hence, time-domain methods can be defined as those operating with $\mathbf{X}$ defined through (6.2). The subspace of $\mathbb{R}^{N_2-N_1+1}$ spanned by rows of $\mathbf{X}$ will be called the *observation space* and $\mathbf{X}$ the *observation matrix*. It contains all signals that can be obtained from the mixed signals by a MISO filtering by filter of the length L.

MISO filters that output independent signals could be searched as rows of a separating matrix computed by an ICA algorithm that is applied to $\mathbf{X}$. In other words, ICA can be used to give the decomposition $\mathbf{X} = \mathbf{AS}$, where $\mathbf{S}$ are independent components that can be used to reconstruct separated signals.

The decomposition can be either complete or partial, which corresponds, respectively, to finding the whole regular $\mathbf{A}$ or some of its columns only. In the former case, there are Ld independent components, but only d original signals should be retrieved. Therefore, the components should form d independent subspaces, which bears relation to ISA (Chapter 3). In the partial decomposition, each original signal is represented by one component.

Independent components usually do not provide useful estimates of audio signals by themselves, because their spectra are flat or randomly colored. A reconstruction procedure that retrieves the original spectra is therefore needed. A different situation is, for example, when communication i.i.d. finite-alphabet signals are to-be separated. Such sources are temporarily white, which is a prior knowledge about their spectra (this task is named the blind equalization [53]).

In partial-decomposition methods, the initialization determines what component is going to be found. However, it is never known in advance which component of what source is found unless any prior knowledge is given. To avoid findings of two or more components of the same source, an efficient constraint should be applied. An example of method for blind separation of audio sources doing the partial decomposition is [54], where the paraunitary constraint on separating filters is used. The reconstruction of signal spectra relies on multichannel spatio-temporal pre-whitening of the input signals prior to the separation and the inverse operation after the separation.

The complete decomposition is computationally expensive when L is very large. Many methods apply some constraint to allow the computation of longer filters. An often used constraint is that the matrix $\mathbf{A}$ has a special structure, for example block-Toeplitz or block-Sylvester [55]. An unconstrained decomposition is done by the T-ABCD algorithm that is briefly described by the following subsection.

6.2.1 The T-ABCD algorithm

The abbreviation "T-ABCD" means Time-domain Audio sources Blind separation based on the Complete Decomposition of the observation space. It reflects the fact that the ICA is used to decompose the whole observation matrix into all independent components. A prototype of T-ABCD was first published in [c7]. A similar concept was proposed by Jafari et al. in [59].

Following the minimal distortion principle, T-ABCD estimates microphone responses of the original signals by the four following steps:

1. The matrix $\mathbf{X}$ is formed, basically, according to (6.2).

2. $\mathbf{X}$ is decomposed into independent components by an ICA algorithm. Let the separating (decomposing) matrix be $\mathbf{W}$, which is a $M \times M$ matrix $\mathbf{W}$ where $M = dL$.

3. The components (rows of) $\mathbf{C} = \mathbf{WX}$ are grouped into clusters so that each cluster contains components that correspond to the same original source. Components of a cluster form an independent subspace. Each cluster, in fact, represents an original source.

4. For each cluster, only components of the cluster are used to estimate microphone responses of a source.

Steps 2 and 3 can be done jointly through ISA (Chapter 3). The drawback of ISA is that the dimensions of independent subspaces must be known in advance. In this respect, the two-step procedure seems to be more robust and flexible.

A related issue is the choice of an appropriate ICA algorithm applied in step 2. It was shown experimentally that most ICA methods can be applied successfully, up to some methods using the spectral diversity principle thus doing the joint diagonalization of cross-covariance matrices (WASOBI, BARBI). Since the problem is, in fact, an ISA problem, the joint block-diagonalization of matrices is needed rather than the diagonalization. Since cross-covariance matrices are not diagonally dominant, their joint diagonalization might yield unwanted solutions. On the other hand, this problem does not arise with methods based on the non-stationarity where the requirement of the block-diagonality is not in conflict with the diagonality. The experimental observations are in accordance with the results of [c15], which also serve as benchmarks for choosing the most efficient algorithms in terms of accuracy and computational burden (BGSEP, EFICA).

The details of step 4 are as follows. For the kth cluster, a matrix

$$\widehat{\mathbf{S}_k} = \mathbf{W}^{-1}\mathrm{diag}[\lambda_1^k, \ldots, \lambda_M^k]\,\mathbf{W}\,\mathbf{X} = \mathbf{W}^{-1}\mathrm{diag}[\lambda_1^k, \ldots, \lambda_M^k]\,\mathbf{C}, \tag{6.4}$$

is defined where $\lambda_1^k, \ldots, \lambda_M^k$ denote positive weights from $[0, 1]$, reflecting degrees of affiliation of components to the kth cluster. $\widehat{\mathbf{S}_k}$ should be, ideally, equal to

$$\mathbf{S}_k = \begin{bmatrix} s_k^1(N_1) & \ldots & \ldots & s_k^1(N_2) \\ s_k^1(N_1-1) & \ldots & \ldots & s_k^1(N_2-1) \\ \vdots & \vdots & \vdots & \vdots \\ s_k^1(N_1-L+1) & \ldots & \ldots & s_k^1(N_2-L+1) \\ s_k^2(N_1) & \ldots & \ldots & s_k^2(N_2) \\ s_k^2(N_1-1) & \ldots & \ldots & s_k^2(N_2-1) \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ s_k^d(N_1-L+1) & \ldots & \ldots & s_k^d(N_2-L+1) \end{bmatrix}, \tag{6.5}$$

which is the contribution of the kth source to $\mathbf{X}$, because $\mathbf{X} = \mathbf{S}_1 + \cdots + \mathbf{S}_d$ as follows from (5.2).

Taking the structure of $\mathbf{S}_k$ into account, the microphone responses are estimated from $\widehat{\mathbf{S}_k}$ as

$$\widehat{s_k^i}(n) = \frac{1}{L}\sum_{\ell=1}^{L} \psi_{k,(i-1)L+\ell}(n+\ell-1), \tag{6.6}$$

where $\psi_{k,p}(n)$ is equal to the (p,n)th element of $\widehat{\mathbf{S}}_k$. To clarify, note that $\psi_{k,p}(n)$ provides an estimate of $s_k^i(n-\ell+1)$ for $p=(i-1)L+\ell$.

An extensive descriptions of T-ABCD and of its generalized version are provided by article [j7], which is included as a part of this thesis. A conference paper on the generalized version was published in [c18].

6.3 Subband Separation

Subband processing is a standard approach of signal processing. In blind separation of audio signals, it provides a way between the FD and TD approaches. The mixed signals are decomposed into subbands via a filter bank. Each subband is processed by a TD separating method. The separated subbands must be re-ordered since the order of separated signals is random and different in each subband (the permutation problem). Finally, separated signals from subbands are synthesized, which gives the full-band separated signals.

The subband approach provides a compromise [56]. Although the permutation problem exists here, if a moderate number of subbands is chosen, it becomes less difficult than in FD methods. Next, the signals of each subband can be decimated prior to the separation since their spectra are narrow. The effective length of separating filters computed by the TD method and applied within subbands is increased (multiplied), consequently.

A subband variant of T-ABCD was proposed in article [c19], included in this thesis.

Reprints

- Z. Koldovský, P.Tichavský, "Time-domain blind separation of audio sources on the basis of a complete ICA decomposition of an observation space, " accepted for publication in *IEEE Trans. on Audio, Language, and Speech Processing*, April 2010.
- Z. Koldovský, P. Tichavský, and J. Málek, "Subband Blind Audio Source Separation Using a Time-Domain Algorithm and Tree-Structured QMF Filter Bank," *Proc. of LVA/ICA 2010*, 2010.

Time-Domain Blind Separation of Audio Sources on the Basis of a Complete ICA Decomposition of an Observation Space

Zbyněk Koldovský[1,2] and Petr Tichavský[2]

[1]Faculty of Mechatronics, Informatics, and Interdisciplinary Studies, Technical University of Liberec, Studentská 2, 461 17 Liberec, Czech Republic. E-mail: zbynek.koldovsky@tul.cz, fax:+420-485-353112, tel:+420-485-353534

[2]Institute of Information Theory and Automation, Pod vodárenskou věží 4, P.O. Box 18, 182 08 Praha 8, Czech Republic. E-mail: tichavsk@utia.cas.cz, fax:+420-2-868-90300, tel. +420-2-66052292

***Abstract*—Time-domain algorithms for blind separation of audio sources can be classified as being based either on a partial or complete decomposition of an observation space. The decomposition, especially the complete one, is mostly done under a constraint to reduce the computational burden. However, this constraint potentially restricts the performance. The authors propose a novel time-domain algorithm that is based on a complete unconstrained decomposition of the observation space. The observation space may be defined in a general way, which allows application of long separating filters, although its dimension is low. The decomposition is done by an appropriate independent component analysis (ICA) algorithm giving independent components that are grouped into clusters corresponding to the original sources. Components of the clusters are combined by a reconstruction procedure after estimating microphone responses of the original sources. The authors demonstrate by experiments that the method works effectively with short data, compared to other methods.**

I. Introduction

Blind separation of simultaneously active audio sources is a popular task of audio signal processing motivated by many emerging applications, such as hands-free and distant-talking speech communication, human/machine interactions, and so on. The goal is to retrieve d audio sources from their convolutive mixtures recorded by m microphones, which is described by

$$x_i(n) = \sum_{j=1}^{d} \sum_{\tau=0}^{M_{ij}-1} h_{ij}(\tau) s_j(n-\tau), \quad i = 1, \ldots, m, \quad (1)$$

where $x_1(n), \ldots, x_m(n)$ are the observed signals on microphones and $s_1(n), \ldots, s_d(n)$ are the original (audio) signals unknown in the "blind" scenario.

In fact, the mixing system is a multi-input multi-output (MIMO) linear filter with source-microphone impulse responses h_{ij}, each of length M_{ij}. The responses characterize propagation of sound in the recording room and are also unknown. It is assumed that the system is time-invariant, which usually means that positions of the sources and the microphones do not change within recording of N samples.

The separation through a linear processing consists in seeking a MIMO filter that inverts the mixing process (1). Any estimate of the jth original signal $s_j(n)$, $j = 1, \ldots, d$, thus has the form

$$\widehat{s}_j(n) = \sum_{i=1}^{m} \sum_{\tau=0}^{L-1} w_{ji}(\tau) x_i(n-\tau), \quad (2)$$

where L is the length of the separating filter. The blind separation that is based on the assumption of statistical independence of the original signals is addressed here. The separating filters will therefore be estimated via Independent Component Analysis (ICA) [1], [2].

Indeterminacies that are inherent to the ICA cause that each original signal is estimated up to an unknown filtering [3], [4]. Without any prior knowledge, that is not available in the blind scenario, an arbitrarily filtered source signal can also be considered a source signal. It is therefore meaningful to aim at estimating responses of sources at microphones, which only have properly defined colorations. Following from (1), the microphone response of the kth source at the ith microphone is

$$s_k^i(n) = \sum_{\tau=0}^{M_{ik}-1} h_{ik}(\tau) s_k(n-\tau). \quad (3)$$

Consequently, each source is estimated m times (all its responses are estimated). Once the responses $s_k^i(n)$, $i = 1, \ldots, m$, are estimated, it might be desirable to combine them in one-channel estimate of the kth signal denoted by $\widehat{s}_k(\cdot)$.

Basically, the blind audio source separation can be performed either in the frequency-domain or in the time-domain (TD). In the frequency-domain approach [5], [6], [7], the signals are transformed by the Discrete Fourier Transform (DFT), and the convolution operation in (1) changes to the

[0]Parts of this work were presented at HSCMA 2008 [21]. The work was supported by Grant Agency of the Czech Republic through the projects 102/07/P384 and 102/09/1278 and by Ministry of Education, Youth and Sports of the Czech Republic through the project 1M0572.

ordinary multiplication[1]. This translates the convolutive model into a set of complex-valued instantaneous mixtures, one for each frequency, that can be separated by complex-domain ICA methods. The frequency-domain approach allows effective computation of long separating filters, which is favorable in audio applications. By contrast, the computation of long filters requires long recordings to generate sufficient amount of data for each frequency [8].

Time-domain approaches transform the convolutive model into an instantaneous one by constructing data vectors or matrices of a special structure, by which the convolution is translated into the vector/matrix product. The data structures, constructed from the only available signals from microphones, define the *observation space*. Most often a matrix is defined so that its rows contain the time-lagged copies of signals from microphones, and the observation space is spanned by these rows. In general, TD methods aim at finding subspaces of the observation space that correspond to separated signals [9].

Decomposition of the observation space can be either complete or partial [10]. In the former case, the original signals are represented by d independent subspaces spanning the whole observation space. In the latter case, the signals are estimated as one-dimensional subspaces (components) of the observation space. A reconstruction procedure must follow the decomposition to retrieve the microphone responses of separated signals.

Performance of methods doing the partial decomposition depends very much on initialization of a convergence scheme [11], [12], [13]. It might also happen that the method finds two components of the same source and skips another source. In this respect, the complete decomposition is more reliable, however, at higher computational demand.

To alleviate these problems, the decomposition may be done with some constraint. The complete decomposition is usually constrained by an assumption that the inverse of the decomposing transform (matrix) has a special structure, for example block-Toeplitz or block-Sylvester; see articles of Kellermann et al. and Belouchrani et al., e.g. [9], [14], [15]. Févotte et al. proposed a two-stage separation procedure in [10] doing the complete decomposition by an algorithm for the independent subspace analysis (ISA) through joint block diagonalization (JBD) [9] utilizing the orthogonal constraint [17]. The algorithm of Douglas et al. [18] is an example of a constrained partial decomposition. It uses a para-unitary filter constraint and is compared in experiments in this article.

A potential drawback of the constrained decomposition is that it assumes all independent subspaces to have the same dimension. The constraint might also cause some restrictions due to the finite length of data or the limited length of separating filters. In this respect, the complete *unconstrained* decomposition provides an effective way to utilize the available data as effectively as possible, but it was considered to be computationally too extensive [10]. For instance, the JBD algorithm applied in [10] appeared to fail with $L > 6$, which, in other words, means that this algorithm cannot work on observation spaces of higher dimension. It is known that the stability and speed issues in high-dimensional spaces are the shortcomings of many ICA/ISA algorithms.

In this article, a novel method based on the complete unconstrained decomposition of the observation space is proposed. It utilizes modern ICA methods that allow fast, accurate and reliable separation of high-dimensional spaces. Especially, very fast ICA algorithms that are based on approximate joint diagonalization (AJD) by Tichavský and Yeredor [20] are used. Next, the method involves an effective reconstruction step, which yields effective results even when separating filters are much shorter than the mixing filter. Moreover, a general construction of the observation space is proposed, which allows the method to apply long (even infinite) separating filters while preserving its computational complexity (dimension of the observation space). In real-world experiments, the proposed method yields very good results in comparison with its competitors. It has several attractive features such as the ability to estimate the number of sources d, and it provides room for further development of its variants in future, such as a sub-band version or an on-line version.

The article is organized as follows. The following Section II provides a comprehensive description of a basic version of the proposed method, first introduced in [21], where classical time-lag construction of the observation space is used. The method is a five-step procedure, where each step can be solved in many alternative ways. A few basic variants are proposed. This also includes a novel *oracle* algorithm [8] that utilizes known responses of the sources and provides a reference solution that depends on the quality of the ICA decomposition only. In Section III, an extension of the method that comes from a generalized definition of the observation space is proposed. A special case of the definition leads to the application of infinite impulse response (IIR) Laguerre separating filters. In Section IV, results of various real-world experiments that demonstrate excellent performance of the proposed method in comparison with other existing methods are presented.

II. Basic Version of the Proposed Method

In the following subsection, a brief description of main steps of the basic variant of the proposed method is given, and in the other subsections each step is further commented and illustrated by an example.

A. Outline

Assume that N samples of simultaneously recorded signals from microphones $x_1(n), \ldots, x_m(n)$, $n = 1, \ldots, N$, are available. The method proceeds in five consecutive steps.

1) *Form a* $M \times (N_2 - N_1 + 1)$ *data matrix* $\mathbf{X}$, *whose rows contain time-lagged copies of the signals from microphones.* Each signal is delayed L times, thus, L rows correspond to each signal, and $M = mL$. The

[1]More precisely, the circular convolution changes to the ordinary multiplication.

matrix $\mathbf{X}$ is given by

$$\mathbf{X} = \begin{bmatrix} x_1(N_1) & \dots & \dots & x_1(N_2) \\ x_1(N_1-1) & \dots & \dots & x_1(N_2-1) \\ \vdots & \vdots & \vdots & \vdots \\ x_1(N_1-L+1) & \dots & \dots & x_1(N_2-L+1) \\ x_2(N_1) & \dots & \dots & x_2(N_2) \\ x_2(N_1-1) & \dots & \dots & x_2(N_2-1) \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ x_m(N_1-L+1) & \dots & \dots & x_m(N_2-L+1) \end{bmatrix}, \tag{4}$$

where N_1 and N_2, $1 \le N_1 < N_2 \le N$, determine a segment of recordings that is used for computations. The subspace of $\mathbb{R}^{N_2-N_1+1}$ spanned by rows of $\mathbf{X}$ will be called the *observation space*.

2) *Apply an ICA method to the mixture given by* $\mathbf{X}$ *to obtain* all *independent components of* $\mathbf{X}$. As there may be up to M (independent) components, the output is a $M \times M$ *de-mixing* (decomposing) matrix $\mathbf{W}$, and the components are given by $\mathbf{C} = \mathbf{WX}$. The rows of $\mathbf{C}$ will be denoted $\mathbf{c}_1^T, \dots, \mathbf{c}_M^T$ and the components (the signals) defined by them will be denoted by $c_1(n), \dots, c_M(n)$.
3) *Group the components* $c_1(n), \dots, c_M(n)$ *into* d_{est} *clusters, so that each cluster contains components that correspond to the same original source.* The number d_{est} is either estimated or equal to an apriori known (if available) number of sources d. The grouping is done subject to a similarity measure between the components.
4) *For each cluster and each component, a weight that characterizes a measure of confidence of the component to belong to the cluster is computed. Then for each cluster, a reconstructed version of the matrix* $\mathbf{X}$ *is computed using weighted components subject to the cluster, and rows of the reconstructed matrix are used for estimation of microphone responses of a source corresponding to the cluster.* Mathematically, the reconstructed matrix, for the kth cluster, $k = 1, \dots, d_{est}$, is

$$\begin{aligned} \widehat{\mathbf{S}_k} &= \mathbf{W}^{-1}\mathrm{diag}[\lambda_1^k, \dots, \lambda_M^k]\,\mathbf{C} \\ &= \mathbf{W}^{-1}\mathrm{diag}[\lambda_1^k, \dots, \lambda_M^k]\,\mathbf{W}\,\mathbf{X}, \end{aligned} \tag{5}$$

where $\lambda_1^k, \dots, \lambda_M^k$ denote the weights, each one from $[0, 1]$, reflecting degrees of affiliation of components to the kth cluster. Their particular selection will be described later in this section. Finally, microphone responses (3) of an original source corresponding to the kth cluster are estimated as

$$\widehat{s}_k^i(n) = \frac{1}{L}\sum_{\ell=1}^{L} \psi_{k,(i-1)L+\ell}(n+\ell-1), \quad i = 1, \dots, m, \tag{6}$$

where $\psi_{k,p}(n)$ is the (p, n)th element of $\widehat{\mathbf{S}}_k$. Obviously $\psi_{k,p}(n)$, $p = (i-1)L+\ell$, provides an estimate of $s_k^i(n-\ell+1)$.
5) *Apply a beamformer to the estimated responses of each source to get the one-channel estimate of the source.*

In the following subsections, the steps of this method are discussed in more details. To make the presentation clearer, an accompanying example is given with three original sources that were artificially mixed into three signals. The mixing system consists of filters of the length $M_{ij} = 4$, $i, j = 1, \dots, 3$, whose coefficients were randomly generated according to Gaussian law with zero mean and unit variance. The original and the mixed signals are, respectively, shown in Figs. 1 and 2.

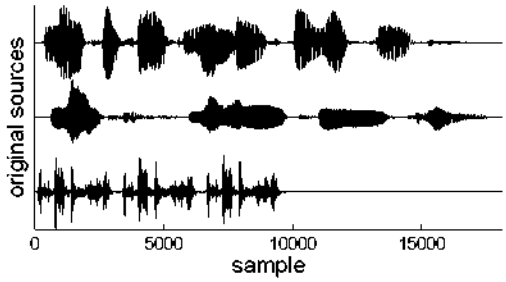


Fig. 1. Original sources considered in the demonstration example. The signals are, respectively, a man's speech, a woman's speech, and a typewriter sound, recorded at the sampling frequency 8kHz.

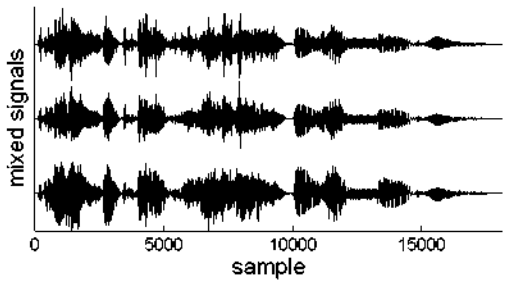


Fig. 2. Three artificial convolutive mixtures of the sources from Fig. 1 simulating signals obtained by three microphones.

B. Step 1: Construction of $\mathbf{X}$

As mentioned in the introduction, constructing $\mathbf{X}$ according to (4) allows to convey the separating convolution operation via multiplying $\mathbf{X}$ by a de-mixing matrix. $\mathbf{X}$ is usually interpreted as an instantaneous mixture, $\mathbf{X} = \mathbf{AS}$, where $\mathbf{S}$ is a matrix constructed of delayed original signals analogously to $\mathbf{X}$, and $\mathbf{A}$ is a mixing matrix that has the block-Sylvester structure. However, such mixture is equivalent with (1) in full if only $\mathbf{A}$ has more columns than rows, $m > d$ and L is sufficiently large; see [9], [10].

In this article, none of the above conditions is assumed. The mixing or, equivalently, the de-mixing matrix is considered to be square without any special structure. The structure of $\mathbf{S}$ is not specified either. It is only assumed that its rows consist of the filtered versions of original signals and form independent subspaces. Consequently, $\mathbf{S}$ can be estimated, up to indeterminacies, as independent subspaces or components of $\mathbf{X}$ via ISA or ICA. This approach proves to be more flexible, among others, because $\mathbf{X}$ may be defined in different ways than (4) as proposed in Section III.

In the accompanying example, consider $L = 4$, $N_1 = L$ and $N_2 = 8000 + L - 1$. This means that the length of

separating filters is 4, and the matrix $\mathbf{X}$ is 12×8000. Note that all computations made with $\mathbf{X}$ use data contained in first 8000 samples (the first second) of recordings only. Once the separating MIMO filter is found, it can be applied to the entire data set.

C. Step 2: ICA Decomposition

At the heart of the proposed separation procedure is a suitable ICA algorithm to be applied to $\mathbf{X}$. Because no constraint is applied to the de-mixing matrix, many of the known ICA and ISA algorithms can be considered including those based on Non-Gaussianity, nonstationarity or spectral diversity (distinct coloration) of signals; a survey of ICA algorithms is provided, for example, by ICALAB [22].

The problem of the selection of ICA/ISA algorithm for this purpose exceeds the scope of this paper. The study in [23] showed that ISA algorithms do not have any obvious advantage over ICA algorithms that are followed by clustering. Potentially, ICA methods are computationally inefficient since they not only separate independent subspaces, but also signals within the subspaces. However, the ICA methods considered here are computationally still much faster then up-to-date ISA methods.

Owing to the need to separate mixtures whose dimension is frequently 40 or more, two algorithms are considered: the Non-Gaussianity based EFICA algorithm from [24] and the nonstationarity based algorithm from [20] called BGSEP.

EFICA is an improved version of the well-known FastICA algorithm [25]. BGSEP consists in a special approximate joint diagonalization of a set of covariance matrices of signals in data matrix divided in blocks. Both methods achieve asymptotical optimality within respective models of signals and perform very well in [23].

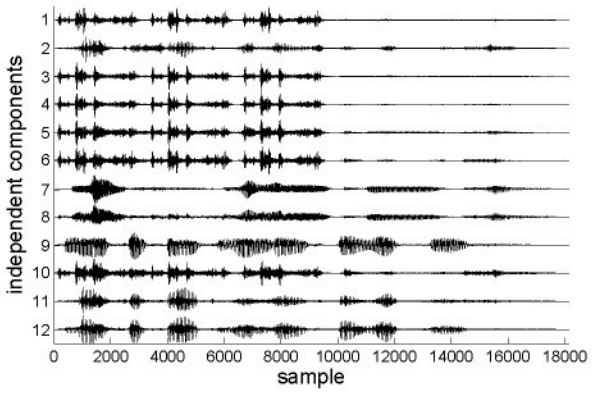


Fig. 3. Independent components obtained by the BGSEP algorithm in the demonstration example. As can be seen by comparing signals from Fig. 1, some components clearly correspond to separated signals.

D. Step 3: Clustering of Components

An independent component obtained by the ICA algorithm equals, in an ideal case, to a filtered copy of an original source. As the number of components is higher than the number of sources, i.e. $M > d$, there should be d clusters of components where each cluster contains components of one source. The utilization of the ICA algorithm in the second step should be therefore followed by the clustering of components.

As already discussed above, the alternative way is to apply an ISA method instead of ICA, which does not need the clustering step [19], [26], [27]. However, the "ICA+clustering" approach used here has the following advantages.

- ICA methods work reliably without knowing or estimating the number of components of clusters.
- The approach is flexible because various criteria of similarity of components and clustering methods can be used.

1) Similarity of components: If the ith and the jth component belong to the same source and contain no interference, it holds that there exists a filter f such that

$$c_i(n) = \sum_{\tau=-\infty}^{+\infty} f(\tau) c_j(n-\tau) = \{f \star c_j\}(n). \qquad (7)$$

In practice, (7) holds approximately only, and f can be searched by minimizing the mean square distance between the two sides of (7). Therefore, the value of

$$\min_f \hat{\mathrm{E}}[c_i(n) - \{f \star c_j\}(n)]^2, \qquad (8)$$

where $\hat{\mathrm{E}}$ denotes the sample mean operator, reveals whether the two components belong to the same source. In practice, the minimization in (8) proceeds over filters of length $2L$.

Therefore, the similarity of the ith and the jth component, $i \neq j$, is defined as the ijth element of matrix $\mathbf{D}$, where

$$\mathbf{D}_{ij} = \hat{\mathrm{E}}[\mathbf{P}_i \mathbf{c}_j]^2 + \hat{\mathrm{E}}[\mathbf{P}_j \mathbf{c}_i]^2, \qquad (9)$$

where $\mathbf{P}_i$ denotes a projector on a subspace spanned by delayed copies of the ith component, that is, by signals $c_i(n-L+1), \ldots, c_i(n+L-1)$. Diagonal elements of $\mathbf{D}$ have no significance here and are set to zero. The computation of (9) can be done efficiently using the FFT and Levinson-Durbin algorithm; see [21]. An example of the similarity matrix $\mathbf{D}$ is shown in Fig. 4.

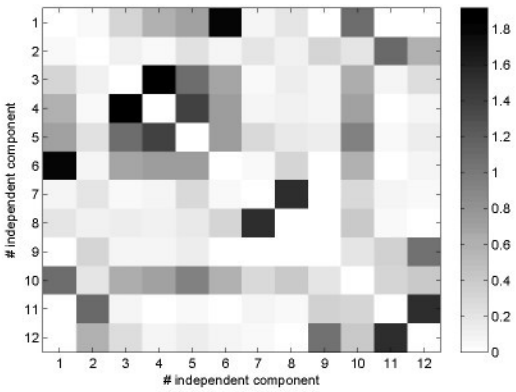


Fig. 4. Similarity matrix $\mathbf{D}$ between components from Fig. 3 computed according to the definition (9).

2) Clustering: The task now is to cluster the M components subject to the similarity matrix $\mathbf{D}$. This general task can be solved by various methods. In this study, attention is restricted to the agglomerative hierarchical clustering algorithm that appears to perform well in this application.

The algorithm consists of M levels, each giving a partitioning of components. In the beginning (the first level), each component forms a cluster, called singleton, thus, there are M clusters. At each subsequent level, the method merges two clusters whose similarity is maximal. The number of clusters is thus always equal to the level. In the last level, all components form one cluster.

Finally, the most satisfactory level (partitioning) should be chosen. If the number of sources d is known in advance, the level giving the desired number of clusters is selected. Otherwise, it is possible to select the level according to a criterion such as

$$\max_p \left\{ \frac{1}{M-p+1} \sum_{k=1}^{M-p+1} \frac{M-|K_k^p|}{|K_k^p|} \frac{\sum_{i\in K_k^p, j\in K_k^p} \mathbf{D}_{ij}}{\sum_{i\in K_k^p, j\notin K_k^p} \mathbf{D}_{ij}} \right\}, \tag{10}$$

where p is the level index within $M-m+1,\ldots,M-1$ (i.e., the maximum number of estimated sources corresponds to the number of microphones m), K_k^p is a set of indices of components in the kth cluster of the pth partitioning level, and $|K_k^p|$ is the number of those indices. The argument of sum in (10) evaluates the ratio between the average intra-similarity of components of the kth cluster to the average inter-similarity of components from the other clusters. The criterion thus reflects the quality of the pth partitioning as it averages the argument over all of its clusters.

Maximization of (10) can be interpreted as a method of estimating the number of sources. However, since the results are not always satisfactory in practice, there is room for further improvement. In this paper we assume, for simplicity, that the number of active sources is known a priori.

What is left is to define the similarity between clusters, called the *linkage strategy*. A modified average linking strategy is to be used, which is defined as follows. Let Q and R contain indices of components of two different clusters. The similarity of the clusters is given by

$$d(Q,R) = \frac{1}{\min(|Q|,|R|)} \frac{1}{|Q|} \frac{1}{|R|} \sum_{q\in Q} \sum_{r\in R} \mathbf{D}_{qr}, \tag{11}$$

where $|Q|$ is the number of indices in Q. The modification of the average linkage strategy consists in the division by $\min(|Q|,|R|)$. It penalizes mutual similarity of "large" clusters and highlights the similarity of "small" clusters with "large" ones, which is preferable to this application. Pseudocode 1 summarizes the clustering algorithm.

The clustering algorithm was applied to the components from Fig. 3. Three clusters shown by Figs. 5(a)-5(c) were found; reordered similarity matrix $\mathbf{D}$ according to the clustering is shown by Fig. 6. This example demonstrates clearly that each source may consist of different number of components. In other words, independent subspaces corresponding to the original sources may have different dimensions.

Pseudocode 1 Hierarchical clustering of components

$K_i^1 = \{i\},\ i = 1,\ldots,M$
$\mathcal{K}^1 = \{K_1^1,\ldots,K_M^1\}$
for $p = 1$ to $M-1$ **do**
  $k,\ell = \arg\min_{k,\ell=1,\ldots,M-p+1} d(K_k^p, K_\ell^p)$
  $\mathcal{K}^{p+1} = \{K_i^p, i\neq k,\ell\} \cup \{K_k^p \cup K_\ell^p\}$
end for
if d is known **then**
  $p = M-d+1$
else
  Select p from $M-m+1$ to $M-1$ according to (10)
end if
return $\mathcal{K}^p = \{K_1^p,\ldots,K_{M-p+1}^p\}$

It can also be seen that some components often exhibit certain closeness to more than one cluster. This is because of the residual interference between components caused by various practical limitations such as the finite length of separating filters. The method takes this important phenomenon into account in the reconstruction step discussed in the following subsection.

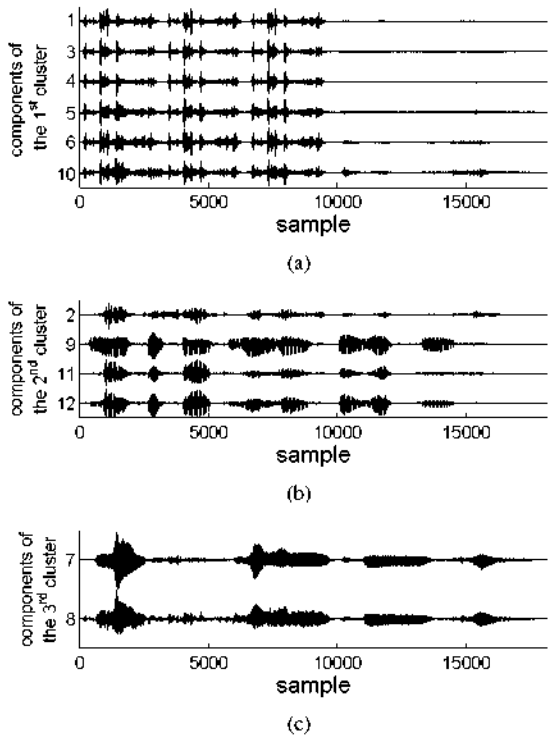


Fig. 5. Components assigned to the three founded clusters.

E. Step 4: Reconstruction

The goal of the reconstruction step is to obtain the responses of sources on microphones defined by (3). The response is a signal observed by the microphone if the source is sounding solo. Since all sources sound simultaneously, it holds that

$$x_i(n) = s_1^i(n) + \cdots + s_d^i(n), \qquad i = 1,\ldots,m. \tag{12}$$

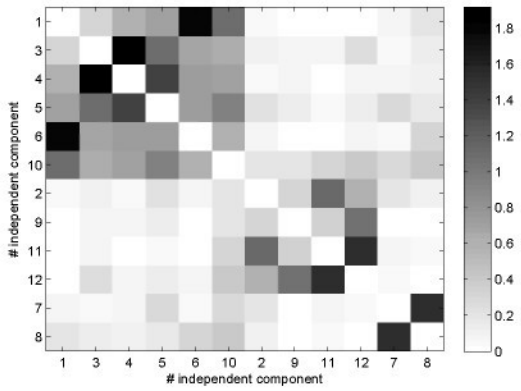


Fig. 6. Reordered similarity matrix $\mathbf{D}$ from Fig. 4 after the hierarchical agglomerative clustering. Three detected clusters can be markedly seen from the picture.

Hence, $\mathbf{X}$ can be written as a sum of matrices $\mathbf{S}_1, \ldots, \mathbf{S}_d$, where $\mathbf{S}_k$ is constructed in the same way as $\mathbf{X}$ but using responses of the kth source only,

$$\mathbf{X} = \mathbf{S}_1 + \cdots + \mathbf{S}_d. \tag{13}$$

First, the binary weighting that reflects the results of the clustering of components is introduced by setting the weights introduced in (5) to

$$\lambda_\ell^k = \begin{cases} 1 & \ell \in K_k \\ 0 & \text{otherwise} \end{cases}, \tag{14}$$

where K_k contains indices of components assigned to the kth cluster.

If there is no interference between components, and the clustering proceeds without errors, then $\widehat{\mathbf{S}_k}$ obtained by (5) satisfies $\widehat{\mathbf{S}_k} = \mathbf{S}_k$, $k = 1, \ldots, d_{est}$, and rows of $\widehat{\mathbf{S}_k}$ contain delayed microphone responses of the kth source $s_k^1(n), \ldots, s_k^m(n)$. Equation (6) means that the pth response is estimated as an average of $(p-1)L+1, \ldots, pL$ rows of $\widehat{\mathbf{S}_k}$ with restored delays.

It is worth noting here that the length of filters found by ICA producing independent components is L. The reconstruction formula (6) can be interpreted as a filtering by another FIR filter of the length L. Therefore the final separating filter has the length up to $2L-1$.

F. Computation of Weights

A natural extension of the "hard" weighting given by (14) is to consider λ_ℓ^k as positive numbers from $[0, 1]$ selected according to an appropriate rule. The rule introduced in [21] is used, which is given by

$$\lambda_\ell^k \leftarrow \left(\frac{\sum_{j \in K_k, j \neq \ell} \mathbf{D}_{\ell j}}{\sum_{j \notin K_k, j \neq \ell} \mathbf{D}_{\ell j}} \right)^\alpha$$
$$\lambda_\ell^k \leftarrow \lambda_\ell^k / \left(\max_{\ell=1,\ldots,M} \lambda_\ell^k \right), \tag{15}$$

where α is an adjustable positive parameter. The denominator in (15) reflects the similarity of the ℓth component to components from different clusters than the kth one. If the component clearly belongs to the kth cluster, the denominator is close to zero, and the value of (15) becomes large.

If $\alpha \to +\infty$, the reconstruction proceeds practically from a single component with the maximum value of the fraction in (15). On the other hand, with α close to zero the weighting becomes uniform, which means no separation.

An example in Section IV. indicates that a good choice of α is $\alpha = 1$. Figure 7 shows resulting weights obtained in the demonstration example for this choice.

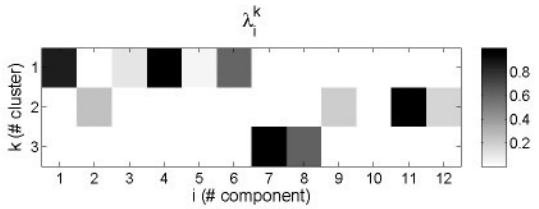


Fig. 7. Weights of components from Fig. 3 computed according to the rule (15) with $\alpha = 1$.

G. Oracle Weighting

It is interesting to know what would be the best possible weights for separation in theory, given the ICA decomposition of the observation space. In other words, what are the best possible weights independent of the similarity given by $\mathbf{D}$ and the clustering algorithm. Such weights can be derived using known responses of sources. The authors call it an "oracle weighting", and the corresponding algorithm an "oracle algorithm", following the work of Vincent et al. [8].

The oracle weighting can be derived as the one that minimizes $\|\mathbf{S}_k - \widehat{\mathbf{S}_k}\|_F^2$, $k = 1, \ldots, d_{est}$, given the true responses of sources on the microphones forming the matrix $\mathbf{S}_k$. Here $\|\cdot\|_F$ denotes the Frobenius norm, and $d_{est} = d$. Using (5), the oracle weights are defined by

$$\mathbf{l}_k = \arg\min_{\mathbf{l}} \|\mathbf{S}_k - \mathbf{W}^{-1}\mathrm{diag}(\mathbf{l})\mathbf{W}\mathbf{X}\|_F^2, \tag{16}$$

where $\mathbf{l}_k = [\lambda_1^k, \ldots, \lambda_M^k]^T$. After some computations it can be shown that

$$\mathbf{l}_k = \left((\mathbf{W}\mathbf{X}\mathbf{X}^T\mathbf{W}^T) \odot (\mathbf{W}\mathbf{W}^T)^{-1} \right)^{-1} \cdot \mathrm{diag}\left[\mathbf{W}^{-T}\mathbf{S}_k\mathbf{X}^T\mathbf{W}^T \right], \tag{17}$$

and $\odot$ denotes the Hadamard (element-wise) product. The rest of the oracle algorithm (reconstruction and beamforming) proceeds normally.

H. Step 5: Beamforming

A beamformer can be applied to the multi-channel estimate of each source (microphone responses) to yield a mono-channel estimate of the source. This problem is not addressed here, because it exceeds the scope of this article. The beamforming requires an additional definition of a principle that is not given in the blind scenario considered here. The reader is referred to [16].

Results obtained in the demonstration example after delay-and-sum beamforming of estimated microphone responses are shown in Fig. 8. The order of estimated signals with respect to the original ones is arbitrary.

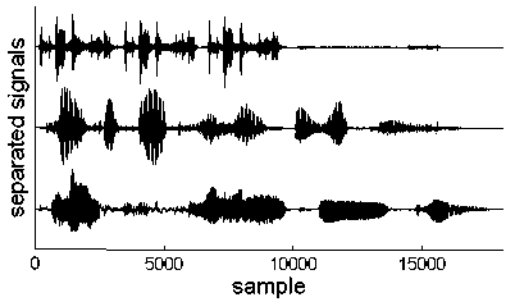


Fig. 8. Estimates of original sources obtained by the delay-and-sum beamformer applied to estimated microphone responses.

III. Generalized Observation Space

In this section, the previous definition of the observation space is altered (generalized) in order to enable the application of long (IIR) separating filters while keeping the same dimension of the observation space. This makes the method computationally affordable even when long separating filters are considered.

It is worth pointing out that maintaining a reasonable dimension of the observation space is also desirable from the probabilistic theory point of view. The ICA is a stochastic method whose accuracy measured in terms of the interference-to-signal ratio is, in theory, proportional to the number of components M; see for example [24], [28], [29].

The original construction of $\mathbf{X}$ given by (4) means that FIR filters are applied to the mixed signals, because the outputs of FIR filters can be seen as linear combinations of time-shifted versions of the input signals (rows of (4)). The proposed generalization consists in constructing $\mathbf{X}$ so that separating filters have a well-known generalized feed-forward (FF) structure [30], [31], which also embodies FIR filters as a special case.

The output of a generalized FF filter applied to an input signal $x(n)$ can be written as

$$\begin{aligned} y(n) &= \sum_{\ell=1}^{L} w_\ell \sum_{\tau=-\infty}^{+\infty} f_\ell(\tau) x(n-\tau) \\ &= \sum_{k=1}^{L} w_\ell \{f_\ell \star x\}(n), \end{aligned} \tag{18}$$

where $\star$ denotes the convolution, w_ℓ are the filter weights, and the filters f_ℓ are called *eigenmodes* of the filter. The definition of MIMO filters with the generalized FF structure is analogous.

A. Generalized Observation Matrix $\mathbf{X}$

For a given set of invertible eigenmodes f_ℓ, the ith block of the observation matrix $\mathbf{X}$ can be defined as

$$\mathbf{X}_i = \begin{bmatrix} \{f_1 \star x_i\}(N_1) & \dots & \dots & \{f_1 \star x_i\}(N_2) \\ \{f_2 \star x_i\}(N_1) & \dots & \dots & \{f_2 \star x_i\}(N_2) \\ \vdots & \vdots & \vdots & \vdots \\ \{f_\ell \star x_i\}(N_1) & \dots & \dots & \{f_\ell \star x_i\}(N_2) \end{bmatrix}, \tag{19}$$

The whole $\mathbf{X}$ is then given by

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \vdots \\ \mathbf{X}_m \end{bmatrix}. \tag{20}$$

If f_ℓ is the all-pass filter that realizes backward time-shift by ℓ samples, the construction in (20) coincides with the one in (4).

Example of perfect separation: Consider the general 2×2 scenario

$$x_1(n) = \{h_{11} \star s_1\}(n) + \{h_{12} \star s_2\}(n) \tag{21}$$

$$x_2(n) = \{h_{21} \star s_1\}(n) + \{h_{22} \star s_2\}(n). \tag{22}$$

Almost perfect separation can be achieved when taking $L = 2$ and applying special eigenmodes for each matrix $\mathbf{X}_1$ and $\mathbf{X}_2$, namely, $f_{11} = g \star h_{22}$, $f_{12} = -g \star h_{21}$, $f_{21} = -g \star h_{12}$, and $f_{22} = g \star h_{11}$, where $g = (h_{11} \star h_{22} - h_{21} \star h_{12})^{-1}$ assuming that the inversion exists. A trivial verification shows that combinations of signals $\{f_{11} \star x_1\}(n) + \{f_{21} \star x_2\}(n)$ and $\{f_{12} \star x_1\}(n) + \{f_{22} \star x_2\}(n)$ are, respectively, equal to the original independent sources s_1 and s_2. In other words, s_1 and s_2 can be found as independent components in the observation space. ■

The example demonstrates the great potential of the general construction of $\mathbf{X}$ in theory. For instance, it is indicative of the possibility to tailor the eigenmodes to room acoustics.

After the ICA decomposition of $\mathbf{X}$ the method proceeds normally up to the fourth reconstruction step. Let $\psi_{k,p}(n)$ be the (p,n)th element of $\widehat{\mathbf{S}}_k$. Then, $\psi_{k,p}(n)$, $p = (i-1)L + \ell$, provides an estimate of $\{f_\ell \star s_k^i\}(n)$. Let f_ℓ^{-1} be the inverse of f_ℓ so that $f_\ell \star f_\ell^{-1} = \delta$. The authors estimate the response of the kth separated source at the ith microphone as

$$\widehat{s}_k^i(n) = \frac{1}{L} \sum_{\ell=1}^{L} \{f_\ell^{-1} \star \psi_{k,(i-1)L+\ell}\}(n). \tag{23}$$

Obviously, (23) is a generalization of (6).

B. Laguerre Filters

A good example of FF filters for this study are Laguerre filters parametrized by μ from $(0, 2)$, which were considered in [31]. They are defined recursively, through their transfer functions

$$F_1(z) = 1, \tag{24}$$

$$F_2(z) = \frac{\mu z^{-1}}{1 - (1-\mu) z^{-1}}, \tag{25}$$

$$F_n(z) = F_{n-1}(z) G(z), \quad n = 3, \dots, L, \tag{26}$$

where

$$G(z) = \frac{(\mu - 1) + z^{-1}}{1 - (1-\mu) z^{-1}}. \tag{27}$$

Note that f_2 is either a low-pass filter (for $0 < \mu < 1$) or a high-pass filter (for $1 < \mu < 2$), and g is an all-pass filter.

The construction discussed here is a generalization of (4), because for $\mu = 1$, $F_2(z) = G(z) = z^{-1}$, that is $f_2(n) =$

$g(n) = \delta(n-1)$. This is the only case where separating filters are FIR of the length L. For $\mu \neq 1$, the filters are IIR.

The so-called *memory depth* denoted by L_* is defined as the minimum length needed to capture 90% of the total energy contained in the impulse response. For the Laguerre filters it approximately holds that [31]

$$L_* = (1 + 0.4|\mu - 1| \log_{10} L)L/\mu. \quad (28)$$

From here on, we will name the proposed method T-ABCD (Time-domain Audio sources Blind separation based on the Complete Decomposition of the observation space) keeping in mind that its given variant must be specified at the place where the acronym is used.

IV. Experiments

A. Experiments with short recordings of two sources

The experiments described in this section were designed to compare T-ABCD in setups for which the method was specialized. Those are mainly situations where only short recordings are available, and short separating filters ($L = 3, \ldots, 40$) are used (more precisely, the dimension of the observation space is in the range of tens).

Data used for testing of T-ABCD consist of nine recordings of two simultaneously talking persons articulating short commands. The length of each recording is 2 s, which gives 16000 samples at 8kHz sampling. Different genders are considered so that there are three recordings of male/male, three of female/male, and three of female/female speakers.

The recordings were obtained by two closely spaced microphones when playing the speakers' commands over two loudspeakers. Microphone responses of each source were obtained by recording the source when the other sources were silent[2].

Three different positions of loudspeakers were considered that differ in distance and angle between sources; see Fig. 9 and Table I. Each scenario was situated in an ordinary living room with the reverberation time of about 300 ms.

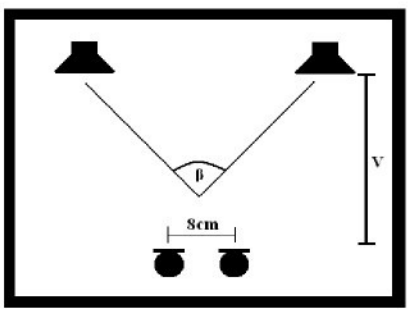


Fig. 9. Illustration of positions of sources (loudspeakers) and microphones.

Two variants of T-ABCD were tested using the BGSEP and the EFICA algorithm in the second step, respectively, marked as *T-ABCDb* and *T-ABCDe*. For theoretical reasons performances of the oracle algorithms (Section II.G) based on these ICA algorithms were also studied. They are denoted as *Oracle-b* and *Oracle-e*, respectively. Finally, an ultimate performance bound determined by the MMSE estimator [31] is shown, which is computed for the separating filter length $2L - 1$.

[2]Equipment used for recordings consists of external sound device EDIROL FA-101 and condenser omnidirectional microphones Audio-technica AT831b.

TABLE I
Technical Details of Scenarios

scenario	β	V [cm]	Average SNR [dB]
1	70°	50	40.6
2	60°	100	38.6
3	75°	200	35.9

Two other algorithms were used for comparisons. The first one is the STFICA algorithm from [18] using two stages of preprocessing (prewhitening) of the length 300. The observation space separated by STFICA is set to have the same dimension as the proposed method, that is, $2L$ for two microphones. The second algorithm is that of Parra from [6] with two lengths of FFT, respectively, 512 (*Parra-1*) and 128 (*Parra-2*); the other parameters had the default values.

Results of these experiments are evaluated by two standard measures [32]: Signal-to-Interference ratio (SIR) and Signal-to-Distortion ratio (SDR). The SIR determines the ratio of energies of the desired signal and the interference in the separated signal. SIR is highly influenced by a filtering of the measured signal, which might be misleading, especially, in audio separation. It is also influenced by the input SIR, which is the SIR measured before the separation. In our experiments, the input SIR was always about 0 dB, which means that both sources were approximately equally loud. The SDR provides a supplementary criterion of SIR that reflects the difference between the desired and the estimated signal in the mean-square sense. SDR is, by contrast, highly sensitive to the filtering, which may yield a rigid evaluation of methods applying long separating filter. It is therefore advisable to consider both criteria.

Hereinafter, all results are evaluated in terms of averaged SIR improvement and SDR over all separated sources and over all their estimated microphone responses. The results are also averaged over the nine recorded combinations of signals to reduce the effect of statistical properties of the recorded signals.

1) Performance versus L: Here, the separation is done for different lengths of separating filter L, and the other parameters are fixed. Namely, only 8000 samples of data (the first second) are used for the computation. T-ABCD utilizes the basic construction of the observation space corresponding to $\mu = 1$.

Figure 10 shows results of separation obtained by processing signals from scenario 1. SDR of T-ABCD improves with growing L, similar to the SDR of the MMSE estimator. SIR does not improve with growing L, but is good for all L. This is explained by the fact that for small L, few components are used to reconstruct sources, and SIR remains good, but SDR is poorer, because the reconstructed sources have different coloration than the original responses. In this respect, the behavior of oracle algorithms is different as they primarily optimize SDR. The gap between the SIR/SDR of the oracle

algorithms and T-ABCD indicates that there might still be a room for improvements of the performance through clustering and weighting.

Separated signals obtained by the other algorithms, STFICA, Parra-1 and Parra-2 are perceptually not bad, but not as good as those of the proposed method. It does not improve with growing L neither in terms of SIR nor in terms of SDR. STFICA failed to converge for $L \geq 20$.

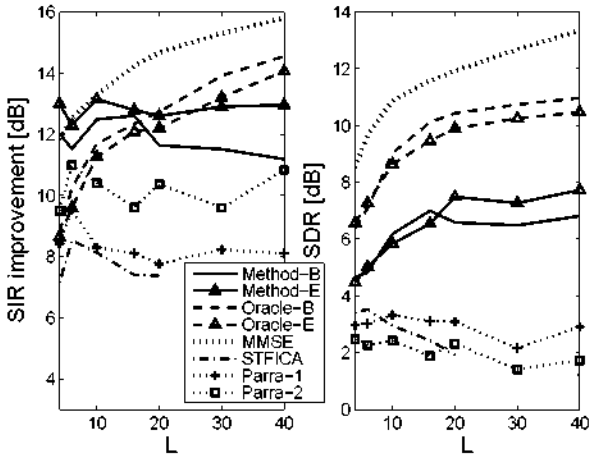


Fig. 10. SIR and SDR as functions of the length of separating filter L. The results were obtained by using data from scenario 1.

2) Performance versus Length of Data: A similar experiment to the previous one was repeated for a fixed $L = 20$ and varying the length of data used for computations of separating filter in scenario 2. This scenario is more difficult for the separation because of the higher distance of sources and lower ratio between the energy of the direct-path source signals and the energy of their reflections. The results are shown in Fig. 11.

It is noted that for a fixed filter length, there is a certain length of the data beyond which performance of the algorithms does not improve at all. In this experiment, the length was about 0.8-1s. T-ABCD performs better than the other algorithms and demonstrates its superior capability to work with short data.

3) Performance versus α: The parameter α was introduced in (15), and provides some trade-off between SIR and SDR. This is demonstrated by separating signals from scenario 2 by T-ABCD with $L = 20$, 8000 samples of data for computations, and various α. Results are shown in Fig. 12.

It is noted that SIR is an increasing function of α, whereas SDR achieves its maximum at a certain value in the vicinity of $\alpha = 1$. This points to the need of using SIR and SDR simultaneously to evaluate the separation fairly.

4) Performance of T-ABCD using Laguerre filters versus μ: The signals recorded in scenario 3 were separated with $L = 20$ using the first second of data. The parameter μ was gradually decreased from 1.9 to 0.1, which corresponds to changing the separating filter memory L_* defined by (28) from 15 to 293 samples; see Fig. 13.

The results indicate a minor (about 0.7 dB) improvement of performance of T-ABCD at the optimum value $\mu = 0.2$ compared to $\mu = 1$. A higher potential improvement in performance is indicated by increased SIR improvement of the MMSE bound and of the oracle algorithms (about 2dB). Again, it indicates a room for improvement through a different clustering and weighting.

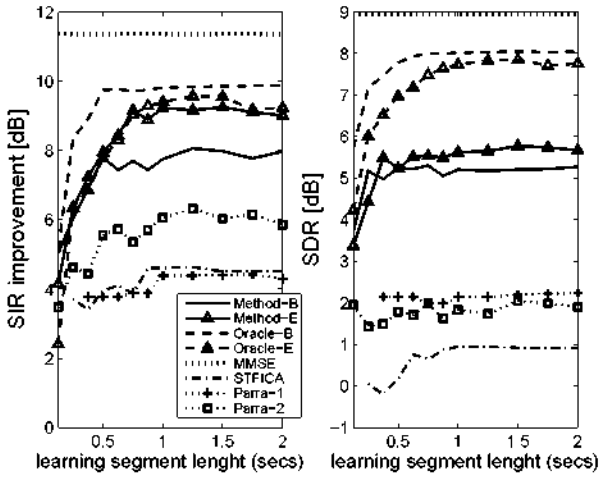


Fig. 11. SIR and SDR as functions of the length of data used for the computation of separating filter. Results corresponds to data from scenario 2.

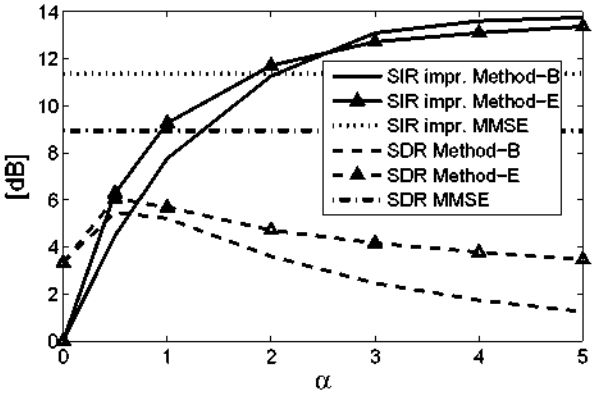


Fig. 12. Performance of T-ABCD as a function of α. In this example, STFICA, Parra-1 and Parra-2 performed, respectively, with 4.6dB, 4.4dB and, 6dB of SIR improvement and 1dB, 1.8dB, and 2.2dB of SDR.

B. Experiments with Hiroshi Sawada's Recordings

In this subsection, the above methods were tested by separating data available on the Internet[3]. These data were recorded in a room with the reverberation time 130ms. A linear microphone array with the distance of 4cm between microphones was used to record 2-4 simultaneous speeches coming from different directions from the distance of 1.2 m at the sampling rate 8kHz. The length of the recordings is 7 s.

The data were processed by T-ABCD with Laguerre filters with $\mu = 0.2$ and $L = 30$. Therefore, the dimension of the observation space was equal to $30m$, where m is the number

[3]http://www.kecl.ntt.co.jp/icl/signal/sawada/demo/bss2to4/index.html

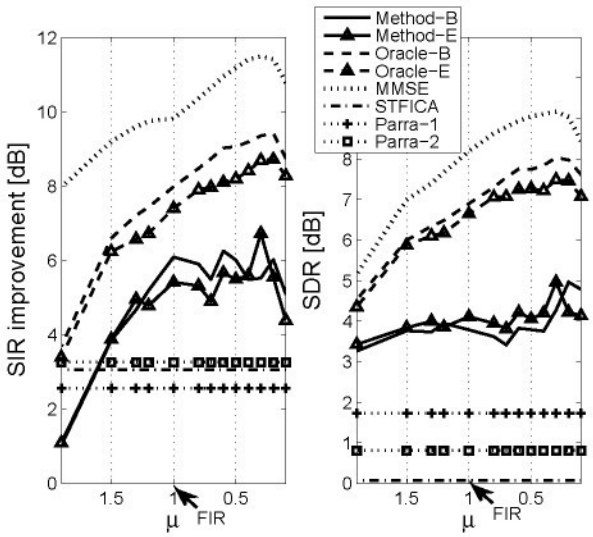


Fig. 13. SIR and SDR as functions of the parameter μ of Laguerre filters. Results corresponds to data from scenario 3. Note that for $\mu = 1$ the Laguerre filters are FIR, and the generalized T-ABCD coincides with the original one described in Section II.

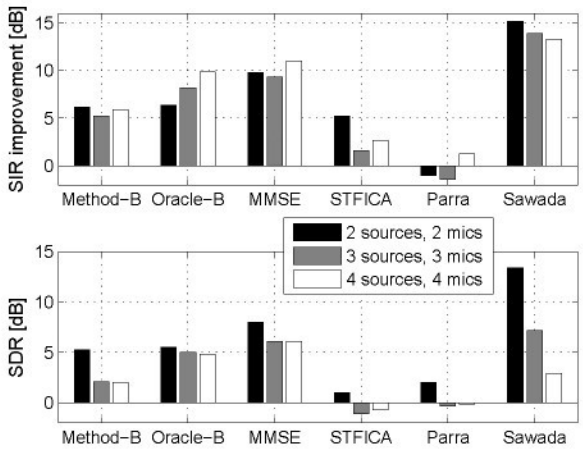


Fig. 14. Performance of separation methods applied to Sawada's data with 2-4 sources and microphones.

of microphones (2-4). The algorithm used 8000 samples (1s) to perform the separation, beginning at 4.6s of the recordings. For the Parra's algorithm, the whole (7 s long) recordings were used, the length of the FFT was 1024 and the time-domain filter had 400 taps (the same setting as in [18]). The STFICA algorithm had the preprocessing length of 50 taps, and the separating system had $L = 15$ taps. (The algorithm did not converge with a larger L and longer preprocessing.)

Results of the comparison are summarized in Fig. 14. It contains performance of the Sawada's algorithm [7], which works in the frequency domain, here, using the FFT length 2048 with the overlap of 512 samples.

It can be seen that T-ABCD outperforms STFICA and the Parra's algorithm, but is worse than that of Sawada's algorithm, whose results were taken from the web-site. The latter algorithms take the advantage of utilizing the whole data for the separation. In the case of four sources, performance of the proposed algorithm and the Sawada's algorithm are almost equal in the terms of SDR.

The SIR and the SDR of the Sawada's algorithm can be higher than these quantities for the MMSE, because the MMSE is computed for the filter length equal to $30m$, while the Sawada's algorithm applies filters of the length 2048 taps.

V. Conclusions

The novel time-domain algorithm has been proposed for blind separation of audio sources that is based on the complete unconstrained ICA decomposition of the observation space. The algorithm, named T-ABCD, is suitable for situations where only short data records are available. In this respect, it outperforms other known time-domain BSS algorithms.

T-ABCD consists of five steps, each one providing a room for other variants and improvements. In particular, the selection of eigenmodes may lead to a more effective definition of the observation space. The comparison with the oracle algorithm showed that the measure of the similarity of components, their clustering, and weighting might be still significantly improved.

Finally, as T-ABCD works with short data segments, it has great potential to be modified for online or batch processing needed in situations with moving sources.

References

[1] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, Wiley-Interscience, New York, 2001.

[2] A. Cichocki and S.-I. Amari, *Adaptive Signal and Image Processing: Learning Algorithms and Applications*, Wiley, New York, 2002.

[3] S. C. Douglas, A. Cichocki, and S. Amari, "Self-Whitening Algorithms for Adaptive Equalization and Deconvolution," *IEEE Trans. Signal Processing*, vol. 47, no. 4, pp. 1161-1165, April 1999.

[4] K. Matsuoka and S. Nakashima, "Minimal distortion principle for blind source separation," *Proceedings of 3rd International Conference on Independent Component Analysis and Blind Source Separation (ICA '01)*, pp. 722-727, San Diego, Calif, USA, Dec. 2001.

[5] N. Mitianoudis and M. E. Davies, "Audio Source Separation of Convolutive Mixtures," *IEEE Trans. on Speech Audio Process.*, vol. 11, no. 5, pp. 489-497, Sep. 2003.

[6] Parra, L., and Spence, C.: "Convolutive Blind Separation of Non-Stationary Sources", *IEEE Trans. on Speech and Audio Processing*, Vol. 8, No. 3, pp. 320-327, May 2000.

[7] H. Sawada, R. Mukai, S. Araki, and S. Makino, "A robust and precise method for solving the permutation problem of frequency-domain blind source separation," *IEEE Trans. Speech Audio Processing*, vol. 12, no. 5, pp. 530-538, 2004.

[8] E. Vincent, R. Gribonval, and M.D. Plumbley, "Oracle estimators for the benchmarking of source separation algorithms," *Signal Processing*, 87(8):1933-1950, 2007.

[9] H. Bousbia-Salah, A. Belouchrani, and K. Abed-Meraim, "Jacobi-like algorithm for blind signal separation of convolutive mixtures, " *IEE Elec. Letters*, vol. 37, no. 16, pp. 1049-1050, Aug. 2001.

[10] C. Févotte, A. Debiolles, and C. Doncarli. "Blind separation of FIR convolutive mixtures: application to speech signals," *Proc. 1st ISCA Workshop on Non-Linear Speech Processing*, 2003.

[11] S. Amari, S. C. Douglas, A. Cichocki, and H. H. Yang, "Multichannel Blind Deconvolution and Equalization Using the Natural Gradient," *Proc. IEEE Workshop on Signal Processing Advances in Wireless Communications*, Paris, France, pp. 101-104, April 1997.

[12] S. C. Douglas, H. Sawada, and S. Makino, "Natural gradient multichannel blind deconvolution and speech separation using causal FIR filters, " *IEEE Trans. Speech Audio Processing*, vol. 13, no. 1, pp. 92-104, Jan. 2005.

[13] M. Gupta and S. C. Douglas, "Beamforming Initialization and Data Prewhitening in Natural Gradient Convolutive Blind Source Separation of Speech Mixtures," *Proc. of ICA 2007*, LNCS 4666, pp. 462-470, Sept. 2007.

[14] H. Buchner, R. Aichner, and W. Kellermann, "A Generalization of Blind Source Separation Algorithms for Convolutive Mixtures Based on Second-Order Statistics," *IEEE Trans. on Speech and Audio Processing*, Vol. 13, No. 1, pp. 120-134, Jan. 2005.

[15] H. Buchner, R. Aichner, and W. Kellermann, "TRINICON: A Versatile Framework for Multichannel Blind Signal Processing," *Proc. of IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, Montreal, Canada, vol. 3, pp. 889-892, May 2004.

[16] J. Benesty, S. Makino and J. Chen, *Speech Enhancement*, Springer, March 2005.

[17] J.-F. Cardoso, "On the performance of orthogonal source separation algorithms," *Proc. EUSIPCO*, pp. 776-779, Edinburgh, Sept. 1994.

[18] S. Douglas, M. Gupta, H. Sawada, and S. Makino, "Spatio-temporal FastICA algorithms for the blind separation of convolutive mixtures, " *IEEE Trans. Audio, Speech and Language Processing*, vol. 15, no. 5, pp. 1511-1520, July 2007.

[19] J.-F. Cardoso, "Multidimensional independent component analysis, "*Proc. of ICASSP 1998*, Vol. 4, pp. 1941-1944, 12-15 May 1998.

[20] P. Tichavský and A. Yeredor, "Fast Approximate Joint Diagonalization Incorporating Weight Matrices, " *IEEE Transactions of Signal Processing*, vol. 57, no.3, pp. 878-891, March 2009.

[21] Z. Koldovský, P. Tichavský, "Time-domain blind audio source separation using advanced component clustering and reconstruction," *Proc. of HSCMA 2008*, Trento, Italy, pp. 216-219, May 2008.

[22] A. Cichocki, S. Amari, K. Siwek, T. Tanaka, Anh Huy Phan, and R. Zdunek, ICALAB MATLAB Toolbox Ver. 3 for Signal Processing, [online], `http://www.bsp.brain.riken.jp/ICALAB`.

[23] Z. Koldovský and P. Tichavský, "A Comparison of Independent Component and Independent Subspace Analysis Algorithms, " *Proc. of EUSIPCO 2009*, pp. 1447-1451, Glasgow, August 24-28, 2009.

[24] Z. Koldovský, P. Tichavský and E. Oja, "Efficient Variant of Algorithm FastICA for Independent Component Analysis Attaining the Cramér-Rao Lower Bound," *IEEE Trans. on Neural Networks*, Vol. 17, No. 5, Sept 2006.

[25] A. Hyvärinen, "Fast and Robust Fixed-Point Algorithms for Independent Component Analysis", *IEEE Transactions on Neural Networks*, 10(3):626-634, 1999.

[26] K. Abed-Meraim and A. Belouchrani, "Algorithms for Joint Block Diagonalization," *Proc. of EUSIPCO 2004*, pp. 209212, Vienna, Austria, 2004.

[27] C. Févotte and F. J. Theis, "Orthonormal approximate joint block-diagonalization," *Technical Report GET/Tlcom Paris 2007D007*, 2007.

[28] Shalvi, O.; Weinstein, E., "Maximum likelihood and lower bounds in system identification with non-Gaussian inputs", *IEEE Tr. Information Theory*, Vol. 40, No. 2, March 1994, pp. 328 - 339.

[29] D-T. Pham and J-F. Cardoso. "Blind separation of instantaneous mixtures of non stationary sources," *IEEE Trans. Signal Processing*, pp. 1837-1848, vol. 49, no. 9, 2001.

[30] J. C. Principe, B. de Vries, Guedes de Oliveira, "Generalized feedforward structures: a new class of adaptive filters, " *Proc. of ICASSP 1992*, Vol. 4, pp. 245-248, March 1992.

[31] K. E. Hild II, D. Erdogmuz, J. C. Principe, "Experimental Upper Bound for the Performance of Convolutive Source Separation Methods", *IEEE Trans. on Signal Processing*, Vol. 54, No. 2, pp. 627-635, Feb. 2006.

[32] D. Schobben, K. Torkkola, and P. Smaragdis, "Evaluation of blind signal separation methods," *Int. Workshop on Independent Component Analysis and Signal Separation (ICA '99)*, Aussois, France, pp. 261-266, Jan. 1999.

[33] Z. Koldovský, Matlab^{TM} p-code of the proposed algorithm, [online], `http://itakura.ite.tul.cz/zbynek/downloads.htm`

[34] Z. Koldovský and P. Tichavský, "Způsob slepé separace akustických signálů z jejich konvolutorní směsi," Czech Patent Application PV 2008-752, Nov. 2008.

Zbyněk Koldovský (S'03-M'04) was born in Jablonec nad Nisou, Czech Republic, in 1979. He received the M.S. degree and Ph.D. degree in mathematical modeling from Faculty of Nuclear Sciences and Physical Engineering at the Czech Technical University in Prague in 2002 and 2006, respectively. He is currently with the Institute of Information Technology and Electronics, Technical University of Liberec. He has also been with the Institute of Information Theory and Automation of the Academy of Sciences of the Czech Republic since 2002.

His main research interests are in audio signal processing, blind source separation and independent component analysis.

Petr Tichavský (M'98, SM'04) received the M.S. degree in mathematics in 1987 from the Czech Technical University, Prague, Czechoslovakia and the Ph.D. degree in theoretical cybernetics from the Czechoslovak Academy of Sciences in 1992. Since that time he is with the Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic in Prague. In 1994 he received the *Fulbright grant* for a 10 month fellowship at Yale University, Department of Electrical Engineering, in New Haven, U.S.A. In 2002 he received the *Otto Wichterle Award* from Academy of Sciences of the Czech Republic.

He is author and co-author of research papers in the area of sinusoidal frequency/frequency-rate estimation, adaptive filtering and tracking of time varying signal parameters, algorithm-independent bounds on achievable performance, sensor array processing, independent component analysis and blind signal separation.

Petr Tichavský served as associate editor of the *IEEE Signal Processing Letters* from 2002 to 2004, and as associate editor of the *IEEE Transactions on Signal Processing* from 2005 to 2009. Since 2009 he is a member of the IEEE Signal Processing Society's Signal Processing Theory and Methods (SPTM) Technical Committee. Petr Tichavský also serves as a general chair of the 36th IEEE Int. Conference on Acoustics, Speech and Signal Processing ICASSP 2011 in Prague, Czech Republic.

Subband Blind Audio Source Separation Using a Time-Domain Algorithm and Tree-Structured QMF Filter Bank

Zbyněk Koldovský[1,2], Petr Tichavský[2], and Jiří Málek[1]

[1] Institute of Information Technology and Electronics
Technical University of Liberec, Studentská 2, 461 17 Liberec, Czech Republic
zbynek.koldovsky@tul.cz,
http://itakura.ite.tul.cz/zbynek
[2] Institute of Information Theory and Automation, Pod vodárenskou věží 4,
P.O. Box 18, 182 08 Praha 8, Czech Republic
p.tichavsky@ieee.org,
http://si.utia.cas.cz/Tichavsky.html

Abstract. T-ABCD is a time-domain method for blind linear separation of audio sources proposed by Koldovský and Tichavský (2008). The method produces short separating filters (5-40 taps) and works well with signals recorded at the sampling frequency of 8-16 kHz. In this paper, we propose a novel subband-based variant of T-ABCD, in which the input signals are decomposed into subbands using a tree-structured QMF filter bank. T-ABCD is then applied to each subband in parallel, and the separated subbands are re-ordered and synthesized to yield the final separated signals. The analysis filter of the filter bank is carefully designed to enable maximal decimation of signals without aliasing. Short filters applied within subbands then result in sufficiently long filters in fullband. Using a reasonable number of subbands, the method yields improved speed, stability and performance at an arbitrary sampling frequency.

1 Introduction

Blind separation (BSS) of simultaneously active audio sources is a challenging problem within audio signal processing. The goal is to retrieve d audio sources from their convolutive mixtures recorded by m microphones. The model is described by

$$x_i(n) = \sum_{j=1}^{d} \sum_{\tau=0}^{M_{ij}-1} h_{ij}(\tau) s_j(n-\tau), \quad i = 1, \ldots, m, \tag{1}$$

where $x_1(n), \ldots, x_m(n)$ are the observed signals on microphones and $s_1(n), \ldots, s_d(n)$ are the unknown original (audio) signals. This means that the mixing system

[0] This work was supported by Ministry of Education, Youth and Sports of the Czech Republic through the project 1M0572 and by Grant Agency of the Czech Republic through the project 102/09/1278.

is a MIMO (multi-input multi-output) linear filter with source-microphone impulse responses h_{ij}'s each of length M_{ij}. Linear separation consists in finding a MIMO filter that inverts the mixing process (1) and yields estimates of the original signals $s_1(n), \ldots, s_d(n)$. It is convenient to assume the independence of $s_1(n), \ldots, s_d(n)$, and the separation can be based on Independent Component Analysis (ICA) [1]. Indeterminacies that are inherent to the ICA cause that the original colorations of $s_1(n), \ldots, s_d(n)$ cannot be retrieved. The goal is therefore to estimate their microphone responses (images), which only have properly defined colorations. The response of the kth source on the ith microphone is

$$s_k^i(n) = \sum_{\tau=0}^{M_{ik}-1} h_{ik}(\tau) s_k(n-\tau). \tag{2}$$

To apply the ICA, the convolutive mixture (1) must be transformed into an instantaneous one. This is done either directly in the time-domain (**TD**) by decomposing a matrix usually constructed of delayed copies of signals from microphones, or in the frequency-domain (**FD**) where the signals are transformed by the Short-Time Fourier Transform (**STFT**) that converts the convolution operation into the ordinary multiplication. Weaknesses of both approaches are well known from literature. The **FD** approach meets the so-called permutation problem [2] due to inherent indeterminacies in ICA and requires long data to generate sufficient number of samples for each frequency bin. On the other hand, TD methods are computationally more expensive due to simultaneous optimization of all filter coefficients, which restrict their ability to compute long filters.

A reasonable compromise is the subband approach [3] that consists in decomposing the mixed signals into subbands via a filter bank, separating each subband by a TD method, permuting the separated subbands, and synthesizing the final signals. If a moderate number of subbands is chosen, the permutation problem becomes less difficult compared to the FD approach. Since the subband signals are decimated, the length of separating filters is multiplied.

Several subband approaches have already been proposed in literature using various filter banks. The method in [5] uses a uniform DFT filter bank. Araki et al. [3] use a polyphase filter bank with a single side-band modulation. In [6, 7], uniform FIR filter banks were used. All the referenced methods do not apply the maximal decimation of signals in order to reduce the aliasing between subbands. This restrains both the computational efficiency and the effective length of separating filters.

We propose a novel subband method designed to be maximally effective in this respect. The signals are decomposed uniformly into 2^M subbands using a two-channel QMF filter bank applied recursively in the full-blown 2-tree structure with M levels [4]. The signals are decimated by 2 in each level of the 2-tree so they are finally decimated by 2^M, which means *maximal decimation*. Through a careful design of a halfband FIR filter, which determines the whole filter bank, the aliasing is avoided. The blind separation within subbands is then carried out independently by the T-ABCD method [10], which is robust and effective in estimating short separating filters. The permutation problem due to the random

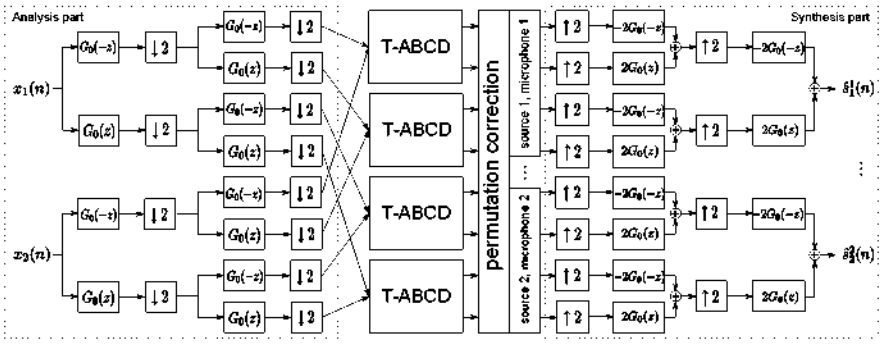


Fig. 1. An illustration of the proposed subband BSS algorithm using the QMF tree-structured filter bank with $M = 2$. Here, two-microphone recording is separated into two responses of each of two original sources.

order of separated signals in each subband is solved by comparing correlations of absolute values of signals [2]. Finally, the reordered signals are synthesized to yield the estimated responses (2). The flow of the method is illustrated in Fig. 1.

The following section gives more details on the proposed method, and Section 3 demonstrates its performance by experiments done with real-world signals.

2 Proposed Subband BSS Method

2.1 Subband Decomposition

The building block of the tree-structured subband decomposition applied in the proposed method is a two-channel bank that separates the input signals into two bands. In general, a two-channel bank consists of two analysis filters and two synthesis filters whose transfer functions are, respectively, $G_0(z)$, $G_1(z)$, $H_0(z)$ and $H_1(z)$. The input signal is filtered by $G_0(z)$ and $G_1(z)$ in parallel, and the outputs are decimated by 2 giving the subband signals. After the subband processing, the signals are expanded by 2 and passed through the synthesis filters and are added to yield the output signal.

The analysis filters of a Quadrature Mirror Filter (QMF) bank satisfy

$$G_1(z) = G_0(-z). \tag{3}$$

$G_0(z)$ should be a low-pass filter with the pass band $[-\pi/2, \pi/2]$ so that the decimated signals are not aliased. The synthesis filters may be defined as

$$H_0(z) = 2G_1(-z), \qquad H_1(z) = -2G_0(-z). \tag{4}$$

Then the whole two-channel QMF bank is determined by $G_0(z)$.

(4) is a sufficient condition for eliminating the aliasing from synthesized signals provided that no subband processing is done, i.e., when the signals are expanded immediately after the decimation (equation (12.58) in [4]). In such special

case, the transfer function of the two-channel QMF bank is $[G_0(z)]^2 - [G_0(-z)]^2$. It follows that the bank does not possess the perfect reconstruction property in general, which is nevertheless not as important in audio applications. While phase distortions are avoided provided that $G_0(z)$ has a linear phase, amplitude distortions can be made inaudible by a careful design of the filter[3].

To decompose the signal into more than two bands, the analysis part of the two-channel QMF bank can be applied recursively to split each band into two subbands etc. If the depth of the recursion is M, the filter bank splits the spectrum uniformly into 2^M subbands. This approach is utilized in the proposed method as demonstrated by Fig. 1. After the processing of subbands, the synthesis is done backwards then the analysis.

2.2 Separation Algorithm: T-ABCD

T-ABCD is an ICA-based method for blind separation of audio signals working in time-domain. It is based on the estimation of all independent components (ICs) of an *observation space* by an incorporated ICA algorithm. The observation space is spanned by rows of a data matrix $\mathbf{X}$ that may be defined in a general way [10]. For simplicity, we will consider the basic definition that is common to other TD methods [12]: Rows of $\mathbf{X}$ contain L time-shifted copies of each observed signal $x_1(n), \ldots, x_m(n)$. The number of rows of $\mathbf{X}$ is mL, which is the dimension of the observation space. Linear combinations of rows of $\mathbf{X}$ correspond to outputs of FIR MISO filters of the length L (hence also the ICs of $\mathbf{X}$). The steps of T-ABCD are as follows.

1. Find all mL independent components of $\mathbf{X}$ by an ICA algorithm.
2. Group the components into clusters so that each cluster contains components corresponding to the same original source.
3. For each cluster, use components of the cluster to reconstruct microphone responses (images) of a source corresponding to the cluster.

For more details on the method see [9] and [10].

A shortcoming of T-ABCD is that its computational complexity grows rapidly with L. On the other hand, T-ABCD is very powerful when L is reasonably low ($L = 1, \ldots, 40$). This is because all ICs of $\mathbf{X}$ are estimated without applying any constraint to the separating MISO filters (step 1), and all ICs are used to reconstruct the sources' responses (steps 2 and 3). The performance of T-ABCD is robust as it is independent of an initialization provided that the applied ICA algorithm in step 1 is equivariant. Consequently, the use of T-ABCD within the subband separation is desirable, because the separating filters in subbands are shorter than those in fullband [3].

[3] We have chosen $G_0(z)$ as an equiripple FIR filter [4] with 159 taps having the minimum attenuation of 60 dB in the stopband. To eliminate the aliasing, the stop-frequency was shifted slightly from $\pi/2$ to the left by $\epsilon \approx 0.01$, which is small enough so that the cut-off band around $\pi/2$ is very narrow and results in inaudible distortions of signals.

2.3 The permutation problem

The estimated responses of sources by T-ABCD are randomly permuted due to indeterminacy of ICA or, more specifically, due to the indeterminacy of the order of clusters identified by step 2. Since the permutation might be different in each subband, the estimated signals in subbands must be aligned before synthesizing them.

Let $\hat{s}^i_{k,j}(n)$, $k = 1, \ldots, d$ be the not yet sorted estimates of responses of the sources at the ith microphone in the jth subband. We wish to find permutations $\pi_j(k)$, $j = 1, \ldots, M$ such that $\hat{s}^i_{\pi_j(k),j}(n)$ is the estimated response of the kth source at the microphone in the subband. We shall assume, for convenience, that the order of the components in one, say in the j_1th subband (e.g. $j_1 = 1$), is correct. Therefore we set $\pi_{j_1}(k) = k$, $k = 1, \ldots, d$. Permutations in all other subbands can be found by maximizing the following criterion,

$$d(p,q,r,s) = \sum_{i=1}^{m} \left|\mathrm{cov}\left(|\widehat{s}^i_{p,q}(n)|, |\widehat{s}^i_{r,s}(n)|\right)\right| =$$

$$= \sum_{i=1}^{m} \frac{1}{T} \sum_{n=1}^{T} \left(|\widehat{s}^i_{p,q}(n)| - \frac{1}{T}\sum_{t=1}^{T} |\widehat{s}^i_{p,q}(t)|\right)\left(|\widehat{s}^i_{r,s}(n)| - \frac{1}{T}\sum_{t=1}^{T} |\widehat{s}^i_{r,s}(t)|\right) \quad (5)$$

that compares dynamic profiles (absolute values) of the signals [2], as follows.

1. Put $\mathcal{S} = \{j_1\}$, a set of already permuted subbands.
2. Find $j_2 = \arg\max_{s \notin \mathcal{S}} \{\max_{p,r} d(p, j_1, r, s)\}$.
3. Use the greedy algorithm to find $\pi_{j_2}(\cdot)$ by maximizing $d(\cdot, j_1, \cdot, j_2)$. Namely, define $\mathcal{P} = \emptyset$ and $\mathcal{R} = \emptyset$, and repeat
   (a) $(p, r) = \arg\max_{p \notin \mathcal{P}, r \notin \mathcal{R}} d(p, j_1, r, j_2)$
   (b) put $\pi_{j_2}(p) = r$
   (c) $\mathcal{P} = \mathcal{P} \cup \{p\}$, $\mathcal{R} = \mathcal{R} \cup \{r\}$
   until $\mathcal{P} \subsetneq \{1, \ldots, M\}$
4. $\mathcal{S} = \mathcal{S} \cup \{j_2\}$, $j_1 = j_2$.
5. If $\mathcal{S} \subsetneq \{1, \ldots, M\}$, go to 2.

3 Experiments

To demonstrate the performance of the proposed method, we test it on selected data from the SiSEC 2010 campaign[4]. The data consists of two-microphone real-world recordings of, respectively, two male and two female speakers played over loudspeakers (signal combinations #1 and #2) placed in room #1 in position #1 shown in Fig. 2. Each source was recorded separately to obtain its microphone responses, and the signals were summed to obtain the mixed signals; the original sampling rate was 44.1kHz.

[4] The task "Robust blind linear/non-linear separation of short two-sources-two-microphones recordings" in the "Audio source separation" category; see `http://sisec.wiki.irisa.fr/tiki-index.php`

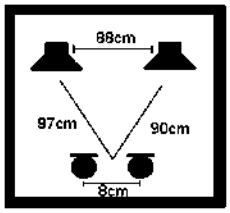


Fig. 2. The position of microphones and loudspeakers in the experiment.

For evaluation of the separation, we use the standard Signal-to-Interference Ratio (SIR), as defined in [13]. The evaluation is computed using the full length of recordings, which is about 2 seconds, but only the first second of the data was used for computations of separating filters.

We compare the original T-ABCD from [9] working in fullband with the proposed subband T-ABCD decomposing signals into 2, 4, 8, and 16 subbands, that is, with $M = 1, \ldots, 4$. The fullband T-ABCD is applied with $L = 20$, while in subbands $L = 10$ is taken. The other parameters of T-ABCD are the same both in fullband and subband; namely, the weighting parameter is $\alpha = 1$, and the BGSEP algorithm from [11] is used for finding ICs of $\mathbf{X}$.

Fig. 3 shows the results of experiments done with signals resampled to the sampling rates f_s =8, 16, 32, and 44.1 kHz, respectively. The performance of the fullband T-ABCD decreases with the growing f_s. This is due to the fact that the effective length of separating filters decreases as L is fixed to 20. A comparable length of filters is applied in the 2-subbands method, where $L = 10$ in each subband. The performance of the 2-subbands method is either comparable ($f_s = 8$ and 32 kHz) to the fullband method or even better ($f_s = 16$ and 44.1 kHz) and does not decrease until $f_s \leq 16$. This points to the fact that the fullband method suffers from increased bandwidth of signals when f_s grows.

As can be seen from Fig. 3, the performance of the subband method does not automatically increase with the number of subbands. This is mainly caused by the permutation problem, which becomes more difficult with the growing number of subbands. The results indicate that the optimal bandwidth of subbands is between 2-5 kHz. Namely, (1) the 4-subbands method performs best at $f_s = 16$ and 32 kHz, (2) the 8-subbands method provides the best results when $f_s = 32$ and 44.1 kHz, and (3) the 16-subbands method seems to be effective if $f_s = 44.1$ kHz. On the other hand, the decomposition of signals into 16 subbands seems to be inadequate when $f_s = 8$ or 16 kHz, as the 16-subbands method yields unstable performance here.

3.1 Computational aspects

The methods were running on a PC with quad-core i7 2.66 GHz processor in Matlab^{TM} with Parallel Computing Toolbox^{TM}. There were four running workers, i.e. one for each core of the processor, which means that up to four T-ABCDs

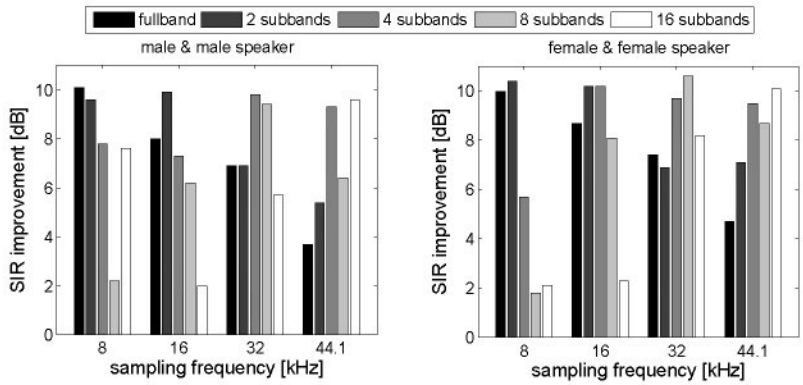


Fig. 3. SIR improvement achieved by the separation. Each value is an average over both separated sources and estimated microphone responses.

may run simultaneously in subbands. The average computational burden summarizes Table 1 in the form A/B, where A and B denote the time needed for separation without and with the aid of parallel computations, respectively. The parallelization was realized through the parallel for-cycle (`parfor`).

The values in Table 1 prove the advantage of the subband method consisting in lower computational complexity. Although the parallelization by means of the Parallel Computing Toolbox^{TM} is not that effective, it points to the potential improvement in terms of speed. For example, the 4-subband method should be almost four-times faster when running in parallel, since about 80% of the computational burden is caused by T-ABCD, while the permutation correction takes about 3% and the rest is due to the filtering operations.

Table 1. Average time needed per separation without and with parallel computations.

	computational time [s]			
	8kHz	16kHz	32kHz	44.1kHz
fullband	0.42/ -	0.90/ -	2.03/ -	2.84/ -
2-subband	0.25/0.17	0.46/0.31	0.90/0.69	1.35/0.97
4-subband	0.30/0.19	0.56/0.33	1.06/0.65	1.51/0.97
8-subband	0.40/0.25	0.66/0.42	1.26/0.83	1.83/1.13
16-subband	0.56/0.35	0.87/0.55	1.50/0.99	2.20/1.39

4 Conclusion

The proposed subband T-ABCD was shown to be an improved variant of T-ABCD in terms of speed and separation performance, especially, when working

with signals sampled at sampling rates higher than 16 kHz. The method is able to separate one second of data in a lower time, which points to its applicability in a batch-online processing. The optimum number of subbands depends on the sampling frequency, which was shown to correspond to the bandwidth of about 2-5kHz per subband. Experiments not shown here due to lack of space show that the subband T-ABCD might be combined with other filter banks (e.g. [3]) as well, but the analysis filters must be adjusted to avoid the aliasing in maximally decimated signals.

References

1. Comon, P.: Independent component analysis: a new concept? Signal Processing, vol. 36, 287-314 (1994)
2. Sawada, H., Mukai, R., Araki, S., Makino, S.: A robust and precise method for solving the permutation problem of frequency-domain blind source separation. IEEE Trans. Speech Audio Processing, vol. 12, no. 5, 530-538 (2004)
3. Araki, S., Makino, S., Aichner, R., Nishikawa, T., Saruwatari, H.: Subband-based blind separation for convolutive mixtures of speech. IEICE Trans. Fundamentals, vol. E88-A, no. 12, 3593-3603 (2005)
4. Porat, B.: A course in digital signal processing. John Wiley & Sons, New York (1997)
5. Grbić, N., Tao, X.-J., Nordholm, S.E., Claesson, I.: Blind signal separation using overcomplete subband representation. IEEE Transactions on Speech and Audio Processing, vol. 9, no. 5, 524-533 (2001)
6. Russell, I., Jiangtao Xi, Mertins, A., Chicharo, J.: Blind source separation of nonstationary convolutively mixed signals in the subband domain. In: ICASSP 2004, vol. 5, pp. 481-484 (2004)
7. Kokkinakis, K., Loizou, P.C.: Subband-based blind signal processing for source separation in convolutive mixtures of speech. In: ICASSP 2007, vol. 4, pp. 917-920 (2007)
8. Duplessis-Beaulieu, F., Champagne, B.: Fast convolutive blind speech separation via subband adaptation. In: ICASSP 2003, vol. 5, pp. 513-516 (2003)
9. Koldovský, Z., Tichavský, P.: Time-domain blind audio source separation using advanced component clustering and reconstruction. In: HSCMA 2008, Trento, Italy, pp. 216-219 (2008)
10. Koldovský, Z., Tichavský, P.: Time-domain blind separation of audio sources on the basis of a complete ICA decomposition of an observation space. Accepted for publication in IEEE Trans. on Audio, Language, and Speech Processing (April 2010)
11. Tichavský, P., Yeredor, A.: Fast approximate joint diagonalization incorporating weight matrices. IEEE Transactions of Signal Processing, vol. 57, no.3, 878-891 (2009)
12. Bousbia-Salah, H., Belouchrani, A., Abed-Meraim, K.: Jacobi-like algorithm for blind signal separation of convolutive mixtures. IEE Elec. Letters, vol. 37, no. 16, 1049-1050 (2001)
13. Vincent, E., Févotte, C., Gribonval, R.: Performance measurement in blind audio source separation. IEEE Trans. Audio, Speech and Language Processing, vol. 14, no. 4, 1462-1469 (2006)

Part III

Present and Future Works

8.1 BSS Methods in Audio Signal Processing

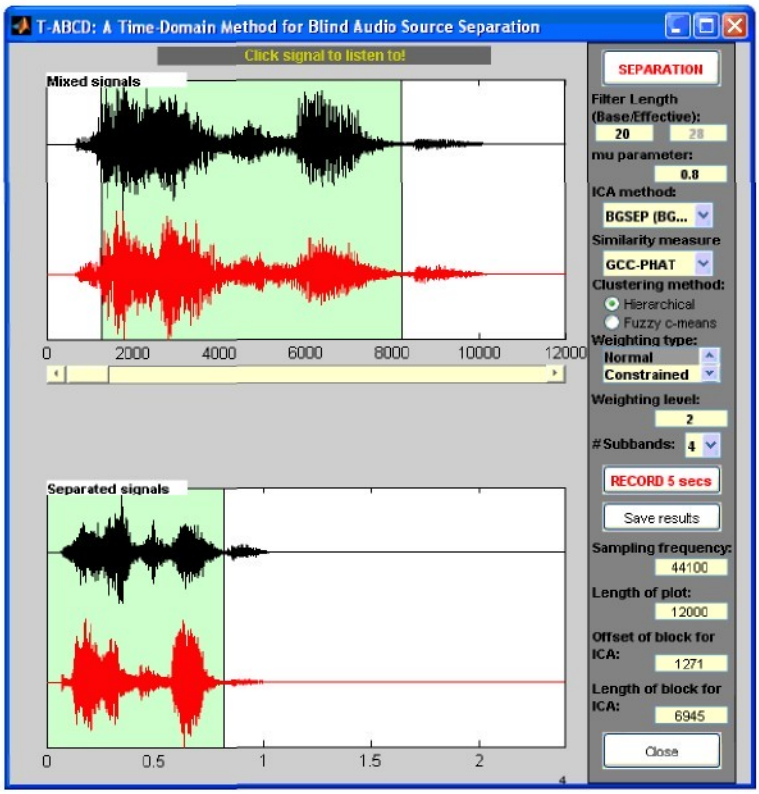


Figure 8.1: An illustration of the Matlab graphical user interface designed for the T-ABCD algorithm and its extensions.

Future works in the field of audio signal processing will be especially focused on extensions and modifications of the T-ABCD method. The concept of T-ABCD is flexible in that other BSS approaches can be applied for

1. a more adaptive construction of the observation matrix (6.2) [c18], especially for its dimension reduction, and

2. the decomposition of the observation matrix into components corresponding to individual sources or directions of arrival.

For instance, methods for Non-negative Matrix or Tensor Factorization or Sparse Component Analysis can be considered for a decomposition of a (transformed) observation matrix as they are, like ICA, able to retrieve individual signals or their parts. Another possibility here is to use methods that are able to work with

underdetermined models, thereby allow T-ABCD to separate more sources from a lower number of microphones.

Another modification of T-ABCD aims at real-time processing of signals. Recently, a prototype of an on-line version has been proposed in [c20] and its variant implemented in C++ in [m9]. In this area, the main need is to simplify the algorithm considerably owing to its computational burden and implementation complexity. A great potential provide parallel computations that are possible within T-ABCD. Other simplifications are possible when considering variants of the algorithm that are tailored to particular applications (e.g. a separator working with two microphones only assuming maximum two sources).

Finally, further audio applications such as speech enhancement and dereverberation can also be considered. A motivation comes from the fact that independent components possess interesting properties that can be measured, namely, directivity, non-Gaussianity, sparsity or non-stationarity. Methods using these features can be developed in order to reconstruct dereverberated and/or enhanced signals.

A project P103/11/1947 devoted to these topics has already been accepted by the Grant Agency of the Czech Republic and is going to start in 2011.

8.2 CP Tensor Decomposition

Recently, signal processing researchers have payed a lot of attention to tensor decomposition techniques. In fact, it is a renewed topic of multi-linear algebra whose origin comes from the beginning of the 20th century [60]. It was revisited in the 1970s by psychometricians and chemometricians [61].

The goal is to analyze multidimensional (three-way and higher-way) data arrays, usually called tensors, through decomposing them into a sum of low rank arrays. The problem can be seen as an extension of a low rank decomposition of matrices. The appealing feature of higher-way arrays is that, compared to matrices, their such decompositions can be essentially unique (i.e. up to some scaling and permutation ambiguities).

Let $\mathcal{X}$ be a three-way tensor of dimensions $I \times J \times K$ whose ijkth element is X_{ijk}. Parallel factor analysis (PARAFAC), or Canonical decomposition (CANDECOMP), or CP decomposition, consists in finding factor matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ of dimensions, respectively, $I \times R$, $J \times R$ and $K \times R$, whose ijth elements are A_{ij}, B_{ij}, and C_{ij}, such that it holds

$$X_{ijk} = \sum_{r=1}^{R} A_{ir} B_{jr} C_{kr}, \tag{8.7}$$

for $i = 1, \ldots, I$, $j = 1, \ldots, J$, and $k = 1, \ldots, K$. The smallest R such that the decomposition exist is the rank of $\mathcal{X}$. The definition for higher-way tensors is analogous.

There are several alternative notations of (8.7); see e.g. [62]. For example, it can be written that

$$\mathcal{X} = \mathcal{I} \times_1 \mathbf{A} \times_2 \mathbf{B} \times_3 \mathbf{C} \tag{8.8}$$

where $\mathcal{I}$ is the identity tensor of the dimensions $R \times R \times R$, and $\times_i$ denotes the matrix multiplication in the mode i, $i = 1, 2, 3$, that is, the matrix multiplication of columns, rows or tubes of $\mathcal{X}$ when these are unfolded as column vectors. It can be shown that the order of operations $\times_i$ can be arbitrary.

Another notation is

$$\mathcal{X} = \sum_{r=1}^{R} \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r, \tag{8.9}$$

where $\circ$ denotes the tensor outer product, and $\mathbf{a}_r$, $\mathbf{b}_r$, and $\mathbf{c}_r$ denote, respectively, the rth column of $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$. The notation (8.9) reveals clearly that $\mathcal{X}$ is decomposed into the sum of R rank-one tensors.

8.2.1 Essential Uniqueness

The CP decomposition is *essentially unique* if the factor matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are unique up to rescaling and jointly permuting their columns. This definition comes from the fact that (8.7) remains unchanged whenever changing $\mathbf{A} \rightarrow \mathbf{APT}_1$, $\mathbf{B} \rightarrow \mathbf{BPT}_2$ and $\mathbf{C} \rightarrow \mathbf{CPT}_3$ where $\mathbf{P}$ is a $R \times R$ permutation matrix, and $\mathbf{T}_1$, $\mathbf{T}_2$ and $\mathbf{T}_3$ are diagonal matrices such that $\mathbf{T}_1\mathbf{T}_2\mathbf{T}_3 = \mathbf{I}_R$. $\mathbf{I}_R$ denotes the $R \times R$ identity matrix.

The essential uniqueness of CP is an important issue as it entails identifiability of the factor matrices from the tensor. A sufficient condition was derived by Kruskal in [61]. He showed that if

$$k_A + k_B + k_C \geq 2R + 2 \tag{8.10}$$

then the CP decomposition is essentially unique. Here k_A is the so-called k-rank (Kruskal rank) of $\mathbf{A}$, which is the largest integer number k such that every subset of k columns of $\mathbf{A}$ is linearly independent. Recently, Sidiropoulos and Bro generalized the condition for higher than three-way arrays in [63]. The problem has been addressed again, for instance, Stegeman et al. derived a condition that is closer to the necessity in [64].

Lim and Comon recently pointed to a simple inequality between the k-rank and the so-called coherence of a matrix in [68]. The coherence of $\mathbf{A}$ is defined as $\mu(\mathbf{A}) = \max_{i,j} |(\bar{\mathbf{a}}_i, \bar{\mathbf{a}}_j)|$ where $\bar{\mathbf{a}}_i$ denotes the normalized column $\mathbf{a}_i$ (assuming

that there are no zero columns). It follows that $0 \leq \mu(\mathbf{A}) \leq 1$, and the coherence is equal to zero if the columns of $\mathbf{A}$ are orthogonal and is equal to one if there are two co-linear columns at least. The inequality says that, if the columns of $\mathbf{A}$ are linearly dependent[1], then

$$k_A \geq \frac{1}{\mu(\mathbf{A})}, \tag{8.11}$$

so the Kruskal's sufficient condition can then be approximated by

$$\frac{1}{2}\left[\frac{1}{\mu(\mathbf{A})} + \frac{1}{\mu(\mathbf{B})} + \frac{1}{\mu(\mathbf{C})}\right] > R. \tag{8.12}$$

Both conditions (8.10) and (8.12) point to the fact that the essential uniqueness of the CP decomposition requires that the columns of factor matrices should be "as linearly independent as possible".

The author of this thesis is presently collaborating on a paper [c22] that addresses stability of the CP decomposition. The study is done by deriving the Cramér-Rao lower bound on variance of an unbiased estimate of the tensor parameters, i.e. elements of its factor matrices, from its noisy observation (the tensor plus a random Gaussian i.i.d. tensor). The existence of the bound reveals necessary conditions for essential uniqueness of the CP decomposition, moreover, for identifiability of each column of each factor matrix separately.

8.2.2 Decomposition Algorithms

To compute the decomposition (8.7), i.e. the factor matrices, several algorithms have been proposed. Given a tensor $\mathcal{X}$, they mostly consist in minimizing a quadratic criterion

$$\left\|\mathcal{X} - \mathcal{I} \times_1 \mathbf{A} \times_2 \mathbf{B} \times_3 \mathbf{C}\right\|_2^2 \tag{8.13}$$

where $\|\cdot\|_2$ denotes the quadratic norm (the square root of sum of squared elements of the argument).

A basic approach is the alternative least square (ALS) algorithm that proceeds iteratively, and minimizes the criterion with respect to individual factor matrices one by one. Convergence of this approach is known to be usually slow, especially, if some of the factor matrices contain nearly co-linear columns. There is also a high risk that the algorithm converges to a local minimum of (8.13). There is a modification of the ALS algorithm using a technique called Enhanced Line Search (ELS), which was proposed in [65].

[1] The condition that columns of the matrix must be linearly dependent is missing in [68], but it follows from the proof of the respective lemma (Lemma 5.5). Note that if the columns are linearly independent, then the k-rank of the matrix is equal to its rank.

More robust methods were derived by optimizing all factor matrices simultaneously. Paatero et al. proposed a method called PMF3 in [66] for the CP decomposition of three-way arrays. It comes from the damped Gauss-Newton or Levenberg-Marquardt optimization algorithms designed for quadratic criteria like (8.13). The problem of these optimization techniques consists in their computational complexity. An iteration of these algorithms requires the computation of the inversion matrix of an approximate hessian matrix whose dimensions are $R(I+J+K)\times R(I+J+K)$. Some simplified versions of PMF3 were proposed in [c21]. Further simplifications and generalizations of PMF3 to higher-way tensors were proposed in [67].

8.3 Underdetermined BSS

The CP decomposition can be used to identify the mixing matrix in the basic instantaneous ICA model $\mathbf{X} = \mathbf{AS}$. Let the number of columns of $\mathbf{A}$ (the number of sources) be R. Consider a set of statistics $\mathbf{T}_t(\mathbf{X})$, $t = 1, \ldots, T$, that all satisfy

$$\mathrm{E}[\mathbf{T}_t(\mathbf{X})] = \mathbf{A}\,\mathrm{E}[\mathbf{T}_t(\mathbf{S})]\,\mathbf{A}^T \tag{8.14}$$

and $\mathrm{E}[\mathbf{T}_t(\mathbf{S})]$ are diagonal. Since (8.14) can be written in the form

$$\mathrm{E}[\mathbf{T}_t(\mathbf{X})] = \sum_{r=1}^{R} \lambda_{tr}\mathbf{a}_r\mathbf{a}_r^T, \tag{8.15}$$

where $\lambda_{tr} = (\mathrm{E}[\mathbf{T}_t(\mathbf{S})])_{rr}$ is the trth element of the matrix $\mathbf{\Lambda}$, it is easily seen that the matrices $\mathbf{A}$ and $\mathbf{\Lambda}$ might be identified as the factor matrices of the tensor

$$\mathcal{X} = \mathcal{I} \times_1 \mathbf{A} \times_2 \mathbf{\Lambda} \times_3 \mathbf{A}, \tag{8.16}$$

whose tth frontal slice, in Matlab notation $\mathcal{X}_{:,:,t}$, is equal to $\mathrm{E}[\mathbf{T}_t(\mathbf{X})]$.

An important fact is that the identifiability of $\mathbf{A}$ and $\mathbf{\Lambda}$ is conditioned by the essential uniqueness of the CP decomposition of $\mathcal{X}$. It follows that the number of columns of $\mathbf{A}$, which is R, can be higher then its number of rows. In other words, there can be more sources than sensors in the ICA mixing model (see Section 1.2.3).

It is good to remind that the estimation of $\mathbf{A}$ and the separation of the original sources $\mathbf{S}$ are not equivalent tasks in the underdetermined case, because $\mathbf{A}$ is not invertible. To retrieve $\mathbf{S}$, some other techniques must be applied, e.g., a beamforming. Note that some information about the sources is contained in the matrix $\mathbf{\Lambda}$, which can be used for the sources retrieval.

In practice, the expectation values of the statistics $\mathbf{T}_t(\mathbf{X})$, $t = 1, \ldots, T$, are not known. Therefore, the tensor is defined so that

$$\mathcal{Y}_{:,:,t} = \mathbf{T}_t(\mathbf{X}), \tag{8.17}$$

which can be seen as a noisy observation of $\mathcal{X}$, that is,

$$\mathcal{Y} = \mathcal{X} + \mathcal{E} \tag{8.18}$$

where $\mathcal{E}$ is a tensor whose entries contain random errors. The CP decomposition of $\mathcal{Y}$ is done in an approximate way, where its rank (the number of sources) is either known in advance or must be estimated. However, the estimation of the rank is a difficult task, because the rank of $\mathcal{Y}$ is usually different from that of $\mathcal{X}$, and even the computation of the rank of $\mathcal{X}$ is likely NP hard [68].

8.3.1 Algorithms

The second-order statistics satisfying (8.14) are the covariance and cross-covariance matrices considered in ICA methods based on the nonstationarity and spectral diversity principles (Sections 2.3.2 and 2.3.3). For example, the SOBIUM algorithm from [69] can be seen as an extension of the SOBI algorithm from [12] to underdetermined mixtures.

Higher-order statistics satisfying (8.14) are the cumulants. For example, fourth-order statictics are used by the FOBIUM algorithm from [70]. Cumulants can also be used to define higher than three-way tensors, and thanks to the multi-linearity of cumulants and their further features, the blind identification od $\mathbf{A}$ is possible with higher number of sources R than in case of methods decomposing three-way tensors only [72, 71].

Similarly to how the WASOBI, BGSEP or BARBI algorithms (Section 2.5.2) were derived, it is possible to introduce weights into the quadratic criterion (8.13), to make the identification of $\mathbf{A}$ and $\mathbf{\Lambda}$ optimum in a statistical sense. The author of this thesis is a co-author of the paper [j9] where an algorithm that works with underdetermined mixtures of nonstationary sources in the optimum way, is proposed.

Bibliography

[1] A. Cichocki and S.-I. Amari, *Adaptive Signal and Image Processing: Learning Algorithms and Applications*, Wiley, New York, 2002.

[2] A. Cichocki, R. Zdunek, A. H. Phan and S. I. Amari, *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-way Data Analysis and Blind Source Separation*, Wiley, 2009.

[3] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, Wiley-Interscience, New York, 2001.

[4] Te-Won Lee, *Independent Component Analysis: Theory and Applications*, MA: Kluwer, Boston, 237 pp., 1998.

[5] P. Comon and C. Jutten, *Handbook of Blind Source Separation: Independent Component Analysis and Applications*, Academic Press, Elsevier Ltd., 859 pp., 2010.

[6] J. Herault and C. Jutten, "Space or time adaptive signal processing by neural network models," *AIP Conf. Proc.*, Vol. 151, pp. 206-211, 1986.

[7] P. Comon, "Independent component analysis: a new concept?", *Signal Processing*, 36(3):287-314, Apr. 1994.

[8] S. Amari, A. Cichocki and H. H. Yang, "A New Learning Algorithm for Blind Signal Separation," *Advances in Neural Information Processing Systems*, 757–763, 1996.

[9] J.-F. Cardoso. "The three easy routes to independent component analysis; contrasts and geometry, " in *Proc. of the ICA 2001 workshop*, San Diego, Dec. 2001.

[10] J.-F. Cardoso and D. T. Pham, Separation of non stationary sources. Algorithms and performance.", in Independent Components Analysis: Principles and Practice, pp. 158 - 180. S. J. Roberts and R. M. Everson (editors), Cambridge University Press, 2001.

[11] J.-F. Cardoso, "Blind signal separation: statistical principles", *Proceedings of the IEEE*, vol. 90, n. 8, pp. 2009-2026, October 1998.

[12] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, and E. Moulines, "A blind source separation technique using second-order statistics," *IEEE Trans. Signal Processing*, vol. 45, pp. 434-444, Feb. 1997.

[13] J.-F. Cardoso and A. Souloumiac, "Blind Beamforming from non-Gaussian Signals", *IEE Proc.-F*, 140(6):362-370, Dec. 1993.

[14] T-W. Lee, M. Girolami and T. J. Sejnowski. "Independent Component Analysis using an Extended Infomax Algorithm for Mixed Sub-Gaussian and Super-Gaussian Sources", *Neural Computation*, Vol.11(2): 417-441, 1999.

[15] A. Hyvärinen and E. Oja, "A Fast Fixed-Point Algorithm for Independent Component Analysis". *Neural Computation*, 9(7):1483-1492, 1997.

[16] A. Hyvärinen, "Fast and Robust Fixed-Point Algorithms for Independent Component Analysis". *IEEE Trans. on Neural Networks*, 10(3):626-634, 1999.

[17] R. Boscolo, H. Pan, and V. P. Roychowdhury, "Independent Component Analysis Based on Nonparametric Density Estimation", *IEEE Trans. on Neural Networks*, vol. 15, no. 1, pp. 55-65, 2004.

[18] D.T. Pham, Garat, P., "Blind separation of mixture of independent sources through a quasi-maximum likelihood approach", *IEEE Trans. on Signal Processing*, Vol. 45, No. 7, July 1997, pp. 1712 - 1725.

[19] X.-L. Li and T. Adali, "A novel entropy estimator and its application to ICA," in Proc. IEEE Workshop on Machine Learning for Signal Processing (MLSP), Grenoble, France, Sep. 2009.

[20] I.F. Gorodnitsky and A. Belouchrani, "Joint Cumulant and Correlation Based Signal Separation with Application to EEG Data", *Proc. of ICA '01*, San Diego, CA, December 2001.

[21] K. R. Müller, P. Philips, and A. Ziehe, "JADE_{TD}: Combining higher-order statistics and temporal information for blind source separation (with noise)", *Proc. of ICA '99*, pp. 87-92, Aussois, January 1999.

[22] S. A. Cruces-Alvarez and A. Cichocki, "Combining blind source extraction with joint approximate diagonalization: thin algorithms for ICA," in *Proc.*

Int. Conf. on Independent Component Analysis and Blind Signal Separation (ICA 2001), San Diego, pp. 740-745, December 2001.

[23] F. Meinecke, A. Ziehe, M. Kawanabe and K. R. Müller, "A resampling approach to estimate the stability of one-dimensional or multidimensional independent components", *IEEE Transactions on Biomedical Engineering*, vol. 49, no. 12, pp. 1514-1525, Dec. 2002.

[24] D-T. Pham, "Blind Separation of Non Stationary Non Gaussian Sources", Proceeding of the EUSIPCO 2002 Conference, Toulouse, France, September 2002.

[25] S. Choi, A. Cichocki, and A. Belouchrani, "Second order nonstationary source separation," *Journal of VLSI Signal Processing*, vol. 32, no. 1-2, pp. 93-104, Aug. 2002.

[26] A. Yeredor, "Blind separation of Gaussian sources via second-order statistics with asymptotically optimal weighting," *IEEE Signal Processing Letters*, vol. 7, pp. 197-200, July 2000.

[27] L. Tong, V. C. Soon, R. Liu, and Y. Huang, "AMUSE: A new blind identification algorithm," *Proc. of ISCAS*, New Orleans, LA, 1990.

[28] P. Tichavský and A. Yeredor, "Fast Approximate Joint Diagonalization Incorporating Weight Matrices," *IEEE Tr. on Signal Processing*, vol. 57, no.3, pp. 878-891, March 2009.

[29] D.-T. Pham, "Joint approximate diagonalization of positive definite Hermitian matrices," *SIAM J. Matrix Anal. and Appl*, 22 (2001), 1136-1152.

[30] B. Afsari, Charleston and J. Rosca, Eds. et al., "Simple LU and QR based non-orthogonal matrix joint diagonalization," in *Proc. 6th Int. Conf. Independent Compon. Analysis and Blind Source Separ.*, Berlin Heidelberg, Germany, 2006, vol. LNCS 3889, pp. 17–24, ICA 2006.

[31] A. Ziehe, P. Laskov, G. Nolte, and K.-R. Müller, "A fast algorithm for joint diagonalization with non-orthogonal transformations and its application to blind source separation," *J. Mach. Learn. Res.*, vol. 5, pp. 777–800, 2004.

[32] D-T. Pham and J-F. Cardoso. "Blind separation of instantaneous mixtures of non stationary sources", *IEEE Trans. Signal Processing*, pp. 1837-1848, vol. 49, no. 9, 2001.

[33] K. Matsuoka, M. Ohya, and M. Kawamoto, "A neural net for blind separation of nonstationary signals," *Neural networks*, 8(3):411–419, 1995.

[34] E. G. Learned-Miller and J. W. Fisher III, "ICA using spacings estimates of entropy", *Journal of Machine Learning Research*, vol. 4, no. 7-8, pp. 1271-1295, 2004.

[35] A. J. Bell and T. J. Sejnowski, "An Information-Maximization Approach to Blind Separation and Blind Deconvolution", *Neural Computation*, 7:1129-1159, 1995.

[36] E. Doron, A. Yeredor and P. Tichavský, "Cramér-Rao Lower Bound for Blind Separation of Stationary Parametric Gaussian Sources", *IEEE Signal Processing Letters*, vol. 14, no.6, pp. 417-420, June 2007.

[37] S. Smekhov and A. Yeredor, "Optimization of JADE Using a Novel Optimally Weighted Joint Diagonalization Approach, " *Proceedings of the 12th European Conference on Signal Processing (EUSIPCO2004)*, Vienna, Sept. 2004.

[38] J.-F. Cardoso, "On the performance of orthogonal source separation algorithms", *Proc. EUSIPCO*, pp. 776-779, Edinburgh, September 1994.

[39] R. C. Rao, *Linear Statistical Inference and Its Applications*, 2nd ed. Wiley, New York, 1973.

[40] J. Cardoso, "Multidimensional independent component analysis," *Proceedings of ICASSP '98*, Seattle, WA, May 12-15, 1998.

[41] L. de Lathauwer, B. de Moor, and J. Vandewalle, "Fetal electrocardiogram extraction by blind source subspace separation", *IEEE Trans. on Biomedical Engineering*, vol. 47, no. 5, pp. 567-572, May 2000.

[42] C. Févotte and F. J. Theis, "Orthonormal approximate joint block-diagonalization," *Technical Report GET/Tlcom Paris 2007D007*, 2007.

[43] Hyvärinen A.: Gaussian Moments for Noisy Independent Component Analysis. *IEEE Signal Processing Letters*, **6** (1999), 145-147.

[44] S. Makino, Te-Won Lee, and H. Sawada, *Blind Speech Separation*, Springer, Sept. 2007.

[45] K. Torkkola, "Blind separation of convolved sources based on information maximization, " *Neural Networks for Signal Processing [1996] VI. Proceedings of the 1996 IEEE Signal Processing Society Workshop*, pp.423-432, 4-6. Sept, 1996.

[46] K. Matsuoka and S. Nakashima, "Minimal distrortion principle for blind source separation," *Proc of ICA-2001*, pp. 722-727, 2001.

[47] P. Smaragdis, "Blind separation of convolved mixtures in the frequency domain, " *Neurocomputing*, Vol 22., pp. 21-31, 1998.

[48] Vincent, E., Gribonval, R., and Févotte, C.: "Performance Measurement in Blind Audio Source Separation", *IEEE Trans. on Speech and Audio Processing*, Vol 14, No 4, pp. 1462 - 1469, July 2006.

[49] D. Schobben, K. Torkkola, and P. Smaragdis, "Evaluation of blind signal separation methods," *Int. Workshop on Independent Component Analysis and Signal Separation (ICA '99)*, Aussois, France, pp. 261-266, Jan. 1999.

[50] Parra, L., and Spence, C.: "Convolutive Blind Separation of Non-Stationary Sources," *IEEE Trans. on Speech and Audio Processing*, Vol. 8, No. 3, pp. 320-327, May 2000.

[51] H. Sawada, R. Mukai, S. Araki, and S. Makino, "A robust and precise method for solving the permutation problem of frequency-domain blind source separation," *IEEE Trans. Speech Audio Processing*, vol. 12, no. 5, pp. 530-538, 2004.

[52] H. Sawada, S. Araki, R. Mukai, and S. Makino, "Grouping separated frequency components by estimating propagation model parameters in frequency-domain blind source separation, " *IEEE Trans. Audio, Speech and Language Processing*, vol. 15, no. 5, pp. 1592-1604, July 2007.

[53] S. Amari, S. C. Douglas, A. Cichocki, and H. H. Yang, "Multichannel Blind Deconvolution and Equalization Using the Natural Gradient," *Proc. IEEE Workshop on Signal Processing Advances in Wireless Communications*, Paris, France, pp. 101-104, April 1997.

[54] S. Douglas, M. Gupta, H. Sawada, and S. Makino, "Spatio-temporal FastICA algorithms for the blind separation of convolutive mixtures, " *IEEE Trans. Audio, Speech and Language Processing*, vol. 15, no. 5, pp. 1511-1520, July 2007.

[55] Xu, X.-F., Feng, D.-Z., Zheng, W.-X., Zhang, H.: Convolutive blind source separation based on joint block Toeplitzation and block-inner diagonalization. *Signal Processing*, vol. 90, no. 1, 119-133 (2010)

[56] Araki, S., Makino, S., Aichner, R., Nishikawa, T., Saruwatari, H.: Subband-based blind separation for convolutive mixtures of speech. IEICE Trans. Fundamentals, vol. E88-A, no. 12, 3593-3603 (2005)

[57] C. Févotte, A. Debiolles, and C. Doncarli. "Blind separation of FIR convolutive mixtures: application to speech signals," *Proc. 1st ISCA Workshop on Non-Linear Speech Processing*, 2003.

[58] F. Nesta, P. Svaizer, M. Omologo, "Convolutive BSS of short mixtures by ICA recursively regularized across frequencies, " (accepted) in *IEEE Transactions on Audio, Speech and Language Processing*, 2010.

[59] M.G. Jafari, E. Vincent, S.A. Abdallah, M.D. Plumbley, and M.E. Davies, "An adaptive stereo basis method for convolutive blind audio source separation, " *Neurocomputing*, vol. 71, pp. 2087-2097, 2008.

[60] F. L. Hitchcock, "The expression of a tensor or a polyadic as a sum of products, " *J. Math. Phys.*, 6 (1), pp. 164-189, 1927.

[61] J. B. Kruskal, "Three-way arrays: Rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics," *Linear Alg. Appl.*, vol. 18, pp. 95-138, 1977.

[62] T. G. Kolda, "Multilinear operators for higher-order decompositions, " technical report SAND2006-2081, April 2006.

[63] N. D. Sidiropoulos and R. Bro, "On the uniqueness of multilinear decomposition of N-way arrays, " *Journal of Chemometrics*, vol. 14, pp. 229-239, 2000.

[64] A. Stegeman, J.M.F. ten Berge, L. De Lathauwer, "Sufficient Conditions for Uniqueness in Candecomp/Parafac and Indscal with Random Component Matrices", *Psychometrika*, Vol. 71, No. 2, pp. 219-229, June 2006.

[65] M. Rajih, P. Comon, and R. A. Harshman, "Enhanced line search: A novel method to accelerate PARAFAC, " *SIAM Journal on Matrix Analysis Appl.*, vol. 30, no. 3, pp. 1148–1171, September 2008.

[66] P. Paatero, "A weighted non-negative least squares algorithm for three-way 'PARAFAC' factor analysis", *Chemometrics and Intelligent Laboratory Systems*, vol. 38, pp. 223-242, 1997.

[67] A.H. Phan, P. Tichavský and A. Cichocki, "Low Complexity Damped Gauss-Newton Algorithms for Parallel Factor Analysis", *submitted for publication*.

[68] L-H. Lim and P. Comon, "Multiarray Signal Processing: Tensor decomposition meets compressed sensing," *Compte-Rendus de l'Academie des Sciences*, section Mecanique, 338(6):311–320, June 2010.

[69] L. De Lathauwer, J. Castaing, "Blind identification of underdetermined mixtures by simultaneous matrix diagonalization", *IEEE Tr. Signal Processing*, Vol. 56, no. 3, pp. 1096 - 1105, March 2008.

[70] A. Ferréol, L. Albera, and P. Chevalier, "Fourth order blind identification of underdetermined mixtures of sources (FOBIUM), " *IEEE Trans. on Signal Processing*, vol. 53, no. 5, pp. 1640-1653, May 2005.

[71] P. Comon, M. Rajih, "Blind identification of under-determined mixtures based on the characteristic function, " *Signal Processing*, vol. 86, no. 9, pp. 2271-2281, Sept. 2006.

[72] A. Taleb, "An algorithm for the blind identification of N independent signals with 2 sensors, " *Proc. 16th Symp. Signal Process. Applicat.*, pp. 5 - 8, 2001.

Author's Publications

Journal papers

[j1] P. Tichavský and Z. Koldovský, "Optimal Pairing of Signal Components Separated by Blind Techniques", *IEEE Signal Processing Letters*, vol. 11, no. 2, pp. 119–122, 2004.

[j2] P. Tichavský, Z. Koldovský, and E. Oja, "Performance Analysis of the FastICA Algorithm and Cramér-Rao Bounds for Linear Independent Component Analysis", *IEEE Trans. on Signal Processing*, vol. 54, no. 4, April 2006.

[j3] Z. Koldovský, P. Tichavský and E. Oja, "Efficient Variant of Algorithm FastICA for Independent Component Analysis Attaining the Cramér-Rao Lower Bound", *IEEE Trans. on Neural Networks*, Vol. 17, No. 5, Sept 2006.

[j4] P. Tichavský, Z. Koldovský, and E. Oja, "Corrections of the 'Performance Analysis of the FastICA Algorithm and Cramér-Rao Bounds for Linear Independent Component Analysis' ", *IEEE Trans. on Signal Processing*, Vol. 56, No.4, pp. 1715-1716, April 2008.

[j5] P. Tichavský, Z. Koldovský, A. Yeredor, G. G. Herrero, and E. Doron, "A Hybrid Technique for Blind Non-Gaussian and Time-Correlated Sources

Using a Multicomponent Approach," *IEEE Trans. on Neural Networks*, Vol. 19, No. 3, pp. 421-430, March 2008.

[j6] Z. Koldovský, J. Málek, P. Tichavský, Y. Deville, and S. Hosseini, "Blind Separation of Piecewise Stationary NonGaussian Sources", *Signal Processing*, Volume 89, Issue 12, Pages 2570-2584, December 2009.

[j7] Z. Koldovský, P. Tichavský, "Time-domain blind separation of audio sources on the basis of a complete ICA decomposition of an observation space, " accepted for publication in *IEEE Trans. on Audio, Language, and Speech Processing*, April 2010.

[j8] P. Tichavský and Z. Koldovský, "Fast and accurate methods of independent component analysis: A Survey, " *submitted*, 2010.

[j9] P. Tichavský and Z. Koldovský, "Weight adjusted tensor method for blind separation of underdetermined mixtures of nonstationary sources, " accepted for publication in *IEEE Trans. on Signal Processing*, 2010.

Papers in international peer-reviewed conference proceedings

[c1] Z. Koldovský, P. Tichavský and E. Oja, "Cramér-Rao lower bound for linear Independent Component Analysis", *Proc. ICASSP 2005*, Philadelphia, vol. III, pp. 581-584, March 2005.

[c2] Z. Koldovský, P. Tichavský, and E. Oja, "Efficient Variant of Algorithm FastICA for Independent Component Analysis Attaining the Cramér-Rao Lower Bound", *Proc. of SSP-2005*, Bordeaux, July 2005.

[c3] P. Tichavský, Z. Koldovský and E. Oja, "Asymptotic Performance Analysis of the Fixed-Point Algorithm (Fast-ICA) for Independent Component Analysis", *Proc. of SSP-2005*, Bordeaux, July 2005.

[c4] P. Tichavský, Z. Koldovský, E. Doron, A. Yeredor, and G. G. Herrero, "Blind signal separation by combining two ICA algorithms: HOS-based EFICA and time structure-based WASOBI", *Proceedings of The 2006 European Signal Processing Conference (EUSIPCO'2006)*, Florence, Sep. 2006.

[c5] Z. Koldovský, J. Nouza, and J. Kolorenč, "Continuous Time-Frequency Masking Method for Blind Speech Separation with Adaptive Choice of

Threshold Parameter Using ICA", *Interspeech 2006*, Pittsburgh PA, USA, 17.-21. September, pp. 2578-2581, 2006.

[c6] Z. Koldovský, P. Tichavský, "Methods of Fair Comparison of Performance of Linear ICA Techniques in Presence of Additive Noise", *Proc. of ICASSP 2006*, Toulouse, no. V., pp. 873-876, May 2006.

[c7] Z. Koldovský and P. Tichavský, "Time-Domain Blind Audio Source Separation Using Advanced ICA Methods", *Proceedings of the 8th Annual Conference of the International Speech Communication Association (Interspeech 2007)*, pp. 846-849, August 2007.

[c8] P. Tichavský, Z. Koldovský, and E. Oja, "Speed and Accuracy Enhancement of Linear ICA Techniques Using Rational Nonlinear Functions", *Proc. of ICA2007*, pp. 285-292, Sept. 2007.

[c9] Z. Koldovský and P. Tichavský, "Blind Instantaneous Noisy Mixture Separation with Best Interference-plus-noise Rejection," *Proceedings of 7th International Conference on Independent Component Analysis (ICA2007)*, pp. 730-737, Sept. 2007.

[c10] Herrero, G.G., Koldovský, Z., Tichavský, P., Egiazarian, K.: A Fast Algorithm for Blind Separation of Non-Gaussian and Time-Correlated Signals. In: Proceedings of 15th European Signal Processing Conference (EUSIPCO 2007), September 2007, pp. 1731–1735 (2007)

[c11] J. Málek, Z. Koldovský, S. Hosseini, and Y. Deville, "A Variant of EFICA Algorithm with Adaptive Parametric Density Estimator", *8th International Workshop on Electronics, Control, Modelling, Measurement, and Signals (ECMS 2007)*, pp. 79-84, Liberec, Czech Republic, May 2007.

[c12] Z. Koldovský, J. Málek, P. Tichavský, Y. Deville, and S. Hosseini, "Extension of EFICA Algorithm for Blind Separation of Piecewise Stationary non Gaussian Sources", in *Proc. IEEE Int. Conf. Acoust., Speech and Signal Processing*, Las Vegas, Nevada, USA, pp. 1913-1916, April 2008.

[c13] Z. Koldovský, P. Tichavský, "Time-domain blind audio source separation using advanced component clustering and reconstruction," *Proc. of HSCMA 2008*, Trento, Italy, pp. 216-219, May 2008.

[c14] J. Málek, Z. Koldovský, J. Žďánský and J. Nouza, "Enhancement of Noisy Speech Recordings via Blind Source Separation", *Proceedings of the 9th*

Annual Conference of the International Speech Communication Association (Interspeech 2008), pp. 159-162, ISSN: 1990-9772, September 22-26, Brisbane, Australia, 2008.

[c15] Z. Koldovský and P. Tichavský, "A Comparison of Independent Component and Independent Subspace Analysis Algorithms, " *Proc. of EUSIPCO 2009*, pp. 1447-1451, Glasgow, August 24-28, 2009.

[c16] P. Tichavský, A. Yeredor, and Z. Koldovský, "A Fast Asymptotically Efficient Algorithm for Blind Separation of a Linear Mixture of Block-Wise Stationary Autoregressive Processes," ICASSP 2009, pp. 3133-3136, Taipei, Taiwan, April 2009.

[c17] J. Petkov and Z. Koldovský, "BSSGUI – A Package for Interactive Control of Blind Source Separation Algorithms in MATLAB," in *Cross-Modal Analysis of Speech, Gestures, Gaze and Facial Expressions* (Eds.: A. Esposito and R. Vích), pp. 386-398, Springer Berlin / Heidelberg, July 2009.

[c18] Z. Koldovský, P. Tichavský, and J. Málek, "Time-Domain Blind Audio Source Separation Method Producing Separating Filters of Generalized Feedforward Structure," in *Latent Variable Analysis and Signal Separation*, Lecture Notes in Computer Science, Vol. 6365, pp. 17-24, Springer, Heidelberg, Sept. 2010.

[c19] Z. Koldovský, P. Tichavský, and J. Málek, "Subband Blind Audio Source Separation Using a Time-Domain Algorithm and Tree-Structured QMF Filter Bank," in *Latent Variable Analysis and Signal Separation*, Lecture Notes in Computer Science, Vol. 6365, pp. 25-32, Springer, Heidelberg, Sept. 2010.

[c20] J. Málek, Z. Koldovský and P. Tichavský, "Adaptive Time-Domain Blind Separation of Speech Signals," in Latent Variable Analysis and Signal Separation, *Lecture Notes in Computer Science*, Vol. 6365, pp. 9-16, Springer, Heidelberg, Sept. 2010.

[c21] P. Tichavský and Z. Koldovský, "Simultaneous Search for All Modes in Multilinear Models," *ICASSP 2010*, pp. 4114-4117, Dallas, USA, March 2010.

[c22] P. Tichavský and Z. Koldovský, "Stability of CANDECOMP-PARAFAC tensor decomposition," *submitted to a conference*, October 2010.

Miscellaneous

[m1] Z. Koldovský, "Analýza nezávislých komponent v EEG datech", Master thesis, Faculty of Nuclear Sciences and Physical Engineering, CTU in Prague, 2002.

[m2] Z. Koldovský, P. Tichavský, "Cramér-Rao Lower Bound for Independent Component Analysis", technical report nr. 2123, ÚTIA, AV ČR, 2004.

[m3] P. Tichavský, J. Nielsen, Z. Koldovský, "Odstraňování artefaktů v EEG datech I", technical report nr. 2140, ÚTIA, AV ČR, 2005.

[m4] Z. Koldovský, "Fast and Accurate Methods for Independent Component Analysis", Ph.D. thesis, Faculty of Nuclear Sciences and Physical Engineering, CTU in Prague, 2006.

[m5] J. Nielsen, P. Tichavský, Z. Koldovský, "Odstraňování artefaktů v EEG datech II", technical report nr. 2176, ÚTIA, AV ČR, 2006.

[m6] Z. Koldovský and P. Tichavský, "Asymptotic analysis of bias of FastICA-based algorithms in presence of additive noise," *Technical report no. 2181*, ÚTIA, AV ČR, 2007.

[m7] Z. Koldovský, J. Málek, P. Tichavský, Y. Deville, and S. Hosseini, "Performance Analysis of Extended EFICA Algorithm", *Technical report no. 2199*, ÚTIA, AV ČR, 2007.

[m8] Z. Koldovský and P. Tichavský, "Způsob slepé separace akustických signálů z jejich konvolutorní směsi," Czech Patent Application PV 2008-752, Nov. 2008.

[m9] O. Hnilička, J. Málek, K. Paleček, Z. Koldovský, "A Fast C++ Implementation of Time-domain Blind Speech Separation Algorithm, " *Proc. of the 20th Czech-German WorkShop on Speech Processing*, Oct. 2010.

Appendix: The Main Author's Results

The following table provides a survey of the main results of the author achieved in 2002-2010 that were already published or accepted for publication. This especially comprises algorithms, codes, theoretical and experimental analyses, and applications. The entries are ordered chronologically and contain all relevant references.

Name	References	Online code*	Year of the first publication
Algorithms			
Optimum pairing	[m1, j1]	✓	2002
EFICA	[j3, c2]	✓	2005
COMBI	[c4]	✓	2006
MULTICOMBI	[j5]	✓	2008
FCOMBI	[c10]	✓	2007
1FICA	[c9]	✓	2007
Block EFICA	[c12, j6]	✓	2008
BARBI	[c16]	✓	2009
Theoretical Results			
FastICA performance	[c3, j2]		2005
Cramér-Rao bound (2.28)	[c1]		2005
Bias analysis	[c9]		2007
Rational nonlinearities	[c8]		2007
Cramér-Rao bound (2.24)	[c12]		2008
Cramér-Rao bound (2.30)	[c16]		2009
Pseudoconvolutive mixtures	[c15]		2009
Audio Applications			
Time-frequency masking	[c5]		2006
T-ABCD	[c7, c13]	✓	2007
Generalized T-ABCD	[c18, j7]	✓	2010
Subband T-ABCD	[c19]	✓	2010
On-line T-ABCD	[c20]	✓	2010

* The codes of algorithms are available at `http://itakura.ite.tul.cz/zbynek/downloads.htm`.

Uchazeč: Ing. Zbyněk Koldovský, Ph.D. Pracoviště: ITE, FM, TUL

Podpis:

Hodnocené období: 2006-2010

A1. Vědecko výzkumná činnost

Základní výzkum (hodnocený především na základě publikací nových poznatků) (uveďte nejvýznamnější výsledky, vědeckovýzkumný přínos, slovní charakterizaci)
1. výsledek (projekt) Postdoktorský projekt GAČR 102/07/P384, „Použití pokročilých metod pro analýzu nezávislých komponent na slepou separaci reálných signálů“, 2007-2009.
Charakterizace (V-V přínos, osobní podíl, …) Hlavním výstupem projektu jsou publikace [1-4] a [12-22], z nichž všechny s výjimkou článků na konferencích EUSIPCO a [22] jsou uvedené v databázi ISI, kde mají 11 citací (ke konci roku 2010). Výsledkem projektu jsou teoretické analýzy, 6 nových ICA algoritmů (COMBI [24], MULTICOMBI a FCOMBI [4, 21], 1FICA [20], Block EFICA [2,17], BARBI [13]), jejich varianty a metoda T-ABCD [1,15,16,18] a její varianty.
2. výsledek (publikace) Z. Koldovský, P. Tichavský and E. Oja, "Efficient Variant of Algorithm FastICA for Independent Component Analysis Attaining the Cramér-Rao Lower Bound", *IEEE Trans. on Neural Networks*, Vol. 17, No. 5, Sept 2006.
Článek popisuje nový algoritmus pro analýzu nezávislých komponent, který je efektnější a přesnější variantou známé metody FastICA (jeho přesnost dosahuje Rao-Cramérovi meze pro daný model signálů). Algoritmus se jmenuje EFICA a jeho m-kód je zveřejněný na internetu. Algoritmus byl použit v mnoha jiných pracích (Google Scholar eviduje již něco kolem 85 citací tohoto článku) a byl také zahrnut do ICALAB toolboxu profesora Andreje Cichockého z RIKEN institutu v Japonsku, což svědčí o jeho kvalitách. Článek byl oceněn jako nejlepší publikace roku 2006 v ÚTIA AV ČR v.v.i. *Osobní podíl na této publikaci 60%*
3. výsledek (publikace) P. Tichavský and Z. Koldovský, “Weight adjusted tensor method for blind separation of underdetermined mixtures of nonstationary sources, ” accepted for publication in *IEEE Trans. on Signal Processing*, Nov. 2010.
Algoritmus navržený v tomto článku slouží k separaci nestacionárních signálů z jejich směsi, je-li měřených signálů méně než signálů obsažených ve směsi. Algoritmus je přesnější než konkurenční metody. Jeho přesnost dosahuje příslušnou Rao-Cramérovu mez, která je též v článku odvozena. Algoritmus je založen na dekompozici tenzoru, který vznikne poskládáním kovariančních matic z bloků měřených signálů. Jelikož jedna z faktorizačních matic tenzoru je nezáporná, může v některých případech algoritmus přesáhnout Rao-Cramérovu mez, která ze své podstaty nemůže brát tuto informaci v potaz. Článek je jedním z nejnovějších výsledků uchazeče. *Osobní podíl na této publikaci 40%.*

Aplikovaný výzkum (hodnocený na základě realizací nových technologií, konstrukcí, apod.) (uveďte nejvýznamnější výsledky, vědeckovýzkumný přínos, slovní charakterizaci)
1. výsledek (realizace) *Algoritmus T-ABCD, jeho varianty a implementace*: Z. Koldovský and P. Tichavský, "Time-Domain Blind Separation of Audio Sources on the basis of a Complete ICA Decomposition of an Observation Space", accepted for publication in *IEEE Trans. on Speech, Audio and Language Processing*, April 2010.
Charakterizace (V-V přínos, uplatnění, patent, osobní podíl, ...) : Prototyp metody T-ABCD pro slepou separaci akustických signálů byl navržen v článku [18] a jeho vylepšené varianty postupně v [16] a [15]. Nejnovější varianty včetně varianty pro on-line zpracování v reálném čase byly publikovány v [1] a [8-10]. Algoritmus by implementován v Matlabu a bylo vytvořeno grafické uživatelské rozhraní pro jeho ovládání. K jedné z jeho upravených variant byla podána česká patentová přihláška č. PV 2008-752 (dosud nebyla vyhodnocena). Jeho on-line varianta byla implementována v C++.
2. výsledek (projekt) Projekt pro Škoda Auto a.s. „Řešení úlohy snímání a objektivního vyhodnocování hluku předních stěračů“.
Charakterizace (V-V přínos, uplatnění, patent, osobní podíl, ...) : Projekt byl řešen v roce 2008 a na základě požadavků zadavatele byly provedeny další úpravy v roce 2010. Výstupem bylo levné zařízení a software specializované na měření hluku stěračů v automobilu. Bylo navrženo několik kritérií pro vyhodnocení hluku.
3. výsledek (projekt, realizace)
Charakterizace (V-V přínos, uplatnění, patent, osobní podíl, ...) :

A2. Pedagogická a vzdělávací činnost

Přednášková činnost
garance a vedení přednášek
1) Biologické a akustické signály (BSI) – garant předmětu, zavedení předmětu, příprava a vedení přednášek i cvičení; pro pátý ročník v posledním semestru magisterského studia na FM (10 týdnů), odučeny 2 semestry: LS 2009, LS 2010.
2) Výpočty, simulace a vizualizace Matlab (MATLB) – garant předmětu, zavedení předmětu, příprava a vedení přednášek i cvičení; pro druhý semestr bakalářského studia na FM (14 týdnů), odučeny 2 semestry: LS 2009, LS 2010.
3) Počítačové zpracování signálů (PZS) – příprava a vedení poloviny přednášek a cvičení; pro dvouleté navazující magisterské studium na FM, odučeny 2 semestry: LS 2007, LS 2008.
Podrobný seznam a popis všech vyučovaných předmětů (pouze prezenční studium) je uveden v příloze „Přehled pedagogické praxe“

Učebnice a výukové pomůcky
charakteristika učebnice, výukové pomůcky
1) Odborná kniha z oblasti zpracování a rozpoznávání řeči: Nouza, J. a kol.: Řeč a počítač: principy hlasové komunikace, úlohy, metody a aplikace, vydavatel: Technická univerzita v Liberci, editoři: Jan Nouza, Zbyněk Koldovský, Robert Vích, číslo publikace: 55-110-09, první vydání, počet stran: 238, ISBN 978-80-7372-548-8, 2009. kapitola – Zpracování záznamu řeči z více mikrofonů metodami pro řešení inverzního problému“, str. 201-216 (80%)

Individuální vzdělávací činnost
vedení projektu, diplomové práce, doktoranda (kvantitativní i kvalitativní hodnocení)
1) Vedení 7 studentských projektů v bakalářských a magisterských studijních programech, v prezenčním studiu. **)*
2) Vedení 3 bakalářských prací a 3 diplomových prací. **)*
3) Vedení jednoho PhD studenta: Ing. Jiří Málek, zahájení studia 1.8.2006, studijní obor: Technická kybernetika, disertační práce má název *Blind Audio Source Separation via Independent Component Analysis*, odevzdáno v listopadu-prosinci 2010.
**) Poznámka: seznam jednotlivých projektů, bakalářských a diplomových prací a studenta PGS je uveden v příloze „Přehled pedagogické praxe" a včetně kvantitativního vyjádření v příloze „Kvantifikovaná kritéria pro habilitační a jmenovací řízení na Fakultě mechatroniky TU v Liberci"*

Podíl na garantování Bc., Mgr. a PhD oboru Přínos k profilu absolventa
Přípravou podkladů pro akreditaci jsem se podílel na úspěšných akreditacích a re-akreditacích studijních oborů na FMIMS a ÚZS. Jsem garantem jednoho předmětu doktorského studia: Metody digitálního zpracování vícerozměrných signálů. Podílím se na přípravě a výuce mezinárodního česko-německého magisterského studijního programu ELECTRICAL ENGINEERING AND INFORMATICS, anglická výuka připravovaného předmětu DSP1 – Digital Signals Processing 1 (od letního semestru 2011)

A3. Ostatní významné aktivity

Jiné aktivity
1) Recenzent článků pro mezinárodní časopisy IEEE (Trans. Signal Processing, Signal Processing Letters, Trans. Neural Networks, Trans. Audio Speech and Language Processing, a další), Technometrics, Signal Processing, a další a ve sbornících mezinárodních konferencí (LVA/ICA 2010, ICASSP 2011).

Publikační aktivity:

Publikace v mezinárodních impaktovaných časopisech

[1] P. Tichavský and Z. Koldovský, "Fast and accurate methods of independent component analysis: A Survey, " accepted for publication in *Kybernetika*, 2010.

[2] P. Tichavský and Z. Koldovský, "Weight Adjusted Tensor Method for Blind Separation of Underdetermined Mixtures of Nonstationary Sources, " *IEEE Trans. on Signal Processing*, Vol. 59, No. 3, pp. 1037-1047, ISSN:1053-587X, March 2011.

[3] Z. Koldovský and P. Tichavský, "Time-Domain Blind Separation of Audio Sources on the basis of a Complete ICA Decomposition of an Observation Space", *IEEE Trans. on Speech, Audio and Language Processing*, Vol. 19, No. 2, pp. 406-416, ISSN 1558-7916, February 2011.

[4] Z. Koldovský, J. Málek, P. Tichavský, Y. Deville, and S. Hosseini, "Blind Separation of Piecewise Stationary NonGaussian Sources", *Signal Processing*, Volume 89, Issue 12, Pages 2570-2584, ISSN 0165-1684, December 2009.

[5] P. Tichavský, Z. Koldovský, and E. Oja, "Corrections to 'Performance Analysis of the FastICA Algorithm and Cramer-Rao Bounds for Linear Independent Component Analysis' TSP 04/06," *IEEE Tr. Signal Processing*, Vol. 56, No.4, pp. 1715-1716, ISSN: 1053-587X, April 2008.

[6] P. Tichavský, Z. Koldovský, A. Yeredor, G. G. Herrero, and E. Doron, "A Hybrid Technique for Blind Separation of Non-Gaussian and Time-Correlated Sources Using a Multicomponent Approach", *IEEE Trans. on Neural Networks*, Vol. 19, No. 3, pp. 421-430, ISSN: 1045-9227, March 2008.

[7] Z. Koldovský, P. Tichavský and E. Oja, "Efficient Variant of Algorithm FastICA for Independent Component Analysis Attaining the Cramér-Rao Lower Bound", *IEEE Trans. on Neural Networks*, Vol. 17, No. 5, Sept 2006.

[8] P. Tichavský, Z. Koldovský and E. Oja, "Performance Analysis of the FastICA Algorithm and Cramér-Rao Bounds for Linear Independent Component Analysis", *IEEE Trans. on Signal Processing*, Vol. 54, No.4, April 2006.

[9] P. Tichavský, Z. Koldovský, "Optimal Pairing of Signal Components Separated by Blind Techniques", *IEEE Signal Processing Letters*, Vol. 11, No. 2, pp. 119-122, Feb 2004.

Publikace v recenzovaných sbornících mezinárodních konferencí

[10] Z. Koldovský, P. Tichavský, and Anh Huy Phan, "Stability Analysis and Fast Damped-Gauss-Newton Algorithm for INDSCAL Tensor Decomposition," *IEEE Workshop on Statistical Signal Processing*, Nice, France, June 2011.

[11] P. Tichavský and Z. Koldovský, "Stability of CANDECOMP-PARAFAC tensor decomposition," *ICASSP 2011*, Prague, Czech Republic, May 2011.

[12] Z. Koldovský, P. Tichavský, and J. Málek, "Time-Domain Blind Audio Source Separation Method Producing Separating Filters of Generalized Feedforward Structure," in *Latent Variable Analysis and Signal Separation*, Lecture Notes in Computer Science Vol. 6365, pp. 17-24, ISBN: 978-3-642-15994-7, Springer, Heidelberg, Sept. 2010.

[13] Z. Koldovský, P. Tichavský, and J. Málek, "Subband Blind Audio Source Separation Using a Time-Domain Algorithm and Tree-Structured QMF Filter Bank," in *Latent Variable Analysis and Signal Separation*, Lecture Notes in Computer Science Vol. 6365, pp. 25-32, ISBN: 978-3-642-15994-7, Springer, Heidelberg, Sept. 2010.

[14] J. Málek, Z. Koldovský, and P. Tichavský, "Adaptive Time-Domain Blind Separation of Speech Signals," in *Latent Variable Analysis and Signal Separation*, Lecture Notes in Computer Science Vol. 6365, pp. 9-16, ISBN: 978-3-642-15994-7, Springer, Heidelberg, Sept. 2010.

[15] P. Tichavský and Z. Koldovský, "Simultaneous Search for All Modes in Multilinear Models," *ICASSP 2010*, pp. 4114-4117, ISBN: 978-1-4244-4296-6, ISSN: 1520-6149, Dallas, USA, March 2010.

[16] Z. Koldovský and P. Tichavský, "A Comparison of Independent Component and Independent Subspace Analysis Algorithms," *EUSIPCO 2009* , pp. 1447-1451, Glasgow, Scotland, August 24-28, 2009.

[17] P. Tichavský, A. Yeredor, and Z. Koldovský, "A Fast Asymptotically Efficient Algorithm for Blind Separation of a Linear Mixture of Block-Wise Stationary Autoregressive Processes," *ICASSP 2009*, pp. 3133-3136, ISBN: 978-1-4244-2354-5, ISSN: 1520-6149, Taipei, Taiwan, April 2009.

[18] J. Petkov and Z. Koldovský, "BSSGUI – A Package for Interactive Control of Blind Source Separation Algorithms in MATLAB," in *Cross-Modal Analysis of Speech, Gestures, Gaze and Facial Expressions* (Eds.: A. Esposito and R. Vích), pp. 386-398, ISBN: 978-3-642-03319-3, ISSN: 0302-9743, Springer Berlin / Heidelberg, July 2009.

[19] J. Málek, Z. Koldovský, J. Žďánský and J. Nouza, "Enhancement of Noisy Speech Recordings via Blind Source Separation", *Proceedings of the 9th Annual Conference of the International Speech Communication Association (Interspeech 2008)*, pp. 159-162, ISSN: 1990-9772, September 22-26, Brisbane, Australia, 2008.

[20] Z. Koldovský and P. Tichavský, "Time-domain Blind Audio Source Separation Using Advanced Component Clustering and Reconstruction", *Proc. of The Joint Workshop on Hands-free Speech Communication and Microphone Arrays (HSCMA 2008)*, pp. 216-219, ISBN: 978-1-4244-2338-5, May 6-8, Trento, Italy, 2008.

[21] Z. Koldovský, J. Málek, P. Tichavský, Y. Deville, and S. Hosseini, "Extension of EFICA Algorithm for Blind Separation of Piecewise Stationary Non Gaussian Sources", *33rd International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2008)*, Las Vegas, Nevada, pp. 1913-1916, ISBN: 1-4244-1484-9, April 2008.

[22] Z. Koldovský and P. Tichavský, "Time-Domain Blind Audio Source Separation Using Advanced ICA Methods", *Proceedings of the 8th Annual Conference of the International Speech Communication Association (Interspeech 2007)*, pp. 846-849, August 2007.

[23] P. Tichavský, Z. Koldovský and E. Oja, "Speed and Accuracy Enhancement of Linear ICA Techniques Using Rational Nonlinear Functions", *Proceedings of 7th International Conference on Independent Component Analysis (ICA2007)*, pp. 285-292, Sept. 2007.

[24] Z. Koldovský and P. Tichavský, "Blind Instantaneous Noisy Mixture Separation with Best Interference-plus-noise Rejection", *Proceedings of 7th International Conference on Independent Component Analysis (ICA2007)*, pp. 730-737, Sept. 2007.

[25] G. G. Herrero, Z. Koldovský, P. Tichavský, and K. Egiazarian, "A Fast Algorithm for Blind Separation of Non-Gaussian and Time-Correlated Signals", *Proceedings of 15th European Signal Processing Conference (EUSIPCO 2007)*, pp. 1731-1735, Sept 2007.

[26] J. Málek, Z. Koldovský, S. Hosseini, and Y. Deville, "A Variant of EFICA Algorithm with Adaptive Parametric Density Estimator", *8th International Workshop on Electronics, Control, Modelling, Measurement, and Signals (ECMS 2007)*, pp. 79-84, Liberec, Czech Republic, May 2007.

[27] Z. Koldovský, J. Nouza, and J. Kolorenč, "Continuous Time-Frequency Masking Method for Blind Speech Separation with Adaptive Choice of Threshold Parameter Using ICA", *Interspeech 2006*, Pittsburgh PA, USA, 17.-21. September, pp. 2578-2581, 2006.

[28] P. Tichavský, Z. Koldovský, E. Doron, A. Yeredor, and G. G. Herrero, "Blind signal separation by combining two ICA algorithms: HOS-based EFICA and time structure-based WASOBI", *Proceedings of The 2006 European Signal Processing Conference (EUSIPCO'2006)*, Florence, Sep. 2006.

[29] Z. Koldovský, P. Tichavský, "Methods of Fair Comparison of Performance of Linear ICA Techniques in Presence of Additive Noise", *Proc. of ICASSP 2006*, Toulouse, no. V., pp. 873-876, May 2006.

[30] Z. Koldovský, P. Tichavský, "Efficient Variant of Algorithm FastICA for Independent Component Analysis Attaining the Cramér-Rao Lower Bound", *Proc. of SSP-2005*, Bordeaux, July 2005.

[31] P. Tichavský, Z. Koldovský and E. Oja, "Asymptotic Performance Analysis of the Fixed-Point Algorithm (Fast-ICA) for Independent Component Analysis", *Proc. of SSP-2005*, Bordeaux, July 2005.

[32] Z. Koldovský, P. Tichavský and E. Oja, "Cramér-Rao Lower Bound for Linear Independent Component Analysis., *Proc. of ICASSP-2005*, Philadelphia, Vol. III, pp. 581 - 584, March 2005.

Uchazeč: Ing. Zbyněk Koldovský, Ph.D. Pracoviště: ITE

Podpis:

Hodnocené období: 2000-2010

B1. Vědecko-výzkumná činnost

Typ aktivity	kvantif. koef.	počet	body
monografie (výstava arch. prací) zahranič.B4	10 – 20		
samostat. kapitola v mezinár. monografii	3 – 6		
monografie (výstava arch. prací) ČR	7 – 12		
samostat. kapitola v české monografii	2 – 4		
článek v mezinár. recenzovaném časopise	3 – 10	9	37
příspěvek na mezinár. konf. (ve sborníku)	2 – 4	23	46
průmyslem potvrzené použití – zahr.	5 – 10		
realiz. dílo většího rozsahu – zahr.	5 – 10		
průmyslem potvrzené použití - ČR	2 – 5		
realiz. dílo většího rozsahu – ČR	2 – 5		
citace zahraniční	6	68	204
citace ČR	2	1	2
uvedení realiz. díla v odb. publ. – zahr.	6		
uvedení realiz. díla v odb. publ. - ČR	2		
úspěšný projekt uplat. v mezinár. soutěži	10 – 15		
úspěšný projekt uplat. v české soutěži	7 – 12		
původní článek v českém věd. a odb. čas.	3		
odborný článek referativní nebo populariza.	1		
pův. příspěvek na čes. konf. (ve sborníku)	2		
oponovaná výzkumná zpráva	1 – 3		
ochranný dokument (vyn., pat., lic.)	5		
udělený grant zahraniční	5 – 10		
udělený grant ČR	2 – 4	5	12
spoluřeš. zahraničního grantu	3 – 6		
spoluřeš. grantu ČR	1 – 2		
realizované dílo	1 – 7	1	5
jiná aktivita	1 – 4		
Vědecko-výzkumná činnost celkem		107	306

B2. Pedagogická a vzdělávací činnost

Typ aktivity	kvantif. koef.	počet	body
přednášení v řád. studiu min. 2 hod/týd.	2/sem.	7	14
pravidelná cv. min. 2 hod/týd.	0,5/sem.	31	15,5
vedení studentského projektu	0,5/sem.	7	5
zavedení nového předmětu v řád. studiu	4	5	16
Vysokoškolská učebnice	6 – 10		
Vysokoškolská skripta předn./cvičení	6/4		
středoškolská učebnice	2 – 5		
jiná knižní publikace, didakt. pomůcka	3 – 6	3	6
překlad učebnice	3		
výukový film, video, výukový software	3		
úspěšný doktorand nebo aspirant	8		
vedení doktoranda před/po dokt. zkoušce	2/4	1	4
stážista se závěrečnou prací	1		
vedení oceněné studentské práce	1		
vedení obhájené diplomové práce	1/dipl.	3	3
garant doktor. PGS, člen obor. rady PGS	3	1	3
vedoucí výzkumného nebo projekt. týmu	4		
vedoucí katedry, ředitel ústavu	3		
Pedagogická a vzdělávací činnost celkem		58	66,5
		Součet B1 + B2	**372,5**

Doporučený limit pro zahájení habilitačního řízení: 60 – 70 bodů / 5 let.
Doporučený limit pro zahájení jmenovacího řízení: 150 – 170 bodů / 10 let.

Zdůvodnění bodového hodnocení

B1. Vědecko-výzkumná činnost

článek v mezinár. recenzovaném časopise

P. Tichavský and Z. Koldovský, "Fast and accurate methods of independent component analysis: A Survey, " accepted for publication in *Kybernetika*, 2010. (bodové ohodnocení publikace: 6 body, podíl 40%, body: 2)

P. Tichavský and Z. Koldovský, "Weight Adjusted Tensor Method for Blind Separation of Underdetermined Mixtures of Nonstationary Sources, " *IEEE Trans. on Signal Processing*, Vol. 59, No. 3, pp. 1037-1047, ISSN:1053-587X, March 2011. (bodové ohodnocení publikace: 10 body, podíl 40%, body: 4)

Z. Koldovský and P. Tichavský, "Time-Domain Blind Separation of Audio Sources on the basis of a Complete ICA Decomposition of an Observation Space", *IEEE Trans. on Speech, Audio and Language Processing*, Vol. 19, No. 2, pp. 406-416, ISSN 1558-7916, February 2011. (bodové ohodnocení publikace: 10 body, podíl 65%, body: 6,5)

Z. Koldovský, J. Málek, P. Tichavský, Y. Deville, and S. Hosseini, "Blind Separation of Piecewise Stationary NonGaussian Sources", *Signal Processing*, Volume 89, Issue 12, Pages 2570-2584, ISSN 0165-1684, December 2009. (bodové ohodnocení publikace: 10 body, podíl 60%, body: 6)

P. Tichavský, Z. Koldovský, and E. Oja, "Corrections to 'Performance Analysis of the FastICA Algorithm and Cramer-Rao Bounds for Linear Independent Component Analysis' TSP 04/06," *IEEE Tr. Signal Processing*, Vol. 56, No.4, pp. 1715-1716, ISSN: 1053-587X, April 2008. (bodové ohodnocení publikace: 5 body, podíl 40%, body: 2)

P. Tichavský, Z. Koldovský, A. Yeredor, G. G. Herrero, and E. Doron, "A Hybrid Technique for Blind Separation of Non-Gaussian and Time-Correlated Sources Using a Multicomponent Approach", *IEEE Trans. on Neural Networks*, Vol. 19, No. 3, pp. 421-430, ISSN: 1045-9227, March 2008. (bodové ohodnocení publikace: 10 body, podíl 30%, body: 3)

Z. Koldovský, P. Tichavský and E. Oja, "Efficient Variant of Algorithm FastICA for Independent Component Analysis Attaining the Cramér-Rao Lower Bound", *IEEE Trans. on Neural Networks*, Vol. 17, No. 5, Sept 2006. (bodové ohodnocení publikace: 10 body, podíl 60%, body: 6)

P. Tichavský, Z. Koldovský and E. Oja, "Performance Analysis of the FastICA Algorithm and Cramér-Rao Bounds for Linear Independent Component Analysis", *IEEE Trans. on Signal Processing*, Vol. 54, No.4, April 2006. (bodové ohodnocení publikace: 10 body, podíl 35%, body: 3,5)

P. Tichavský, Z. Koldovský, "Optimal Pairing of Signal Components Separated by Blind Techniques", *IEEE Signal Processing Letters*, Vol. 11, No. 2, pp. 119-122, Feb 2004. (bodové ohodnocení publikace: 10 body, podíl 40%, body: 4)

Příspěvek na mezinárodní konferenci (ve sborníku)

Z. Koldovský, P. Tichavský, and Anh Huy Phan, "Stability Analysis and Fast Damped-Gauss-Newton Algorithm for INDSCAL Tensor Decomposition," *IEEE Workshop on Statistical Signal Processing*, Nice, France, June 2011. (bodové ohodnocení publikace: 4 body, podíl 60%, body: 2)

P. Tichavský and Z. Koldovský, "Stability of CANDECOMP-PARAFAC tensor decomposition," *ICASSP 2011*, Prague, Czech Republic, May 2011. (bodové ohodnocení publikace: 4 body, podíl 40%, body: 2)

Z. Koldovský, P. Tichavský, and J. Málek, "Time-Domain Blind Audio Source Separation Method Producing Separating Filters of Generalized Feedforward Structure," in *Latent Variable Analysis and Signal Separation*, Lecture Notes in Computer Science Vol. 6365, pp. 17-24, ISBN: 978-3-642-15994-7, Springer, Heidelberg, Sept. 2010. (bodové ohodnocení publikace: 4 body, podíl 60%, body: 3)

Z. Koldovský, P. Tichavský, and J. Málek, "Subband Blind Audio Source Separation Using a Time-Domain Algorithm and Tree-Structured QMF Filter Bank," in *Latent Variable Analysis and Signal Separation*, Lecture Notes in Computer Science Vol. 6365, pp. 25-32, ISBN: 978-3-642-15994-7, Springer, Heidelberg, Sept. 2010. (bodové ohodnocení publikace: 4 body, podíl 60%, body: 3)

J. Málek, Z. Koldovský, and P. Tichavský, "Adaptive Time-Domain Blind Separation of Speech Signals," in *Latent Variable Analysis and Signal Separation*, Lecture Notes in Computer Science Vol. 6365, pp. 9-16, ISBN: 978-3-642-15994-7, Springer, Heidelberg, Sept. 2010. (bodové ohodnocení publikace: 4 body, podíl 30%, body: 1)

P. Tichavský and Z. Koldovský, "Simultaneous Search for All Modes in Multilinear Models," *ICASSP 2010*, pp. 4114-4117, ISBN: 978-1-4244-4296-6, ISSN: 1520-6149, Dallas, USA, March 2010. (bodové ohodnocení publikace: 4 body, podíl 30%, body: 1)

Z. Koldovský and P. Tichavský, "A Comparison of Independent Component and Independent Subspace Analysis Algorithms," *EUSIPCO 2009* , pp. 1447-1451, Glasgow, Scotland, August 24-28, 2009. (bodové ohodnocení publikace: 4 body, podíl 60%, body: 3)

P. Tichavský, A. Yeredor, and Z. Koldovský, "A Fast Asymptotically Efficient Algorithm for Blind Separation of a Linear Mixture of Block-Wise Stationary Autoregressive Processes," *ICASSP 2009*, pp. 3133-3136, ISBN: 978-1-4244-2354-5, ISSN: 1520-6149, Taipei, Taiwan, April 2009. (bodové ohodnocení publikace: 4 body, podíl 15%, body: 1)

J. Petkov and Z. Koldovský, "BSSGUI – A Package for Interactive Control of Blind Source Separation Algorithms in MATLAB," in *Cross-Modal Analysis of Speech, Gestures, Gaze and Facial Expressions* (Eds.: A. Esposito and R. Vích), pp. 386-398, ISBN: 978-3-642-03319-3, ISSN: 0302-9743, Springer Berlin / Heidelberg, July 2009. (bodové ohodnocení publikace: 4 body, podíl 50%, body: 2)

J. Málek, Z. Koldovský, J. Žďánský and J. Nouza, "Enhancement of Noisy Speech Recordings via Blind Source Separation", *Proceedings of the 9th Annual Conference of the International Speech Communication Association (Interspeech 2008)*, pp. 159-162, ISSN: 1990-9772, September 22-26, Brisbane, Australia, 2008. (bodové ohodnocení publikace: 4 body, podíl 30%, body: 2)

Z. Koldovský and P. Tichavský, "Time-domain Blind Audio Source Separation Using Advanced Component Clustering and Reconstruction", *Proc. of The Joint Workshop on Hands-free Speech*

Communication and Microphone Arrays (HSCMA 2008), pp. 216-219, ISBN: 978-1-4244-2338-5, May 6-8, Trento, Italy, 2008. (bodové ohodnocení publikace: 4 body, podíl 60%, body: 2)

Z. Koldovský, J. Málek, P. Tichavský, Y. Deville, and S. Hosseini, "Extension of EFICA Algorithm for Blind Separation of Piecewise Stationary Non Gaussian Sources", *33rd International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2008)*, Las Vegas, Nevada, pp. 1913-1916, ISBN: 1-4244-1484-9, April 2008. (bodové ohodnocení publikace: 4 body, podíl 60%, body: 2)

Z. Koldovský and P. Tichavský, "Time-Domain Blind Audio Source Separation Using Advanced ICA Methods", *Proceedings of the 8th Annual Conference of the International Speech Communication Association (Interspeech 2007)*, pp. 846-849, August 2007. (bodové ohodnocení publikace: 4 body, podíl 70%, body: 3)

P. Tichavský, Z. Koldovský and E. Oja, "Speed and Accuracy Enhancement of Linear ICA Techniques Using Rational Nonlinear Functions", *Proceedings of 7th International Conference on Independent Component Analysis (ICA2007)*, pp. 285-292, Sept. 2007. (bodové ohodnocení publikace: 4 body, podíl 30%, body: 1)

Z. Koldovský and P. Tichavský, "Blind Instantaneous Noisy Mixture Separation with Best Interference-plus-noise Rejection", *Proceedings of 7th International Conference on Independent Component Analysis (ICA2007)*, pp. 730-737, Sept. 2007. (bodové ohodnocení publikace: 4 body, podíl 70%, body: 3)

G. G. Herrero, Z. Koldovský, P. Tichavský, and K. Egiazarian, "A Fast Algorithm for Blind Separation of Non-Gaussian and Time-Correlated Signals", *Proceedings of 15th European Signal Processing Conference (EUSIPCO 2007)*, pp. 1731-1735, Sept 2007. (bodové ohodnocení publikace: 4 body, podíl 20%, body: 1)

J. Málek, Z. Koldovský, S. Hosseini, and Y. Deville, "A Variant of EFICA Algorithm with Adaptive Parametric Density Estimator", *8th International Workshop on Electronics, Control, Modelling, Measurement, and Signals (ECMS 2007)*, pp. 79-84, Liberec, Czech Republic, May 2007. (bodové ohodnocení publikace: 2 body, podíl 40%, body: 1)

Z. Koldovský, J. Nouza, and J. Kolorenč, "Continuous Time-Frequency Masking Method for Blind Speech Separation with Adaptive Choice of Threshold Parameter Using ICA", *Interspeech 2006*, Pittsburgh PA, USA, 17.-21. September, pp. 2578-2581, 2006. (bodové ohodnocení publikace: 4 body, podíl 75%, body: 3)

P. Tichavský, Z. Koldovský, E. Doron, A. Yeredor, and G. G. Herrero, "Blind signal separation by combining two ICA algorithms: HOS-based EFICA and time structure-based WASOBI", *Proceedings of The 2006 European Signal Processing Conference (EUSIPCO'2006)*, Florence, Sep. 2006. (bodové ohodnocení publikace: 4 body, podíl 30%, body: 2)

Z. Koldovský, P. Tichavský, "Methods of Fair Comparison of Performance of Linear ICA Techniques in Presence of Additive Noise", *Proc. of ICASSP 2006*, Toulouse, no. V., pp. 873-876, May 2006. (bodové ohodnocení publikace: 4 body, podíl 60%, body: 2)

Z. Koldovský, P. Tichavský, "Efficient Variant of Algorithm FastICA for Independent Component Analysis Attaining the Cramér-Rao Lower Bound", *Proc. of SSP-2005*, Bordeaux, July 2005. (bodové ohodnocení publikace: 4 body, podíl 60%, body: 2)

P. Tichavský, Z. Koldovský and E. Oja, "Asymptotic Performance Analysis of the Fixed-Point Algorithm (Fast-ICA) for Independent Component Analysis", *Proc. of SSP-2005*, Bordeaux, July 2005. (bodové ohodnocení publikace: 4 body, podíl 30%, body: 2)

Z. Koldovský, P. Tichavský and E. Oja, "Cramér-Rao Lower Bound for Linear Independent Component Analysis., *Proc. of ICASSP-2005*, Philadelphia, Vol. III, pp. 581 - 584, March 2005. (bodové ohodnocení publikace: 4 body, podíl 50%, body: 2)

Citace zahraniční

Citace v ISI Web of Knowledge po vyloučení všech přímých a nepřímých autocitací. Bodově ohodnoceno 3 body / citace, vzhledem k průměrnému 50 % podílu na článcích.

Citující publikace:

Sui, J; Adali, T; Pearlson, G; Yang, HH; Sponheim, SR; White, T; Calhoun, VD A CCA plus ICA based model for multi-task brain imaging data fusion and its application to schizophrenia, NEUROIMAGE, MAY 15, 2010

Li, HL; Adali, T, A class of complex ICA algorithms based on the kurtosis cost function, IEEE TRANSACTIONS ON NEURAL NETWORKS, MAR, 2008

Novey, M; Adali, T; Roy, A, A Complex Generalized Gaussian Distribution-Characterization, Generation, and Estimation, IEEE TRANSACTIONS ON SIGNAL PROCESSING, MAR, 2010

Chan, TH; Ma, WK; Chi, CY; Wang, Y, A convex analysis framework for blind separation of non-negative sources, IEEE TRANSACTIONS ON SIGNAL PROCESSING, OCT, 2008

Shimizu, S; Hoyer, PO; Hyvarinen, A; Kerminen, A, A linear non-Gaussian acyclic model for causal discovery, JOURNAL OF MACHINE LEARNING RESEARCH, OCT, 2006

Vrins, F; Lee, JA; Verleysen, M A minimum-range approach to blind extraction of bounded sources, IEEE TRANSACTIONS ON NEURAL NETWORKS, MAY, 2007

Taghia, J; Doostari, MA; Taghia, J, A new algorithm in blind source separation for high dimensional data sets such as MEG data, PROCEEDINGS OF THE WSEAS INTERNATIONAL CONFERENCE ON CIRCUITS, SYSTEMS, ELECTRONICS, CONTROL & SIGNAL PROCESSING - SELECTED TOPICS ON CIRCUITS, SYSTEMS, ELECTRONICS, CONTROL & SIGNAL PROCESSING , 2007

Karvounis, EC; Tsipouras, MG; Papaloukas, C; Tsalikakis, DG; Naka, KK; Fotiadis, DI A Non-invasive Methodology for Fetal Monitoring during Pregnancy,METHODS OF INFORMATION IN MEDICINE, 2010

Li, XL; Adali, T, A NOVEL ENTROPY ESTIMATOR AND ITS APPLICATION TO ICA, 2009 IEEE INTERNATIONAL WORKSHOP ON MACHINE LEARNING FOR SIGNAL PROCESSING, 2009

Xue, YF; Ju, F; Wang, YJ; Yang, J, A source adaptive independent component analysis algorithm through solving the estimating equation, EXPERT SYSTEMS WITH APPLICATIONS, DEC 2009.

Hild, KE; Attias, HT; Nagarajan, SS, An expectation-maximization method for spatio-temporal blind source separation using an AR-MOG source model, IEEE TRANSACTIONS ON NEURAL NETWORKS, MAR, 2008

Phlypo, R; Zarzoso, V; Lemahieu, I, Atrial activity estimation from atrial fibrillation ECGs by blind source extraction based on a conditional maximum likelihood approach MEDICAL & BIOLOGICAL ENGINEERING & COMPUTING, MAY, 2010

Du, Q; Kopriva, I, Automated target detection and discrimination using constrained kurtosis maximization, IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, JAN, 2008

Shan, XM; Apley, DW, Blind Identification of Manufacturing Variation Patterns by Combining Source Separation Criteria, TECHNOMETRICS, AUG, 2008

Cabras, G; Carniel, R; Wassermann, J, BLIND SOURCE SEPARATION: AN APPLICATION TO THE MT. MERAPI VOLCANO, INDONESIA, FLUCTUATION AND NOISE LETTERS, DEC, 2008

Frolov, AA; Husek, D; Muraviev, IP; Polyakov, PY, Boolean factor analysis by autractor neural network, IEEE TRANSACTIONS ON NEURAL NETWORKS, MAY, 2007

Ollila, E; Kim, HJ; Koivunen, V, Compact Cramer-Rao bound expression-for independent component analysis, IEEE TRANSACTIONS ON SIGNAL PROCESSING, APR, 2008

Zarzoso, V; Comon, P, Comparative speed analysis of FastICA, Independent Component Analysis and Signal Separation, Proceedings, 2007

Novey, M; Adali, T, Complex ICA by negentropy maximization, IEEE TRANSACTIONS ON NEURAL NETWORKS, APR, 2008

Adali, T; Li, HL; Novey, M; Cardoso, JF, Complex ICA using nonlinear functions, IEEE TRANSACTIONS ON SIGNAL PROCESSING, SEP, 2008

Li, XL; Adali, T, Complex Independent Component Analysis by Entropy Bound Minimization, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS I-REGULAR PAPERS, JUL, 2010

Lv, JC; Tan, KK; Yi, Z; Huang, SN, Convergence Analysis of a Class of Hyvarinen-Oja's ICA Learning Algorithms With Constant Learning Rates, IEEE TRANSACTIONS ON SIGNAL PROCESSING, MAY, 2009

Wibral, M; Turi, G; Linden, DEJ; Kaiser, J; Bledowski, C, Decomposition of working memory-related scalp ERPs: Crossvalidation of fMRI-constrained source analysis and ICA, INTERNATIONAL JOURNAL OF PSYCHOPHYSIOLOGY, MAR, 2008

Yao, ZX; Zhang, K; Liu, HB; Su, H, Eliminate indeterminacies of independent component analysis for chemometrics, PROGRESS IN NATURAL SCIENCE, AUG 10, 2008

Filippi, C; Romanin-Jacur, G, Exact and approximate algorithms for high-multiplicity parallel machine scheduling, JOURNAL OF SCHEDULING, OCT, 2009

Wang, Y; Chi, YL; Wu, X; Liu, C, Extracting Acoustical Impulse Signal of Faulty Bearing Using Blind Deconvolution Method, ICICTA: 2009 SECOND INTERNATIONAL CONFERENCE ON INTELLIGENT COMPUTATION TECHNOLOGY AND AUTOMATION, VOL I, PROCEEDINGS, 2009

Kopriva, I; Jeric, I; Smrecki, V, Extraction of multiple pure component H-1 and C-13 NMR spectra from two mixtures: Novel solution obtained by sparse component analysis-based blind decomposition, ANALYTICA CHIMICA ACTA, OCT 27, 2009

Solvang, HK; Nagahara, Y; Araki, S; Sawada, H; Makino, S, Frequency-Domain Pearson Distribution Approach for Independent Component Analysis (FD-Pearson-ICA) in Blind Source Separation, IEEE TRANSACTIONS ON AUDIO SPEECH AND LANGUAGE PROCESSING, MAY, 2009

Cantero, JL; Atienza, M; Gomez-Herrero, G; Cruz-Vadell, A; Gil-Neciga, E; Rodriguez-Romero, R; Garcia-Solis, D Functional Integrity of Thalamocortical Circuits Differentiates Normal Aging from Mild Cognitive Impairment, HUMAN BRAIN MAPPING, DEC, 2009

Tiilikainen, J; Bosund, V; Tilli, JM; Sormunen, J; Mattila, M; Hakkarainen, T; Lipsanen, H, Genetic algorithm using independent component analysis in x-ray reflectivity curve fitting of periodic layer structures, JOURNAL OF PHYSICS D-APPLIED PHYSICS, OCT 7, 2007

Liu, CJ; Yang, J, ICA Color Space for Pattern Recognition, IEEE TRANSACTIONS ON NEURAL NETWORKS, FEB, 2009

Fadili, MJ; Starck, JL; Bobin, J; Moudden, Y, Image Decomposition and Separation Using Sparse Representations: An Overview, PROCEEDINGS OF THE IEEE, JUN, 2010

Sarmiento, A; Cruces, S; Duran, I, Improvement of the Initialization of ICA Time-Frequency Algorithms for Speech Separation, INDEPENDENT COMPONENT ANALYSIS AND SIGNAL SEPARATION, PROCEEDINGS, 2009

Cantero, JL; Atienza, M; Cruz-Vadell, A; Suarez-Gonzalez, A; Gil-Neciga, E, Increased synchronization and decreased neural complexity underlie thalamocortical oscillatory dynamics in mild cognitive impairment, NEUROIMAGE, JUL 15, 2009

Xue, YF; Wang, YJ; Yang, J, Independent component analysis based on gradient equation and kernel density estimation, NEUROCOMPUTING, MAR, 2009

Nath, MK; Sahambi, JS Independent Component Analysis of Functional MRI Data, 2008 IEEE REGION 10 CONFERENCE: TENCON 2008, VOLS 1-4, 2008

Nath, MK, Independent Component Analysis of Real Data, ICAPR 2009: SEVENTH INTERNATIONAL CONFERENCE ON ADVANCES IN PATTERN RECOGNITION, PROCEEDINGS, 2009

Gomez-Herrero, G; Atienza, M; Egiazarian, K; Cantero, JL, Measuring directional coupling between EEG sources, NEUROIMAGE, NOV 15, 2008

Haufe, S; Tomioka, R; Nolte, G; Muller, KR; Kawanabe, M, Modeling Sparse Connectivity Between Underlying Brain Sources for EEG/MEG, IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING, AUG, 2010

Wang, FY; Chi, CY; Chan, TH; Wang, Y, Nonnegative Least-Correlated Component Analysis for Separation of Dependent Sources by Volume Maximization, IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE , MAY, 2010

Erdogan, AT, On the Convergence of ICA Algorithms With Symmetric Orthogonalization, IEEE TRANSACTIONS ON SIGNAL PROCESSING, JUN, 2009

Zayyani, H; Babaie-zadeh, M; Haddadi, F; Jutten, C, On the Cramer-Rao Bound for Estimating the Mixing Matrix in Noisy Sparse Component Analysis, IEEE SIGNAL PROCESSING LETTERS, 2008

Taghia, J; Taghia, J, One-Channel Audio Source Separation of Convolutive Mixture, ADVANCES IN COMPUTER AND INFORMATIOM SCIENCES AND ENGINEERING, 2008

Yang, XN; Yao, JL; Li, JD; Li, Z, PERFORMANCE ANALYSIS OF THE FASTICA ALGORITHM IN ICA-BASED CO-CHANNEL COMMUNICATION SYSTEM, 2009 5TH INTERNATIONAL CONFERENCE ON WIRELESS COMMUNICATIONS, NETWORKING AND MOBILE COMPUTING, VOLS 1-8, 2009

Chen, HB; Tse, CK; Feng, JC, Performance evaluation of source extraction in wireless sensor networks, COMPUTER COMMUNICATIONS, OCT 25, 2008

Pokharel, PP; Ozertem, U; Erdogmus, D; Principe, JC, Recursive complex BSS via generalized eigendecomposition and application,in image rejection for BPSK, SIGNAL PROCESSING, JUN, 2008
Sadeghi, MH; Aghabozorgi, MR; Sadeghi, MT, Removing Reflection from Image Using ICA, 2008 INTERNATIONAL SYMPOSIUM ON TELECOMMUNICATIONS, VOLS 1 AND 2, 2008

Kopriva, I; Persin, A; Puizina-Ivic, N; Miric, L, Robust demarcation of basal cell carcinoma by dependent component analysis-based segmentation of multi-spectral fluorescence images, JOURNAL OF PHOTOCHEMISTRY AND PHOTOBIOLOGY B-BIOLOGY, JUL 2, 2010

Zarzoso, V; Comon, P, Robust Independent Component Analysis by Iterative Maximization of the Kurtosis Contrast With Algebraic Optimal Step Size, IEEE TRANSACTIONS ON NEURAL NETWORKS, FEB, 2010

Zarzoso, V; Comon, P, Robust Independent Component Analysis for Blind Source Separation and Extraction with Application in Electrocardiography, 2008 30TH ANNUAL INTERNATIONAL CONFERENCE OF THE IEEE ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY, VOLS 1-8, 2008

Cabras, G; Carniel, R; Wasserman, J, Signal enhancement with generalized ICA applied to Mt. Etna volcano, Italy, BOLLETTINO DI GEOFISICA TEORICA ED APPLICATA, MAR, 2010

Petrochilos, N; Hafner, N; Host-Madsen, A; Boric-Lubecke, O; Lubecke, V, Source separation applied to heartbeat Doppler radar, PROCEEDINGS OF THE NINTH IASTED INTERNATIONAL CONFERENCE ON SIGNAL AND IMAGE PROCESSING, 2007

Gaito, S; Grossi, G, Speeding up FastICA by mixture random pruning, Independent Component Analysis and Signal Separation, Proceedings, 2007

Herrmann, JM; Theis, FJ, Statistical analysis of sample-size effects in ICA, INTELLIGENT DATA ENGINEERING AND AUTOMATED LEARNING - IDEAL 2007, 2007

Huang, B; Yan, GZ; Zan, P; Li, QR, Study on gastric interdigestive pressure activity based on phase space reconstruction and FastICA algorithm, MEDICAL ENGINEERING & PHYSICS, APR, 2009

Ollila, E, The Deflation-Based FastICA Estimator: Statistical Analysis Revisited, IEEE TRANSACTIONS ON SIGNAL PROCESSING, MAR, 2010

Kleffner, MD; Jones, DL, Tighter mean-squared error bounds on kurtosis-based fast-ICA, IEEE/SP 14TH WORKSHOP ON STATISTICAL SIGNAL PROCESSING, VOLS 1 AND 2, 2007

Sadeghi, MH; Aghabozorgi, MR; Sadeghi, MT, Using Basis Vectors of ICA for Blind Separation of Reflection from a Single Image, ICSP: 2008 9TH INTERNATIONAL CONFERENCE ON SIGNAL PROCESSING, VOLS 1-5, PROCEEDINGS, 2008

Sui, J; Adali, T; Li, YO; Yang, HH; Calhoun, VD, A review of multivariate methods in brain imaging data fusion, MEDICAL IMAGING 2010: BIOMEDICAL APPLICATIONS IN MOLECULAR, STRUCTURAL, AND FUNCTIONAL IMAGING Proceedings of SPIE-The International Society for Optical Engineering, Conference on Medical Imaging 2010 - Biomedical Applications in Molecular, Structural, and Functional Imaging, FEB 14-16, 2010

Kim, J; Yang, HJ; Jung, BW; Chun, J, Blind Calibration for a Linear Array With Gain and Phase Error Using Independent Component Analysis, IEEE ANTENNAS AND WIRELESS PROPAGATION LETTERS, 2010

Yeredor, A, Blind Separation of Gaussian Sources With General Covariance Structures: Bounds and Optimal Estimation, IEEE TRANSACTIONS ON SIGNAL PROCESSING, 2010

Li, XL; Adali, T, Independent Component Analysis by Entropy Bound Minimization, IEEE TRANSACTIONS ON SIGNAL PROCESSING, 2010

Kopriva, I; Cichocki, A, Nonlinear Band Expansion and 3D Nonnegative Tensor Factorization for Blind Decomposition of Magnetic Resonance Image of the Brain, LATENT VARIABLE ANALYSIS AND SIGNAL SPARATION, Lecture Notes in Computer Science, 9th International Conference on Latent Variable Analysis and Signal Separation, SEP 27-30, 2010

Adamo, A; Grossi, G, Random Pruning of Blockwise Stationary Mixtures for Online BSS, LATENT VARIABLE ANALYSIS AND SIGNAL SPARATION, Lecture Notes in Computer Science, 9th International Conference on Latent Variable Analysis and Signal Separation, SEP 27-30, 2010

Sadeghi, MH; Aghabozorgi, MR; Sadeghi, MT, Removing Reflection from Image Using ICA, 2008 INTERNATIONAL SYMPOSIUM ON TELECOMMUNICATIONS, VOLS 1 AND 2, International Symposium on Telecommunications, AUG 27-28, 2008

Araki, S; Ozerov, A; Gowreesunker, V; Sawada, H; Theis, F; Nolte, G; Lutter, D; Duong, NQK, The 2010 Signal Separation Evaluation Campaign (SiSEC2010): Audio Source SeparationLATENT VARIABLE ANALYSIS AND SIGNAL SPARATION, Lecture Notes in Computer Science, 9th International Conference on Latent Variable Analysis and Signal Separation, SEP 27-30, 2010

Oja, E; Yuan, ZJ, The FastICA algorithm revisited: Convergence analysis, IEEE TRANSACTIONS ON NEURAL NETWORKS, 2006

Granegger, M; Werther, T; Gilly, H, Use of independent component analysis for reducing CPR artefacts in human emergency ECGs, RESUSCITATION, 2011

Citace ČR (po vyloučení všech přímých a nepřímých autocitací):

L. RUČKAY, J. ŠŤASTNÝ, P. SOVKA, „ICA Model Order Estimation Using Clustering Method," *RADIOENGINEERING,* VOL. 16, NO. 4, DECEMBER 2007.

citovaná publikace:

Z. Koldovský, Fast and Accurate Methods for Independent Component Analysis. Ph.D. thesis, CTU Prague, Faculty of Nuclear Sciences and Physical Engineering, Dept. of Mathematics, 2006. (body: 2)

Udělený grant ČR

GAČR 102/07/P384, Použití pokročilých metod pro analýzu nezávislých komponent na slepou separaci reálných signálů (ukončený 2009). (body: 4)

GAČR P103/11/1947, Metody analýzy latentních proměnných ve slepém zpracování řečových a akustických signálů (schválen od roku 2011). (body: 4)

CTU0508214 (interní grant ČVUT v Praze), Analýza nezávislých komponent (ukončený 2005) (body: 1)

Interní grant na rok 2007 FM, TUL, Vývoj rychlých a eficientních ICA algoritmů a jejich aplikace na slepou separaci audio signálů v časové oblasti (ukončený 2007). (body: 1)

SGS TUL 2010, Pokročilé metody zpracování signálů a návrhu elektronických systémů (probíhající do listopadu 2010) (body: 2)

Realizované dílo

T-ABCD – Algoritmus pro slepou separaci reálných nahrávek akustických signálů byl implementován v Matlabu a bylo vytvořeno grafické uživatelské rozhraní pro jeho ovládání. Jeho on-line varianta byla implementována v C++ a publikována v:

O. Hnilička, J. Málek, K. Paleček, Z. Koldovský, "A Fast C++ Implementation of Time-domain Blind Speech Separation Algorithm, " *Proc. of the 20th Czech-German WorkShop on Speech Processing*, Oct. 2010.
(body: 5)

Zdůvodnění bodového hodnocení

B2. Pedagogická a vzdělávací činnost

Přednášení v řádném studiu minimálně 2 hod/týd.

Biologické a akustické signály (BSI) – garant předmětu, zavedení předmětu, příprava a vedení přednášek i cvičení; pro pátý ročník v posledním semestru magisterského studia na FM (10 týdnů), odučeny 2 semestry: LS 2009, LS 2010. (body: 4)

Výpočty, simulace a vizualizace Matlab (MATLB) – garant předmětu, zavedení předmětu, příprava a vedení přednášek i cvičení; pro druhý semestr bakalářského studia na FM (14 týdnů), odučeny 2 semestry: LS 2009, LS 2010. (body: 4)

Počítačové zpracování signálů (PZS) – příprava a vedení poloviny přednášek a cvičení; pro dvouleté navazující magisterské studium na FM, odučeny 2 semestry: LS 2007, LS 2008. (body: 2)

Základy matematické statistiky – zavedení předmětu, příprava a vedení přednášek i cvičení; pro druhý ročník bakalářského studia na FJFI, odučeny 2 semestry: ZS 2003, ZS 2004. (body: 4)

Pravidelná cvičení minimálně 2 hod/týd.

Signály a informace (SGI) – vedení cvičení (minimálně jeden kruh); pro tříleté bakalářské studium na FM, ZS 2006-2010. (body: 2.5)

Biologické a akustické signály (BSI) –příprava a vedení cvičení; pro pátý ročník v posledním semestru magisterského studia na FM (10 týdnů), odučeny 2 semestry: LS 2009, LS 2010. (body: 1)

Výpočty, simulace a vizualizace Matlab (MATLB) –příprava a vedení cvičení; pro druhý semestr bakalářského studia na FM (14 týdnů), odučeny 2 semestry: LS 2009, LS 2010. (body: 1)

Počítačové zpracování signálů (PZS) –vedení cvičení; pro dvouleté navazující magisterské studium na FM, odučeny 2 semestry: LS 2007, LS 2008. (body: 0.5)

Matematické hry –vedení cvičení; pro první ročník bakalářského studia na FJFI, odučen 1 semestr: ZS 2002. (body: 0.5)

Lineární algebra A1 a A2 – vedení cvičení; pro první ročník magisterského studia na FJFI, 1x ZS 2000, 1x ZS a LS 2001, 1x LS 2002, 2x ZS 2002, 2x ZS a LS 2003 a 2004 a 2x LS 2005. (body: 8)

Vybrané partie z matematiky – vedení cvičení; pro druhý ročník magisterského studia na FJFI, LS 2004 a LS 2005. (body: 1)

Základy matematické statistiky – vedení cvičení; pro druhý ročník bakalářského studia na FJFI, odučeny 2 semestry: ZS 2003, ZS 2004. (body: 1)

Vedení studentského projektu

Martin Marek: Odstraňování artefaktů z EEG záznamů pomocí metod slepé separace, 2007.
Jakub Petkov: Advanced Control of Blind Source Separation Algorithms in Matlab (BSSGUI), 2008.
Daniel Hančil: Databáze zvuků pro slepou separaci, 2008.
Petr Schovanec a Jan Pišta: Vyhodnocování hlučnosti zvukového signálu, 2009. (2 semestry)
Radek Hlávka: Realizace audioefektů v C#, 2010.

Augustin Bernard: Detekce základní frekvence v akustickém signálu, 2010. (2 semestry)
David Botka: Návrh a realizace audio efektů, 2010. (2 semestry)
(celkem body: 5)

Zavedení nového předmětu v řádném studiu

Biologické a akustické signály (BSI) – garant předmětu, zavedení předmětu pro pátý ročník v posledním semestru magisterského studia na FM (10 týdnů). (body: 4)

Výpočty, simulace a vizualizace Matlab (MATLB) – garant předmětu, zavedení předmětu pro druhý semestr bakalářského studia na FM (14 týdnů). (body: 4)

Počítačové zpracování signálů (PZS) – příprava poloviny přednášek a cvičení; pro dvouleté navazující magisterské studium na FM. (body: 2)

Základy matematické statistiky – zavedení předmětu pro druhý ročník bakalářského studia na FJFI. (body: 4)

Matematické hry – zavedení předmětu, pro první ročník bakalářského studia na FJFI. (body: 2)

Jiná knižní publikace, didakt. pomůcka

Odborná kniha z oblasti zpracování a rozpoznávání řeči:
Nouza, J. a kol.: Řeč a počítač: principy hlasové komunikace, úlohy, metody a aplikace, vydavatel: Technická univerzita v Liberci, editoři: Jan Nouza, Zbyněk Koldovský, Robert Vích, číslo publikace: 55-110-09, první vydání, počet stran: 238, ISBN 978-80-7372-548-8, 2009.
kapitola – Zpracování záznamu řeči z více mikrofonů metodami pro řešení inverzního problému“, str. 201-216 (80%)
(body: 2)

Poznámky k předmětu Počítačové zpracování signálů (PZS) zveřejněné na internetu (http://itakura.ite.tul.cz/zbynek/teaching.htm), 31 stran. (body: 2)

Poznámky k předmětu Biologické a akustické signály (BSI), zveřejněné na internetu (http://itakura.ite.tul.cz/zbynek/teaching.htm), 32 stran. (body: 2)

Vedení doktoranda před/po zkoušce

Ing. Jiří Málek, zahájení studia 1.8.2006, studijní obor: Technická kybernetika, disertační práce má název *Blind Audio Source Separation via Independent Component Analysis,* odevzdáno listopad-prosinec 2010.
(body: 4)

Vedení obhájené diplomové práce

Filip Žižka: Aplikace metod pro slepou separaci dat ve zpracování EKG záznamů, 2007.
Michal Kuna: Slepá separace řeči ze stereofonního záznamu, 2007.
Luděk Sviták: Zpracování EEG záznamů s použitím nejnovějších technik pro slepou separaci, 2009.
(body: 3)

Garant doktor. PGS, člen obor. rady PGS

Garant předmětu: Metody digitálního zpracování vícerozměrných signálů (body: 3)

Coauthors' Approval of Zbyněk Koldovský's Contributions in Joint Papers (Since 2006)

Journal papers

Z. Koldovský and P. Tichavský, "Time-Domain Blind Separation of Audio Sources on the basis of a Complete ICA Decomposition of an Observation Space", accepted for publication in *IEEE Trans. on Speech, Audio and Language Processing*, April 2010. 65 %

Z. Koldovský, J. Málek, P. Tichavský, Y. Deville, and S. Hosseini, "Blind Separation of Piecewise Stationary NonGaussian Sources", Signal Processing, Volume 89, Issue 12, Pages 2570-2584, ISSN 0165-1684, December 2009. 60 %

P. Tichavský, Z. Koldovský, A. Yeredor, G. G. Herrero, and E. Doron, "A Hybrid Technique for Blind Separation of Non-Gaussian and Time-Correlated Sources Using a Multicomponent Approach", IEEE Trans. on Neural Networks, Vol. 19, No. 3, pp. 421-430, ISSN: 1045-9227, March 2008. 30 %

Conference papers

P. Tichavský and Z. Koldovský, "Simultaneous Search for All Modes in Multilinear Models," ICASSP 2010, pp. 4114-4117, ISBN: 978-1-4244-4296-6, ISSN: 1520-6149, Dallas, USA, March 2010. 30 %

Z. Koldovský and P. Tichavský, "A Comparison of Independent Component and Independent Subspace Analysis Algorithms," EUSIPCO 2009 , pp. 1447-1451, Glasgow, Scotland, August 24-28, 2009. 60 %

P. Tichavský, A. Yeredor, and Z. Koldovský, "A Fast Asymptotically Efficient Algorithm for Blind Separation of a Linear Mixture of Block-Wise Stationary Autoregressive Processes," ICASSP 2009, pp. 3133-3136, Taipei, Taiwan, April 2009. 15 %

Z. Koldovský and P. Tichavský, "Time-domain Blind Audio Source Separation Using Advanced Component Clustering and Reconstruction", Proc. of The Joint Workshop on Hands-free Speech Communication and Microphone Arrays (HSCMA 2008), pp. 216-219, ISBN: 978-1-4244-2338-5, May 6-8, Trento, Italy, 2008. 60 %

Z. Koldovský, J. Málek, P. Tichavský, Y. Deville, and S. Hosseini, "Extension of EFICA Algorithm for Blind Separation of Piecewise Stationary Non Gaussian Sources", 33rd International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2008), Las Vegas, Nevada, pp. 1913-1916, ISBN: 1-4244-1484-9, April 2008. 60 %

Z. Koldovský and P. Tichavský, "Time-Domain Blind Audio Source Separation Using Advanced ICA Methods", Proceedings of the 8th Annual Conference of the International Speech Communication Association (Interspeech 2007), pp. 846-849, August 2007. 70 %

P. Tichavský, Z. Koldovský and E. Oja, "Speed and Accuracy Enhancement of Linear ICA Techniques Using Rational Nonlinear Functions", Proceedings of 7th International Conference on Independent Component Analysis (ICA2007), pp. 285-292, Sept. 2007. 30 %

Z. Koldovský and P. Tichavský, "Blind Instantaneous Noisy Mixture Separation with Best Interference-plus-noise Rejection", Proceedings of 7th International Conference on Independent Component Analysis (ICA2007), pp. 730-737, Sept. 2007. 70 %

G. G. Herrero, Z. Koldovský, P. Tichavský, and K. Egiazarian, "A Fast Algorithm for Blind Separation of Non-Gaussian and Time-Correlated Signals", Proceedings of 15th European Signal Processing Conference (EUSIPCO 2007), pp. 1731-1735, Sept 2007. 20 %

Z. Koldovský, J. Nouza, and J. Kolorenč, "Continuous Time-Frequency Masking Method for Blind Speech Separation with Adaptive Choice of Threshold Parameter Using ICA", Interspeech 2006, Pittsburgh PA, USA, 17.-21. September, pp. 2578-2581, 2006. 75 %

P. Tichavský, Z. Koldovský, E. Doron, A. Yeredor, and G. G. Herrero, "Blind signal separation by combining two ICA algorithms: HOS-based EFICA and time structure-based WASOBI", Proceedings of The 2006 European Signal Processing Conference (EUSIPCO'2006), Florence, Sep. 2006. 30 %

Z. Koldovský, P. Tichavský, "Methods of Fair Comparison of Performance of Linear ICA Techniques in Presence of Additive Noise", Proc. of ICASSP 2006, Toulouse, no. V., pp. 873-876, May 2006. 60 %

Coauthors

Yannick Deville .....

Jiří Málek .....

Jan Nouza .....

Erkki Oja .....

Petr Tichavský .....

Arie Yeredor .....

Prof. Ing. Václav Kopecký, CSc.
děkan Fakulty mechatroniky, informatiky a mezioborových studií
Technická univerzita v Liberci
Studentská 2
461 17 Liberec

Vážený pane děkane,

na základě zákona o vysokých školách – 111/98 Sb., dle §72 čl. 2 žádám tímto o zahájení habilitačního řízení pro obor Technická kybernetika. K žádosti přikládám materiály vyžadované výše uvedeným zákonem.

S pozdravem

Ing. Zbyněk Koldovský, Ph.D.

Přílohy:

1. Životopis obsahující přehled o odborné a pedagogické činnosti
2. Doklady o dosaženém vysokoškolském vzdělání a získaných příslušných titulech
3. Přehled pedagogické praxe
4. Seznam vědeckých a odborných prací
5. Přehled absolvovaných vědeckých a odborných stáží
6. Habilitační práce (4x)
7. Podklady pro habilitační a jmenovací řízení na Fakultě mechatroniky, informatiky a mezioborových studií TU v Liberci (A-kvalitativní hodnocení)
8. Podklady pro habilitační a jmenovací řízení na Fakultě mechatroniky, informatiky a mezioborových studií TU v Liberci (B-kvantitativní hodnocení)
9. Potvrzení podílů spoluautorství
10. CD s elektronickou verzí všech výše uvedených dokumentů

V Liberci dne 30.11.2010

Ing. Zbyněk Koldovský, Ph.D.

Narozen: 25.4.1979 v Jablonci nad Nisou, ženatý
Národnost: Česká

Pracoviště: Ústav informačních technologií a elektroniky
Fakulta mechatroniky, informatiky a mezioborových studií
Technická univerzita v Liberci
Studentská 2
46 117, Liberec

Vedlejší pracoviště: Ústav teorie informace a automatizace AV ČR v.v.i.
Akademie věd České republiky
Pod vodárenskou věží 4
182 08, Praha 8

Telefon: +420 48 535 3534
Fax: +420 48 535 3112
E-mail: zbynek.koldovsky@tul.cz

VZDĚLÁNÍ

Ing. rok 2002
Fakulta jaderná a fyzikálně inženýrská
České vysoké učení technické v Praze
školitel: Ing. Petr Tichavský, CSc.
diplomová práce: Analýza nezávislých komponent v EEG datech

Ph.D. rok 2006
disertační práce: Fast and Accurate Methods for Independent Component Analysis
Fakulta jaderná a fyzikálně inženýrská
České vysoké učení technické v Praze
školitel: Ing. Petr Tichavský, CSc.

VĚDECKÁ A PEDAGOGICKÁ PRAXE

2000 – 2002 Lektor cvičení z lineární algebry A1 a A2 na FJFI ČVUT.

léto 2001 Pomocný vědecký pracovník v Ústavu termodynamiky AV ČR. Řešení diferenciálních rovnic mezní vrstvy atmosféry.

léto 2000 Dohoda o provedení práce v Ústavu teorie informace a automatizace AV ČR. Implementace algoritmu FastICA a algoritmu pro odhad vzájemné informace náhodných proměnných v C++.

2002 – 2005 VŠ asistent na Katedře matematiky FJFI, ČVUT v Praze

2002 – dosud Odborný pracovník VaV
Ústav teorie informace a automatizace AV ČR v.v.i. (0.3 úvazek)

2005 – dosud VŠ asistent na Katedře elektroniky a zpracování signálů
(dnes Ústav informačních technologií a elektroniky, ITE), FM, TUL.

OBLASTI VÝZKUMU

Digitální zpracování signálů, metody slepé separace, analýza nezávislých komponent, lineární a multilineární algebra

Pedagogické aktivity (VŠ předměty)

Současné:
Biologické a akustické signály (garant, přednášky, cvičení, FM, TUL)
Výpočty, simulace a vizualizace Matlab (garant, přednášky, cvičení, FM, TUL)
Digital Signal Processing 1 (garant, přednášky, cvičení, FM, TUL)

Minulé:
Počítačové zpracování signálů (přednášky, cvičení, FM, TUL)
Signály a informace (cvičení, FM, TUL)
Lineární algebra A1,A2 (cvičení, FJFI, ČVUT)
Vybrané partie z matematiky (cvičení, FJFI, ČVUT)
Matematické hry (cvičení, FJFI, ČVUT)
Základy matematické statistiky (přednášky a cvičení, FJFI, ČVUT)

vedení obhájených 3 diplomových a 3 bakalářských prací a 7 projektů

Účast na řešení projektů

řešitel:

- CTU0508214 (interní grant ČVUT v Praze), Analýza nezávislých komponent (ukončený 2005).
- GAČR 102/07/P384, Použití pokročilých metod pro analýzu nezávislých komponent na slepou separaci reálných signálů (ukončený 2009).
- Interní grant na rok 2007 FM, TUL, Vývoj rychlých a eficientních ICA algoritmů a jejich aplikace na slepou separaci audio signálů v časové oblasti (ukončený 2007).
- Projekt pro Škoda a.s., Vyhodnocování hluku předních stěračů (2008-2010).
- GAČR P103/11/1947, Metody analýzy latentních proměnných ve slepém zpracování řečových a akustických signálů (schválen od roku 2011).
- SGS TUL 2010 a 2011, Pokročilé metody zpracování signálů a návrhu elektronických systémů (březen-listopad 2010 a 2011).

spolupracovník:

- 1M0572, Data-Algoritmy-Rozhodování (2005-dosud).
- GAČR 102/09/1278, Pokročilé metody slepé separace signálu a slepé dekonvoluce (2009-2013).
- GAČR 102/05/0278, Nové směry ve výzkumu a využití hlasových technologií (2005-2007).
- GAAV (PPCV) 1QS108040569, Asistenční, informační a komunikační služby s podporou vyspělých hlasových technologií (2005-2009).

Zahraniční návštěvy

- červen 2004, Návštěva Neural Networks Research Centre na Technické univerzitě v Helsinkách v rámci výměnných meziakademických dohod.
- březen 2005, Návštěva Neural Networks Research Centre na Technické univerzitě v Helsinkách s přednáškou na téma Efficient Version of Algorithm FastICA Attaining the Cramer-Rao Lower Bound, seminář Laboratory of Computer and Information Science.
- listopad a prosinec 2009, návštěva oddělení Univerzity ve Stuttgartu (Německo), International doctorate School in Information and Communication Technology of Trento (Itálie) a Laboratoire Astrophysique de Toulouse (Francie)
- prosinec 2010, návštěva Lab. for Advanced Brain Signal Processing, RIKEN, Tokyo (Japonsko)

Publikační aktivity

V období 2005-2010 autorství a spoluautorství 8 článků v impaktovaných zahraničních časopisech, 23 příspěvků v recenzovaných sbornících mezinárodních konferencí, kapitola v odborné knize.

Reference

Ing., Petr Tichavský, CSc., ÚTIA AV ČR, Pod vodárenskou věží 4, 182 08 Praha 8.
Prof. Ing. Jan Nouza, CSc., FM TUL, Studentská 2, 461 17 Liberec 1.
Doc. RNDr. Jan Mareš, CSc., FJFI ČVUT, Trojanova 13, 120 00 Praha 2.

V Liberci dne 30.11.2010 Ing. Zbyněk Koldovský, Ph.D

ČESKÁ REPUBLIKA

Vysoká škola **ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE**

Fakulta **FAKULTA JADERNÁ A FYZIKÁLNĚ INŽENÝRSKÁ**

Číslo diplomu 003532

Č. **2800/2002**

VYSOKOŠKOLSKÝ DIPLOM

Zbyněk Koldovský

(jméno a příjmení)

narozen(a) dne **25. dubna 1979** v **Jablonci nad Nisou** okres **Jablonec nad Nisou**

ukončil(a) studium vykonáním státní závěrečné zkoušky a získal(a) vysokoškolské vzdělání v magisterském studijním programu

Aplikace přírodních věd

studijní obor **Matematické inženýrství**

Podle § 46 odst. 4 písm. a) zákona č. 111/1998 Sb. o vysokých školách se mu (jí) uděluje akademický titul

inženýr

ve zkratce **Ing.**

V Praze dne **17. června 2002**

rektor vysoké školy

ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE L.S. -3-

děkan fakulty

VSP/2524

ČESKÁ REPUBLIKA

Vysoká škola **ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE** Číslo **2800/2002**
Fakulta **JADERNÁ A FYZIKÁLNĚ INŽENÝRSKÁ**
Příloha k diplomu číslo **003532**

VYSVĚDČENÍ
o státní závěrečné zkoušce

Zbyněk Koldovský
(jméno a příjmení)

narozen(a) dne **25. dubna 1979** v **Jablonci nad Nisou** okres **Jablonec nad Nisou**

ukončil(a) podle § 46, odst. 3 zákona č. 111/1998 Sb. o vysokých školách studium

v magisterském studijním programu **Aplikace přírodních věd**

zaměření Matematické modelování

studijní obor **Matematické inženýrství**

Vykonal(a) státní závěrečnou zkoušku
1. z **funkcionální analýzy**
dne **17. 6. 2002** s prospěchem **výborně**
2. z **teorie pravděpodobnosti a matematické statistiky**
dne **17. 6. 2002** s prospěchem **výborně**
3. z **stochastických procesů**
dne **17. 6. 2002** s prospěchem **výborně**
4. z ---
dne --- s prospěchem ---

a obhájil(a) diplomovou práci na téma **Analýza nezávislých komponent v EEG datech, slepá separace konvolutorních směsí**

dne **17. 6. 2002** s prospěchem **výborně**

Celkový výsledek státní závěrečné zkoušky: **v ý b o r n ě**

V Praze dne **17. 6. 2002**

ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE Fakulta jaderná a fyzikálně inženýrská
L.S.

Havlíček
.....
děkan fakulty

STUPNICE ZNÁMEK

jednotlivé zkoušky	výborně	velmi dobře	dobře	nevyhověl
celkový výsledek státní zkoušky	výborně	velmi dobře	dobře	nevyhověl

004110

VSP/2525

ČESKÁ REPUBLIKA

ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE

Číslo diplomu HF 0007607

DIPLOM

Ing. Zbyněk Koldovský

titul, jméno a příjmení

narozen(a) dne 25. dubna 1979 v Jablonci nad Nisou

ukončil(a) studium v doktorském studijním programu

Aplikace přírodních věd kód P3913

ve studijním oboru Matematické inženýrství

na/v Fakultě jaderné a fyzikálně inženýrské

Podle § 47 odst. 5 zákona č. 111/1998 Sb. o vysokých školách a o změně a doplnění dalších zákonů se mu (jí) přiznává akademický titul

DOKTOR

ve zkratce „**Ph.D.**“ uváděné za jménem

rektor

děkan

V Praze dne 9. května 2006

ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE -4-

L.S.

SEVT – **92 575 1**

B.N.B. 991 2001

ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE
CZECH TECHNICAL UNIVERSITY IN PRAGUE

DODATEK K DIPLOMU / DIPLOMA SUPPLEMENT

Diplom č. / Diploma No: HF 0007607

Tento dodatek k diplomu odpovídá modelu vytvořenému Evropskou komisí, Radou Evropy a organizací UNESCO/CEPES. Účelem dodatku je poskytnout odpovídající nezávislé údaje, které přispějí ke zlepšení mezinárodní "průhlednosti" a spravedlivosti akademického a profesního uznávání kvalifikací (diplomů, titulů, osvědčení atd.). Dodatek je určen pro popis podstaty, obsahu, úrovně a postavení studií, která byla uskutečněna a úspěšně dokončena držitelem kvalifikace, ke které je tento dodatek připojen. Dodatek nemá obsahovat žádné ocenění, prohlášení o rovnocennosti nebo doporučení k uznání. Je třeba, aby všech osm částí dodatku bylo vyplněno. Tam, kde informace poskytnuty nebudou, mělo by být uvedeno vysvětlení proč.

This Diploma Supplement follows the model developed by the European Commission, Council of Europe and UNESCO/CEPES. The purpose of the supplement is to provide sufficient independent data to improve the international "transparency" and fair academic and professional recognition of qualifications (diplomas, degrees, certificates etc.). It is designed to provide a description of the nature, level, context, content and status of the studies that were pursued and successfully completed by the individual named on the original qualification to which this supplement is appended. It should be free from any value judgements, equivalence statements or suggestions about recognition. Information in all eight sections should be provided. Where information is not provided, an explanation should give the reason why.

1. Informace o totožnosti držitele kvalifikace / Information identifying the holder of the qualification

1.1 Příjmení, titul / Family name, title: **Koldovský,** Ing.
1.2 Rodné jméno (jména) / First name (names): **Zbyněk**
1.3 Datum narození (den/měsíc/rok) / Date of birth (day/month/year): 25/04/1979
Místo narození / Place of birth: Jablonec Nad Nisou
1.4 Identifikační číslo (kód) studenta / Student identification number or code: 138179

2. Informace o druhu kvalifikace / Information identifying the qualification

2.1 Název kvalifikace a propůjčený titul (v původním jazyce) v plném znění a ve zkratce / Name of the qualification and title conferred (full, abbreviated): doktor - ve zkr. Ph.D. / Doctor - in abbr. Ph.D.
2.2 Hlavní studijní obor(y) v rámci kvalifikace / Main field(s) of study for qualification: Aplikace přírodních věd - Matematické inženýrství / Applied Natural Sciences - Mathematical Engineering
2.3 Název a postavení udělující instituce (v původním jazyce/ v angličtině) / Name and status of awarding institution (in original language/ in English): České vysoké učení technické v Praze, ve zkratce ČVUT v Praze, veřejná vysoká škola / Czech Technical University in Prague, in abbr. CTU in Prague, public higher education institution
2.4 Název a postavení instituce (pokud je jiný, než v bodě 2.3) zajišťující studium/ Name and status of institution providing study (if different from 2.3):
2.5 Studijní jazyk nebo jazyky, ve kterém je realizována výuka a zkoušky / Language(s) of instruction / examination: čeština / Czech

3. Informace o úrovni kvalifikace / Information on the level of the qualification

3.1 Úroveň kvalifikace / Level of qualification: vysokoškolské vzdělání - doktorské studium / higher education - doctoral study programme
3.2 Standardní délka programu / Official length of programme: 6 semestrů / 6 semesters
3.3 Požadavky pro přístup ke studiu / Access requirements: magisterský stupeň a úspěšná přijímací zkouška / master degree and successful entrance exam

4. Informace o obsahu a dosažených výsledcích / Information on the contents and results gained

4.1 Forma studia / Mode of study: prezenční / full-time study
4.2 Požadavky v rámci programu / Programme requirements: předmětové zkoušky, státní doktorská zkouška a disertační práce / subject examinations, state doctoral examination and dissertation

4.3 Podrobné údaje o programu a o jednotlivých dosažených hodnoceních / Programme details and individual grades / ECTS obtained:

Kód předmětu / Code	Název předmětu / Subject	Známka / Exam Mark	Datum zkoušky / Date of exam	Kredity / ECTS
11	Metody MCMC / Markov Chain Monte Carlo methods	1	22.01.2003	
12	Robustní a neparametrické metody / Robust and nonparametric methods	1	05.02.2003	
13	Základy waveletové transformace / Introduction to the wavelet transformations	1	30.04.2003	
14	Lineární problémy s nepřesnými daty / Linear problems with inexact data	1	21.05.2003	
15	Číslicové zpracování signálů / Digital signal processing	1	27.04.2004	
16	Anglický jazyk / English language	2	23.09.2003	
Celkem kreditů / Total ECTS				**0**

Státní doktorská zkouška/ State doctoral examination	Známka / Grade	Datum / Date
Ústní státní doktorská zkouška / Oral state doctoral examination	1	25.11.2004

Disertační práce / Dissertation	Výsledek / Result	Datum / Date
Fast and Accurate Methods for Independent Component Analysis / Fast and Accurate Methods for Independent Component Analysis	obhájil / defended	17.01.2006

4.4 Klasifikační stupnice a vysvětlení jejího významu / Grading scheme and grade distribution guidance:

Výsledek předmětové zkoušky / Result of subject examination

ECTS stupnice / grading scale	A-B	B-C-D-E	FX-F
stupnice ČVUT / CTU grade	1	2	3
slovní hodnocení / verbal assessment	výborně / excellent	prospěl / passed	neprospěl / not passed

Výsledek státní doktorské zkoušky / Result of state doctoral examination

ECTS stupnice / grading scale	A-B	B-C-D-E	FX-F
stupnice ČVUT / CTU grade	1	2	3
slovní hodnocení / verbal assessment	prospěl s vyznamenáním / passed with honours	prospěl / passed	neprospěl / not passed

4.5 Celková klasifikace kvalifikace/ Overall classification of the qualification: prospěl / graduated

5. Informace o funkci kvalifikace / Information on the function of the qualification

5.1 Přístup k dalšímu studiu / Access to further study: nejvyšší stupeň vysokoškolského vzdělání / top level of higher education

5.2 Profesní postavení / Professional status conferred: doktor / Doctor

6. Doplňující informace / Additional information

http://www.cvut.cz / http://www.cvut.cz/en
http://www.msmt.cz, http://www.naric.cz
adresa / address: České vysoké učení technické v Praze
Zikova 4
166 36 Praha 6
Czech Republic

7. Potvrzení dodatku / Certification of the supplement

7.1 Datum / Date: 09.05.2006

7.2 Podpis / Signature: Prof. Ing. Václav Havlíček, CSc.

7.3 Funkce / Capacity: Rektor / Rector

7.4 Oficiální razítko nebo pečeť / Official stamp or seal:

A 000734

8. Informace o národním vysokoškolském systému

Vysokoškolský vzdělávací systém v ČR

Vysoké školy představují nejvyšší úroveň vzdělávacího systému v České republice. Nabízejí akreditované studijní programy tří typů – bakalářské, magisterské a doktorské a také celoživotní vzdělávání. Vysoké školy jsou univerzitního a neuniverzitního typu.

Vysoká škola univerzitní uskutečňuje všechny tři typy studijních programů a v souvislosti s tím vědeckou a výzkumnou, vývojovou, uměleckou nebo další tvůrčí činnost. Vysoká škola neuniverzitní uskutečňuje převážně bakalářské, může též uskutečňovat magisterské studijní programy a v souvislosti s tím výzkumnou, vývojovou, uměleckou nebo další tvůrčí činnost. Vysoká škola neuniverzitní se nečlení na fakulty.

Vysoké školy jsou veřejné, státní a soukromé. Veřejné a soukromé vysoké školy spadají do působnosti Ministerstva školství mládeže a tělovýchovy, státní instituce (vojenské vysoké školy a policejní akademie) spadají do působnosti Ministerstva obrany a Ministerstva vnitra. Seznamy vysokých škol jsou k dispozici na: http://www.msmt.cz/files/htm/Vswwwser1.htm a na www.csvs.cz.

Vysokoškolské vzdělání se získává studiem v rámci akreditovaného studijního programu podle studijního plánu stanovenou formou studia. Forma studia může být prezenční, distanční nebo jejich kombinace.

Podmínkou přijetí ke studiu v bakalářském a magisterském studijním programu je dosažení úplného středního nebo úplného středního odborného vzdělání. Ke studiu v oblasti umění mohou být přijati též uchazeči s vyšším odborným vzděláním poskytovaným na konzervatořích. Podmínkou přijetí ke studiu v magisterském studijním programu, který navazuje na bakalářský studijní program, je rovněž řádné ukončení studia v bakalářském studijním programu. Podmínkou přijetí ke studiu v doktorském studijním programu je řádné ukončení studia v magisterském studijním programu a v oblasti umění též získání akademického titulu. Vysoká škola nebo fakulta může stanovit další podmínky přijetí ke studiu týkající se určitých znalostí, schopností nebo nadání apod.

Vysokoškolské kvalifikace

Kvalifikační struktura rozlišuje tři typy studijních programů – bakalářský studijní program, magisterský studijní program a doktorský studijní program.

Bakalářský studijní program je zaměřen na přípravu k výkonu povolání nebo ke studiu magisterského studijního programu. Standardní doba studia je 3 až 4 roky (180-240 ECTS kreditů). Studium se řádně ukončuje státní závěrečnou zkouškou, jejíž součástí je zpravidla obhajoba bakalářské práce. Absolventi získají akademický titul

- v oblasti umění "bakalář umění" ("BcA."),
- v ostatních oblastech "bakalář" ("Bc.").

Tituly se uvádějí před jménem.

Magisterský studijní program navazuje na bakalářský studijní program. Standardní doba studia je 1 až 3 roky (60-180 ECTS kreditů). V případech, kdy to vyžaduje charakter studijního programu, může být udělena akreditace magisterskému studijnímu programu, který nenavazuje na bakalářský studijní program. Standardní doba studia je potom 4–6 let (240–360 ECTS kreditů). Studium se řádně ukončuje státní závěrečnou zkouškou, jejíž součástí je obhajoba diplomové práce. V oblasti lékařství, zubního lékařství a veterinárního lékařství a hygieny se studium řádně ukončuje státní rigorózní zkouškou.
Absolventům se udělují tyto akademické tituly:

- v oblasti ekonomie, technických věd a technologií, zemědělství, lesnictví a vojenství "inženýr" (ve zkratce "Ing"),
- v oblasti architektury "inženýr architekt" (ve zkratce "Ing. arch."),
- v oblasti lékařství "doktor medicíny" (ve zkratce "MUDr."),
- v oblasti zubního lékařství "zubní lékař" (ve zkratce "MDDr."),
- v oblasti veterinárního lékařství a hygieny "doktor veterinární medicíny" (ve zkratce "MVDr."),
- v oblasti umění "magistr umění" (ve zkratce "MgA."),
- v ostatních oblastech "magistr" (ve zkratce "Mgr.").

Absolventi magisterských studijních programů, kteří získali akademický titul "magistr", mohou vykonat v téže oblasti studia státní rigorózní zkoušku, jejíž součástí je obhajoba rigorózní práce. Po jejím vykonání se udělují tyto akademické tituly:

- v oblasti práva "doktor práv" (ve zkratce "JUDr."),
- v oblasti humanitních, pedagogických a společenských věd "doktor filozofie" (ve zkratce "PhDr."),
- v oblasti přírodních věd "doktor přírodních věd" (ve zkratce "RNDr."),
- v oblasti farmacie "doktor farmacie" (ve zkratce "PharmDr."),
- v oblasti teologie "licenciát teologie" (ve zkratce "ThLic.") nebo "doktor teologie" (ve zkratce "ThDr."), pro oblast katolické teologie "licenciát teologie".

Všechny dosud zmíněné tituly se uvádějí před jménem.

Standardní doba **doktorského studijního programu** je tři roky. Doktorské studium se řádně ukončuje státní doktorskou zkouškou a obhajobou disertační práce. Absolventům doktorského studijního programu se přiznává akademický titul:

- "doktor" (ve zkratce "Ph.D."),
- "doktor teologie" (ve zkratce "Th.D.") v oblasti teologie.

Zkratky těchto titulů se uvádějí za jménem.

Akademický rok

Akademický rok začíná obvykle 1. září a končí 31. srpna následujícího roku. Sestává ze zimního a letního semestru. Organizace akademického roku je určena statutem vysoké školy.

Kreditní systém

Kreditní systém (ECTS nebo srovnatelný systém) je užíván většinou vysokých škol.

Information on national higher education system

Higher Education System in the Czech Republic
Higher Education Institutions

Higher education institutions form the highest level of Czech education. They offer accredited study programmes at three levels - bachelor, master, and doctoral, as well as lifelong learning. Higher education institutions are either university-type or non university-type.

University-type higher education institutions may offer all three types of study programmes (bachelor, master and doctoral) and carry out associated scholarly, research, developmental, artistic or other creative activities. Non university-type higher education institutions offer mainly bachelor study programmes, but may also provide master study programmes and carry out associated scholarly, research, developmental, artistic or other creative activities. Non university-type higher education institutions are not organised to faculties.

There are public, state and private higher education institutions. The public and private higher education institutions come under the responsibility of the Ministry of Education, Youth and Sports, while state institutions (military universities and the police academy) are under the responsibility of the Ministry of Defence and the Ministry of the Interior. The list is available at: http://www.msmt.cz/files/htm/Vswwwser1.htm or www.csvs.cz.

Higher education is realised within the framework of accredited study programmes in accordance with their curriculum and given form of studies. The form of study can be full-time, part-time (distance) or a combination of both.

Access to a bachelor study programme is conditional on completing a full secondary general or vocational education with a GSCE - "maturita" examination and a "maturita" school-leaving general certificate of education (maturitní vysvědčení). In respekt to studies of arts also applicants with completed academy education at the conservatoires are eligible for enrolment. Admission for the master study programme as continuation of the bachelor study is conditioned also by regular completion, i.e. graduation at the bachelor study programme level, while admission for doctoral study programme is condiotional on graduating from a master study programme, and in sphere of arts also award of degrese. University or college may determine another conditions for admission for study in respekt to certain knowledge, skills, talent or endowment etc.

Higher Education Qualifications

The qualification structure recognises bachelor, master and doctoral levels.

A **bachelor study programme** aims at qualifying to enter a profession or a master study programme. It takes from 3 to 4 years (180-240 ECTS credits). The study programme must be completed in due form with a final state examination, which usually includes the presentation and defence of a bachelor thesis. Graduates receive the academic degree

- "bakalář umění" ("BcA.") - Bachelor of Arts in the field of art,
- "bakalář" ("Bc.") – Bachelor in other fields.

Degrees are denoted in front of the name.

A **master study programme** follows a bachelor study programme. The length is 1–3 years (60–180 ECTS credits). In selected fields, where the nature of the study programme so requires, a master study programme need not follow on from a bachelor programme. In this case, the programme lasts 4–6 years (240–360 ECTS credits). Graduates in a master study programme have to take a final state examination and publicly present and defend a master thesis. Studies in the field of medicine, veterinary medicine and hygiene are completed by passing a rigorous state examination including the presentation and defence of a rigorous thesis.
The studies in a master programme graduate with academic degrees classified as follows:

- "inženýr" (abbreviated "Ing".) – Engineer in the field of economics, technical sciences and technologies, agriculture, forestry and in military fields of study;
- "inženýr architekt" ("Ing. arch.") – Engineer Architect in the field of architecture;
- "doktor medicíny" ("MUDr.") – Doctor of Medicine in the field of medicine;
- "zubní lékař" ("MDDr.") – Doctor of Dental Medicine in the field of dental medicine;
- "doktor veterinární medicíny" ("MVDr.") – Doctor of Veterinary Medicine in the field of veterinary medicine and hygiene.
- "magistr umění" ("MgA.") – Master of Arts in the field of art;
- "magistr" ("Mgr.") – Master in humanities, sciences, law, teacher education, pharmacy and theology;

[NB: Courses in human medicine and veterinary medicine last 6 years. They are run by the university medical faculties and faculties of veterinary medicine.]

Graduates - holders of the academic degree "magistr/ Master" can pass a rigorous state examination in the same field and present and defend a rigorous thesis. If completed successfully the following academic degrees are awarded:

- "doktor práv" ("JUDr.") in the field of law;
- "doktor filozofie" ("PhDr.") in the field of humanities, teacher education and social sciences;
- "doktor přírodních věd" ("RNDr.") in the field of natural sciences;
- "doktor farmacie" ("PharmDr.") in the field of pharmacy;
- "licenciát teologie" ("ThLic.") in the field of Catholic theology; "licenciát teologie" ("ThLic.") or "doktor teologie" ("ThDr.") in the field of theology except Catholic theology.

All above mentioned academic titles are at master level and are written in front of the name.

The standard length of a **doctoral study programme** lasts for 3 years. Doctoral studies are completed by the state doctoral examination and the public presentation and defence of a doctoral thesis (dissertation), based on original results, which must be published. Graduates of a doctoral study programme are awarded the academic degree of:

- "doktor" ("Ph.D.")
- "doktor teologie" ("Th.D.") in the field of theology.

These academic titles in abbreviated form are written behind the name.

Academic year

The academic year starts on September 1 and ends on August 31 of the following year. It comprises a winter semester and a summer semester. The organization of academic years is regulated by statute.

Credit system

A credit system (ECTS or a comparable system) is used at most of the higher education institutions.

Uchazeč: Ing. Zbyněk Koldovský, Ph.D. Pracoviště: ITE, FM, TUL

Podpis:

Pedagogická praxe uchazeče na FM, TU Liberec, v letech 2006 až 2010

Vedení přednášek v řádném studiu:

Biologické a akustické signály (BSI) – garant předmětu, zavedení předmětu, příprava a vedení přednášek i cvičení; pro pátý ročník v posledním semestru magisterského studia na FM (10 týdnů), odučeny 2 semestry: LS 2009, LS 2010.

Výpočty, simulace a vizualizace Matlab (MATLB) – garant předmětu, zavedení předmětu, příprava a vedení přednášek i cvičení; pro druhý semestr bakalářského studia na FM (14 týdnů), odučeny 2 semestry: LS 2009, LS 2010.

Počítačové zpracování signálů (PZS) – příprava a vedení poloviny přednášek a cvičení; pro dvouleté navazující magisterské studium na FM, odučeny 2 semestry: LS 2007, LS 2008.

Digital Signal Processing 1 (DSP1) – garant předmětu, zavedení předmětu, příprava a vedení přednášek i cvičení; česko-německý magisterský studijní program, vyučování v angličtině, plánováno na LS 2011.

Vedení cvičení v řádném studiu:

Signály a informace (SGI) – vedení cvičení (minimálně jeden kruh); pro tříleté bakalářské studium na FM, ZS 2006-2010.

Vedení dokončených studentských projektů:

Martin Marek: Odstraňování artefaktů z EEG záznamů pomocí metod slepé separace, 2007.
Jakub Petkov: Advanced Control of Blind Source Separation Algorithms in Matlab (BSSGUI), 2008.
Daniel Hančil: Databáze zvuků pro slepou separaci, 2008.
Petr Schovanec a Jan Pišta: Vyhodnocování hlučnosti zvukového signálu, 2009.
Radek Hlávka: Realizace audioefektů v C#, 2010.
Augustin Bernard: Detekce základní frekvence v akustickém signálu, 2010.
David Botka: Návrh a realizace audio efektů, 2010.
celkem 7 projektů

Vedení obhájených diplomových prací:

Filip Žižka: Aplikace metod pro slepou separaci dat ve zpracování EKG záznamů, 2007.
Michal Kuna: Slepá separace řeči ze stereofonního záznamu, 2007.
Luděk Sviták: Zpracování EEG záznamů s použitím nejnovějších technik pro slepou separaci, 2009.
celkem 3 diplomové práce

Vedení obhájených bakalářských prací:

Daniel Hančil: Jednoduchá lokalizace zvukového zdroje, 2007.
Jakub Petkov: Interaktivní ovládání algoritmů pro slepou separaci signálů v Matlabu, 2007.
Martin Češka: Slepá separace akustických zdrojů pomocí lokalizace, 2009.
celkem 3 bakalářské práce

Vedení doktorandů:

Ing. Jiří Málek, zahájení studia 1.8.2006, studijní obor: Technická kybernetika, disertační práce má název *Blind Audio Source Separation via Independent Component Analysis,* odevzdána v listopadu-prosinci 2010.

Výuka v doktorském studiu

Garant předmětu: Metody digitálního zpracování vícerozměrných signálů, tento předmět do roku 2010 absolvovali: Ing. Lukáš Zedek, Ing. Miloš Kodejška

Pedagogická praxe uchazeče na FJFI, ČVUT v Praze, v letech 2000 až 2005

Vedení přednášek v řádném studiu:

Základy matematické statistiky – zavedení předmětu, příprava a vedení přednášek i cvičení; pro druhý ročník bakalářského studia na FJFI, odučeny 2 semestry: ZS 2003, ZS 2004.

Vedení cvičení v řádném studiu:

Matematické hry – zavedení předmětu, příprava a vedení cvičení; pro první ročník bakalářského studia na FJFI, odučen 1 semestr: ZS 2002.

Lineární algebra A1 a A2 – vedení cvičení; pro první ročník magisterského studia na FJFI, ZS a LS 2000-2005.

Vybrané partie z matematiky – vedení cvičení; pro druhý ročník magisterského studia na FJFI, LS 2004 a LS 2005.

Seznam vědeckých a odborných prací

Uchazeč: Ing. Zbyněk Koldovský, Ph.D. Pracoviště: ITE, FM, TUL

Podpis:

Publikační činnost za období 2005 až 2010

Publikace v mezinárodních impaktovaných časopisech

[1] P. Tichavský and Z. Koldovský, "Fast and accurate methods of independent component analysis: A Survey, " accepted for publication in *Kybernetika*, 2010.

[2] P. Tichavský and Z. Koldovský, "Weight Adjusted Tensor Method for Blind Separation of Underdetermined Mixtures of Nonstationary Sources, " *IEEE Trans. on Signal Processing*, Vol. 59, No. 3, pp. 1037-1047, ISSN:1053-587X, March 2011.

[3] Z. Koldovský and P. Tichavský, "Time-Domain Blind Separation of Audio Sources on the basis of a Complete ICA Decomposition of an Observation Space", *IEEE Trans. on Speech, Audio and Language Processing*, Vol. 19, No. 2, pp. 406-416, ISSN 1558-7916, February 2011.

[4] Z. Koldovský, J. Málek, P. Tichavský, Y. Deville, and S. Hosseini, "Blind Separation of Piecewise Stationary NonGaussian Sources", *Signal Processing*, Volume 89, Issue 12, Pages 2570-2584, ISSN 0165-1684, December 2009.

[5] P. Tichavský, Z. Koldovský, and E. Oja, "Corrections to 'Performance Analysis of the FastICA Algorithm and Cramer-Rao Bounds for Linear Independent Component Analysis' TSP 04/06," *IEEE Tr. Signal Processing*, Vol. 56, No.4, pp. 1715-1716, ISSN: 1053-587X, April 2008.

[6] P. Tichavský, Z. Koldovský, A. Yeredor, G. G. Herrero, and E. Doron, "A Hybrid Technique for Blind Separation of Non-Gaussian and Time-Correlated Sources Using a Multicomponent Approach", *IEEE Trans. on Neural Networks*, Vol. 19, No. 3, pp. 421-430, ISSN: 1045-9227, March 2008.

[7] Z. Koldovský, P. Tichavský and E. Oja, "Efficient Variant of Algorithm FastICA for Independent Component Analysis Attaining the Cramér-Rao Lower Bound", *IEEE Trans. on Neural Networks*, Vol. 17, No. 5, Sept 2006.

[8] P. Tichavský, Z. Koldovský and E. Oja, "Performance Analysis of the FastICA Algorithm and Cramér-Rao Bounds for Linear Independent Component Analysis", *IEEE Trans. on Signal Processing*, Vol. 54, No.4, April 2006.

[9] P. Tichavský, Z. Koldovský, "Optimal Pairing of Signal Components Separated by Blind Techniques", *IEEE Signal Processing Letters*, Vol. 11, No. 2, pp. 119-122, Feb 2004.

Publikace v recenzovaných sbornících mezinárodních konferencí

[10] Z. Koldovský, P. Tichavský, and Anh Huy Phan, "Stability Analysis and Fast Damped-Gauss-Newton Algorithm for INDSCAL Tensor Decomposition," *IEEE Workshop on Statistical Signal Processing*, Nice, France, June 2011.

[11] P. Tichavský and Z. Koldovský, "Stability of CANDECOMP-PARAFAC tensor decomposition," *ICASSP 2011*, Prague, Czech Republic, May 2011.

[12] Z. Koldovský, P. Tichavský, and J. Málek, "Time-Domain Blind Audio Source Separation Method Producing Separating Filters of Generalized Feedforward Structure," in *Latent Variable Analysis and Signal Separation*, Lecture Notes in Computer Science Vol. 6365, pp. 17-24, ISBN: 978-3-642-15994-7, Springer, Heidelberg, Sept. 2010.

[13] Z. Koldovský, P. Tichavský, and J. Málek, "Subband Blind Audio Source Separation Using a Time-Domain Algorithm and Tree-Structured QMF Filter Bank," in *Latent Variable Analysis and Signal Separation*, Lecture Notes in Computer Science Vol. 6365, pp. 25-32, ISBN: 978-3-642-15994-7, Springer, Heidelberg, Sept. 2010.

[14] J. Málek, Z. Koldovský, and P. Tichavský, "Adaptive Time-Domain Blind Separation of Speech Signals," in *Latent Variable Analysis and Signal Separation*, Lecture Notes in Computer Science Vol. 6365, pp. 9-16, ISBN: 978-3-642-15994-7, Springer, Heidelberg, Sept. 2010.

[15] P. Tichavský and Z. Koldovský, "Simultaneous Search for All Modes in Multilinear Models," *ICASSP 2010*, pp. 4114-4117, ISBN: 978-1-4244-4296-6, ISSN: 1520-6149, Dallas, USA, March 2010.

[16] Z. Koldovský and P. Tichavský, "A Comparison of Independent Component and Independent Subspace Analysis Algorithms," *EUSIPCO 2009* , pp. 1447-1451, Glasgow, Scotland, August 24-28, 2009.

[17] P. Tichavský, A. Yeredor, and Z. Koldovský, "A Fast Asymptotically Efficient Algorithm for Blind Separation of a Linear Mixture of Block-Wise Stationary Autoregressive Processes," *ICASSP 2009*, pp. 3133-3136, ISBN: 978-1-4244-2354-5, ISSN: 1520-6149, Taipei, Taiwan, April 2009.

[18] J. Petkov and Z. Koldovský, "BSSGUI – A Package for Interactive Control of Blind Source Separation Algorithms in MATLAB," in *Cross-Modal Analysis of Speech, Gestures, Gaze and Facial Expressions* (Eds.: A. Esposito and R. Vích), pp. 386-398, ISBN: 978-3-642-03319-3, ISSN: 0302-9743, Springer Berlin / Heidelberg, July 2009.

[19] J. Málek, Z. Koldovský, J. Žďánský and J. Nouza, "Enhancement of Noisy Speech Recordings via Blind Source Separation", *Proceedings of the 9th Annual Conference of the International Speech Communication Association (Interspeech 2008)*, pp. 159-162, ISSN: 1990-9772, September 22-26, Brisbane, Australia, 2008.

[20] Z. Koldovský and P. Tichavský, "Time-domain Blind Audio Source Separation Using Advanced Component Clustering and Reconstruction", *Proc. of The Joint Workshop on Hands-free Speech Communication and Microphone Arrays (HSCMA 2008)*, pp. 216-219, ISBN: 978-1-4244-2338-5, May 6-8, Trento, Italy, 2008.

[21] Z. Koldovský, J. Málek, P. Tichavský, Y. Deville, and S. Hosseini, "Extension of EFICA Algorithm for Blind Separation of Piecewise Stationary Non Gaussian Sources", *33rd International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2008)*, Las Vegas, Nevada, pp. 1913-1916, ISBN: 1-4244-1484-9, April 2008.

[22] Z. Koldovský and P. Tichavský, "Time-Domain Blind Audio Source Separation Using Advanced ICA Methods", *Proceedings of the 8th Annual Conference of the International Speech Communication Association (Interspeech 2007)*, pp. 846-849, August 2007.

[23] P. Tichavský, Z. Koldovský and E. Oja, "Speed and Accuracy Enhancement of Linear ICA Techniques Using Rational Nonlinear Functions", *Proceedings of 7th International Conference on Independent Component Analysis (ICA2007)*, pp. 285-292, Sept. 2007.

[24] Z. Koldovský and P. Tichavský, "Blind Instantaneous Noisy Mixture Separation with Best Interference-plus-noise Rejection", *Proceedings of 7th International Conference on Independent Component Analysis (ICA2007)*, pp. 730-737, Sept. 2007.

[25] G. G. Herrero, Z. Koldovský, P. Tichavský, and K. Egiazarian, "A Fast Algorithm for Blind Separation of Non-Gaussian and Time-Correlated Signals", *Proceedings of 15th European Signal Processing Conference (EUSIPCO 2007)*, pp. 1731-1735, Sept 2007.

[26] J. Málek, Z. Koldovský, S. Hosseini, and Y. Deville, "A Variant of EFICA Algorithm with Adaptive Parametric Density Estimator", *8th International Workshop on Electronics, Control, Modelling, Measurement, and Signals (ECMS 2007)*, pp. 79-84, Liberec, Czech Republic, May 2007.

[27] Z. Koldovský, J. Nouza, and J. Kolorenč, "Continuous Time-Frequency Masking Method for Blind Speech Separation with Adaptive Choice of Threshold Parameter Using ICA", *Interspeech 2006*, Pittsburgh PA, USA, 17.-21. September, pp. 2578-2581, 2006.

[28] P. Tichavský, Z. Koldovský, E. Doron, A. Yeredor, and G. G. Herrero, "Blind signal separation by combining two ICA algorithms: HOS-based EFICA and time structure-based WASOBI", *Proceedings of The 2006 European Signal Processing Conference (EUSIPCO'2006)*, Florence, Sep. 2006.

[29] Z. Koldovský, P. Tichavský, "Methods of Fair Comparison of Performance of Linear ICA Techniques in Presence of Additive Noise", *Proc. of ICASSP 2006*, Toulouse, no. V., pp. 873-876, May 2006.

[30] Z. Koldovský, P. Tichavský, "Efficient Variant of Algorithm FastICA for Independent Component Analysis Attaining the Cramér-Rao Lower Bound", *Proc. of SSP-2005*, Bordeaux, July 2005.

[31] P. Tichavský, Z. Koldovský and E. Oja, "Asymptotic Performance Analysis of the Fixed-Point Algorithm (Fast-ICA) for Independent Component Analysis", *Proc. of SSP-2005*, Bordeaux, July 2005.

[32] Z. Koldovský, P. Tichavský and E. Oja, "Cramér-Rao Lower Bound for Linear Independent Component Analysis., *Proc. of ICASSP-2005*, Philadelphia, Vol. III, pp. 581 - 584, March 2005.

Kapitola v odborné knize:

[33] J. Nouza a kol.: Řeč a počítač: principy hlasové komunikace, úlohy, metody a aplikace, vydavatel: Technická univerzita v Liberci, editoři: Jan Nouza, Zbyněk Koldovský, Robert Vích, číslo publikace: 55-110-09, první vydání, počet stran: 238, ISBN 978-80-7372-548-8, kapitola „Zpracování záznamu řeči z více mikrofonů metodami pro řešení inverzního problému", str. 201-216, 2009.

Články jsou s výjimkou [16], [25-26] a [28] zahrnuty v mezinárodní databázi ISI nebo SCOPUS.

Citace:

Po vyloučení autocitací a citací od spoluautorů zhruba 69 citací v databázi ISI a 112 citací v databázi Scopus (březen 2011).

Ocenění:

Ocenění publikace [7] jako nejlepší publikace roku 2006, ÚTIA AV ČR v.v.i.

Uchazeč: Ing. Zbyněk Koldovský, Ph.D. Pracoviště: ITE

Podpis:

Zahraniční návštěvy

- červen 2004, Návštěva Neural Networks Research Centre na Technické univerzitě v Helsinkách v rámci výměnných meziakademických dohod s přednáškou na téma „Asymptotic Analysis of the Symmetric Fast-ICA and Cramer-Rao Lower Bound for Independent Component Analysis."
- březen 2005, Návštěva Neural Networks Research Centre na Technické univerzitě v Helsinkách s přednáškou na téma „Efficient Version of Algorithm FastICA Attaining the Cramer-Rao Lower Bound," seminář Laboratory of Computer and Information Science.
- listopad a prosinec 2009, návštěva oddělení Univerzity ve Stuttgartu (Německo), International doctorate School in Information and Communication Technology of Trento (Itálie) a Laboratoire Astrophysique de Toulouse (Francie).
- prosinec 2010, návštěva Laboratory for Advanced Brain Signal Processing, RIKEN, Tokyo (Japonsko).

Odborné přednášky a vystoupení

- červen 2003, „Optimální párování signálových komponent separovaných slepými metodami," výjezdní seminář ÚTIA AV ČR, Zlenice.
- prosinec 2003, „Slepá dekonvoluce," seminář Katedry matematiky FJFI ČVUT, Praha.
- duben 2006, „Úvod do analýzy nezávislých komponent a jiných metod pro slepou separaci dat," 668. kolokvium teorie obvodů, systémů a signálů, URE AV ČR.
- duben 2008, "Slepá separace promluv současně hovořících osob", seminář Katedry matematiky FJFI ČVUT, Praha.
- září 2008, „Analýza nezávislých komponent a její užití pro slepou separaci akustických signálů," seminář VUT Brno.
- únor 2011, „Slepá separace signálů pomocí rozkladu tenzorů," seminář VUT Brno.