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

Petr Tichavský[1], Zbyněk Koldovský[1], Arie Yeredor[2], Germán Gómez-Herrero[3], and Eran Doron[2]

Submitted to the IEEE Transactions on Neural Networks

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 sub-optimal. 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

I. INTRODUCTION

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.

[1]Institute of Information Theory and Automation, Pod vodárenskou věží 4, P.O.Box 18,182 08 Prague 8, Czech Republic E-mail: tichavsk@utia.cas.cz, zbynek.koldovsky@tul.cz

[2]School of Electrical Engineering, Tel-Aviv University P.O.Box 39040, Tel-Aviv 69978, Israel, E-mail: Eran@amimon.com, arie@eng.tau.ac.il

[3] Institute of Signal Processing, Tampere University of Technology, P.O.Box 553, 33101 TUT, Finland, E-mail: german.gomezherrero@tut.fi

The goal is to estimate the mixing matrix $\mathbf{A}$ or, equivalently, the de-mixing matrix $\mathbf{W} \stackrel{\triangle}{=} \mathbf{A}^{-1}$ 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 re-ordering 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", "non-whiteness" or "non-stationarity" of the source signals[1]. For each of these models there exist algorithms which are asymptotically optimal (in some sense, to be discussed shortly) under certain conditions: 1) Efficient Fast-ICA (EFICA, [18]) for independent white generalized-Gaussian-distributed sources, 2) Weights-Adjusted SOBI (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 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], [27]. Merits in combining these two kinds of methods were already demonstrated on an example with an 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 JADE$_{\text{TD}}$ [23], JCC [12], and Thin ICA (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 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],

---

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

[3], [7]. Note that unlike [1], [3], [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 the next section. 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 de-mixing matrix $\widehat{\mathbf{W}}$, the "realization-ISR" matrix $\mathbf{rISR}$ is given (element-wise) 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 k-th 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 shall 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, it will be shown that

- The EFICA ISR expression, derived assuming temporally-white sources, remains approximately valid when the sources are mildly colored. Likewise:
- 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 the following section.

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 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 element-wise 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

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

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,

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

where $\mathbf{x}[n]$ denotes the n-th 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],

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

---

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

where σ_k^2 is the variance of the innovation sequence of the k-th 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 ℓ-th source with $a_{0\ell} = 1$ for $k, \ell = 1, \ldots, d$, and $R_k[m]$ is the autocorrelation of the k-th source at time lag m (we use a unit-variance scaling assumption $R_k[0] = R_\ell[0] = 1$ in our model).

C. 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 isrs estimates $\widehat{isr}_k^{EF}$ and $\widehat{isr}_k^{WA}$ (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 $\widehat{isr}^{EF}$ and $\widehat{isr}^{WA}$.
3) Let $E = \min_k \widehat{isr}_k^{EF}$ and $W = \min_k \widehat{isr}_k^{WA}$
4) **If** $E < W$,
   a) accept those signals $\mathbf{s}^{EF}$ for which $\widehat{isr}_k^{EF} < W$ and redefine $\mathbf{z}$ as the rejected signals of $\mathbf{s}^{EF}$

   else,

   b) accept those signals $\mathbf{s}^{WA}$ for which $\widehat{isr}_k^{WA} < E$ 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), (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 above, 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 above 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 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 it is essentially unaware of the group-separation ability of the first algorithm (because only the individual sources' isr-s are accounted for).

We therefore propose (in the following section) 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 the Simulation Section.

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 [3], [1]. 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 co-exist.

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 $\widehat{isr}$ of a cluster $\mathcal{I}$ (with respect to all the other

sources) can be defined as

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

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 $isr(\mathcal{I})$ 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 the following subsection for a further discussion.

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

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

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

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

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

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

until either:

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

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 $\mathcal{J} := \{1, \ldots, \dim(\mathbf{z})\} - \cup\mathcal{I}_k$. 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 $\widehat{isr}^{WA}$.

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 the Simulation Section.

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 $\{1, \ldots, \dim(\mathbf{z})\}$. 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, consider 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 most one cluster of non-distinguishable (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're 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 ISR matrices, $\mathbf{D} \overset{\triangle}{=} \widehat{\mathbf{ISR}} + \widehat{\mathbf{ISR}}^T$.
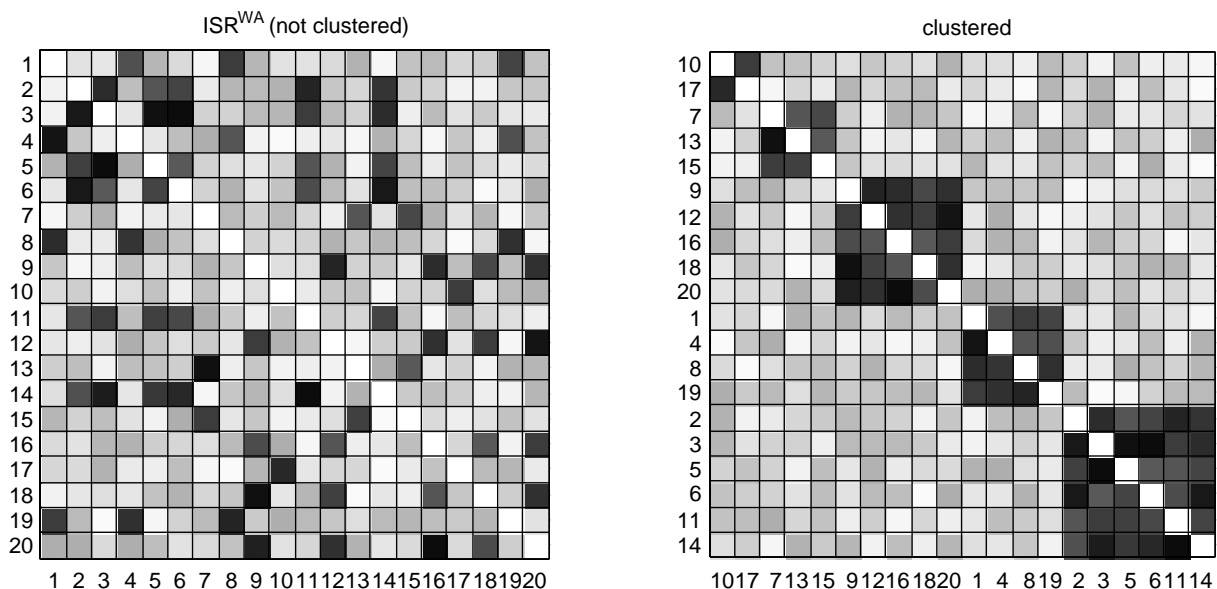
Fig. 1. Example of hierarchial clustering of a WASOBI ISR matrix (in gray-scale). Left: before, Right: after.

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 $\widehat{isr}(\mathcal{I})$ (calculated using (4)). This selection is required on step 5.

The clustering scheme described above 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 Figure 1. On the left-hand side we show the ISR matrix for 20 sources in gray-scale colors, where lighter colors denote low ISR (good separation) and darker colors denote high ISR (poor separation). The resulting re-ordering and partition into clusters is clearly observed on the right-hand side.

---

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

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 non-asymptotic 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.

Example 1 - Figure 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 (BPSK) sequences using all-pole filters. For each $k = 1, 2, \ldots, 5$, the k-th filter was constructed of k poles, located at all k roots of the real-valued parameter $-\rho$. In other words, the filters' autoregression 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 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 are 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 over-optimistic, 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$_{\text{TD}}$ [23] with parameters 0:5, JCC [12] with parameters $p = 5$ and $\tau = [2, 3, 4, 5, 6]$, and Thin ICA (TICA, [6]) with parameters $(d1, d2, d3) = (5, 0, 5)$.

Example 2 - Figure 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

---

[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.
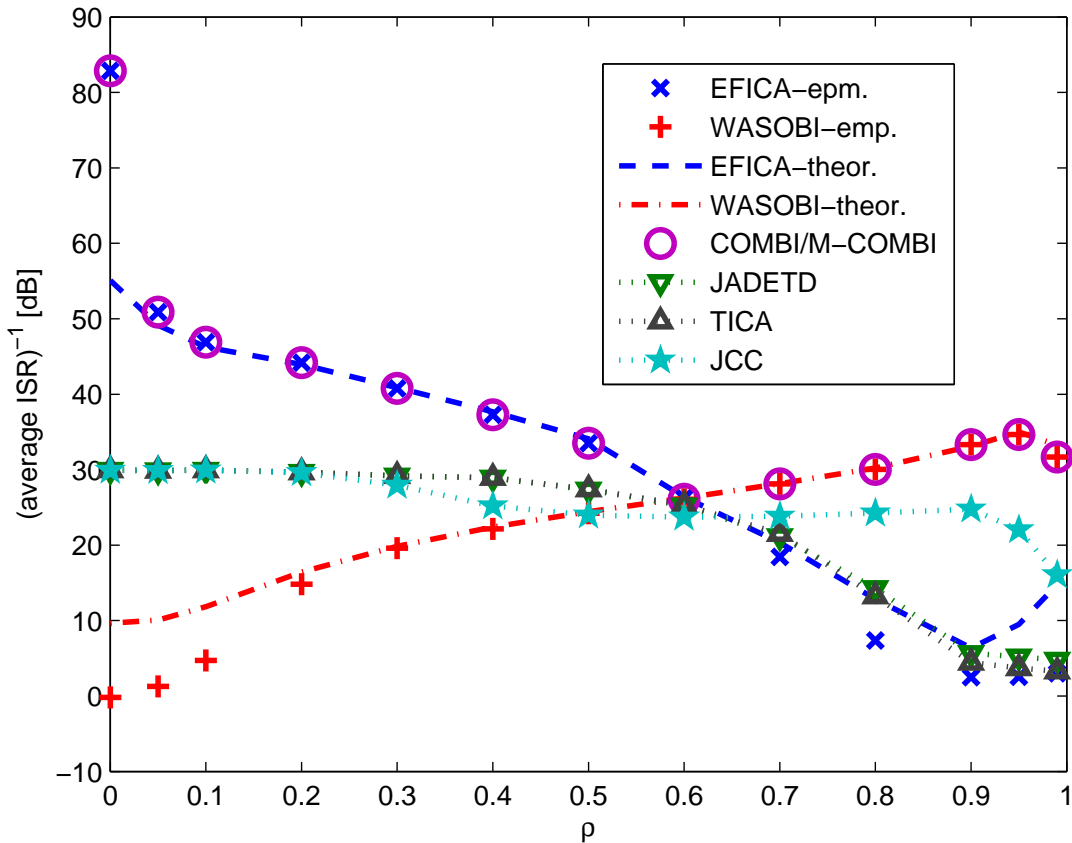
Fig. 2. Inverted average ISR achieved in separation of five AR signals obtained by passing binary (BPSK) i.i.d. sequences of length $N = 1000$ through all-pole filters with autoregression coefficients $[1, \rho]$, $[1, 0, \rho]$, $[1, 0, 0, \rho]$, $[1, 0, 0, 0, \rho]$, and $[1, 0, 0, 0, 0, \rho]$, respectively, 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.

distribution is denoted GG(α). Diagram (a) in Figure 3 shows the result for $\alpha = 0.5$ as a function of parameter ρ and diagram (b) the result for $\rho = 0.5$ and varying α. Each simulation point is an average of 100 trials.

The general behavior in diagram (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.

Diagram (b) of Figure 3 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.

Example 3 - Figure 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$_{\text{TD}}$ [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 iid 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 non-separable by WASOBI, while the second and fourth sources are non-separable 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 dB)

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

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

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,

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

and

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

In this situation, WASOBI cannot accurately resolve individual components but it separates the two two-dimensional 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 uknown type (supergaussian/subgaussian), because each type requires a different built-in nonlinear function. (In this example, we have used "pow3" to achieve a good performance.)
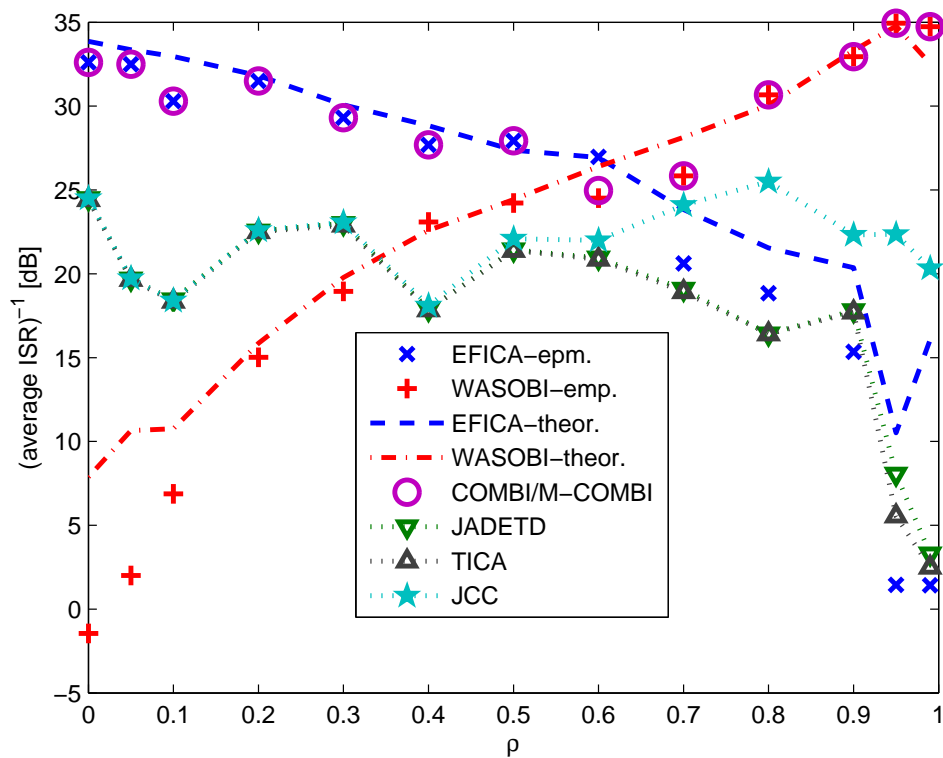
(a)
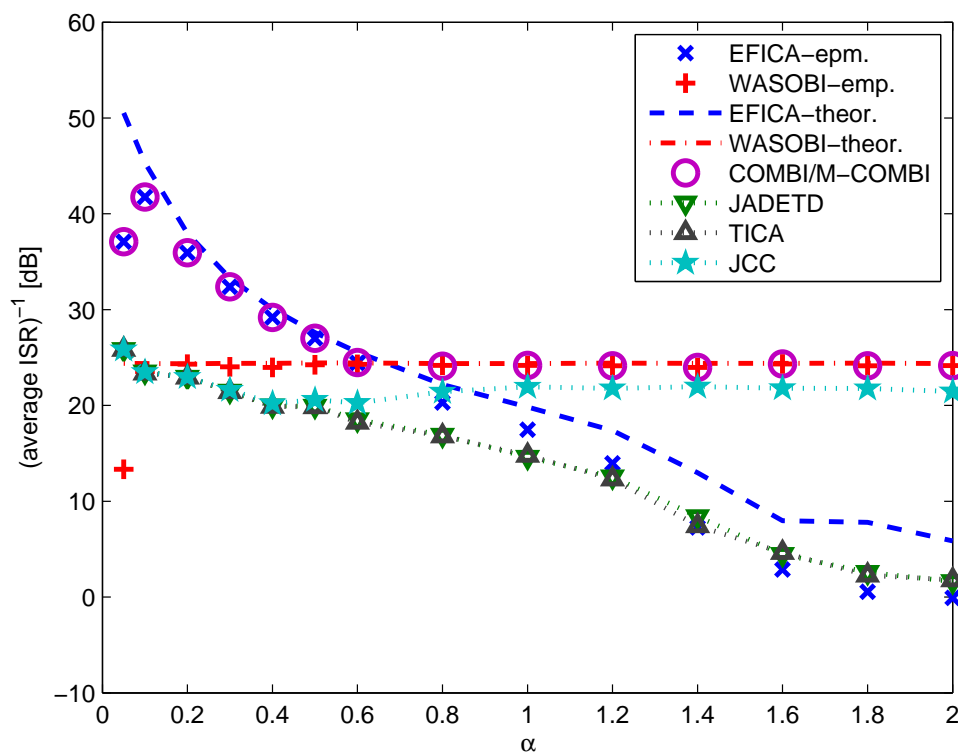
(b)

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 autoregression 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. Diagram (a) shows the result for a fixed $\alpha = 0.5$ versus varying ρ, diagram (b) is for $\rho = 0.5$ versus varying α.
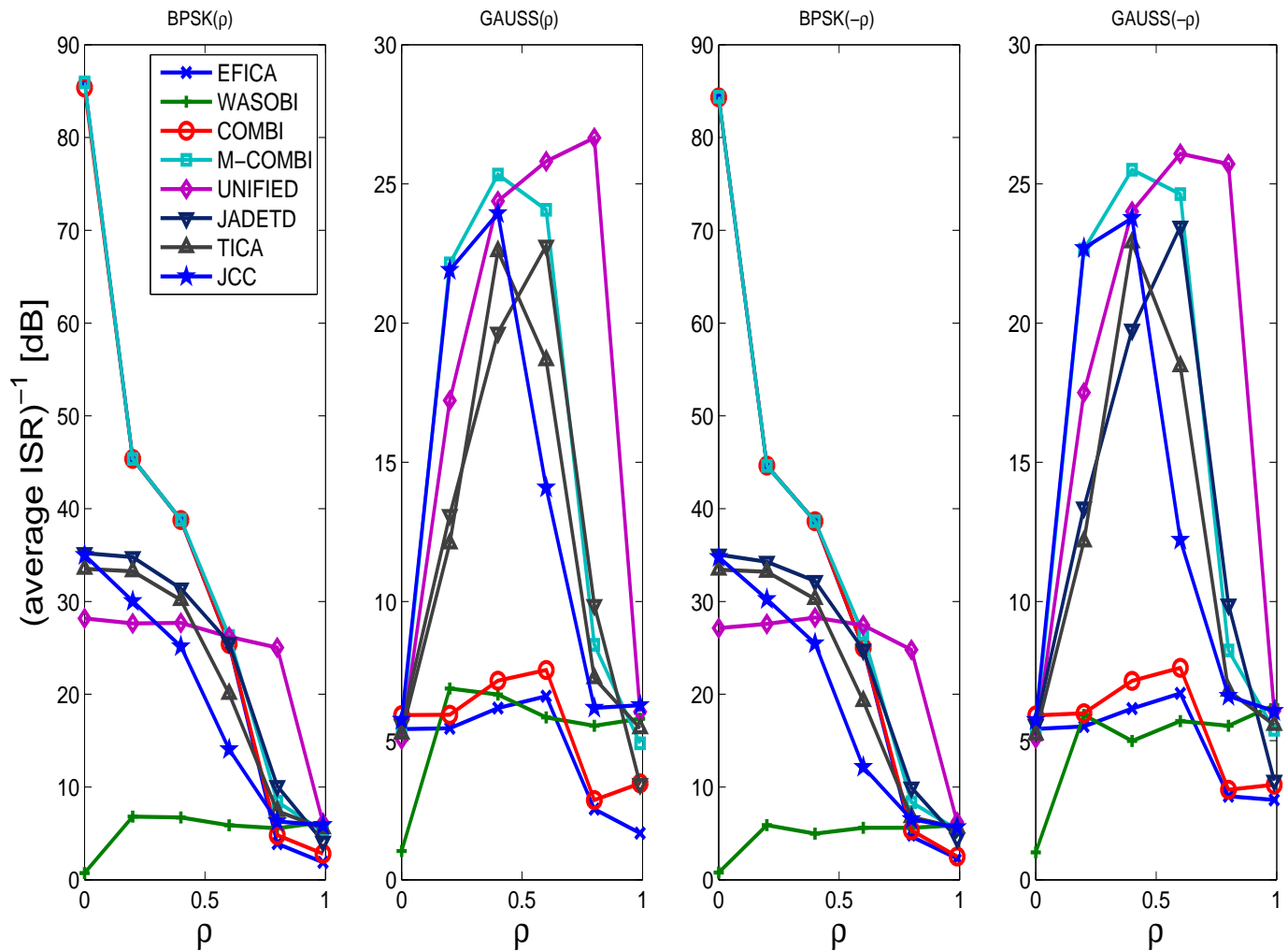
Fig. 4. Inverted average ISR achieved in separation of four AR signals obtained by passing BPSK, Gaussian, BPSK and Gaussian i.i.d. sequences of length $N = 1000$ through all-pole filters whose autoregression coefficients were $[1, \rho]$, $[1, \rho]$, $[1, -\rho]$ and $[1, -\rho]$, respectively. Each simulation point is an average of 100 trials.

Example 4 - Figure 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 was Gaussian for the first group, BPSK for the second, Laplace ($=GG(1)$) for the third and Uniform ($=\lim GG(\alpha)$ for $\alpha \to \infty$) for the fourth. Thus, for EFICA the first group of five Gaussian signals comprise a non-separable 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 - Figure 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 0dB 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+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.

Example 6 - Figure 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 hierarchial 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 PC (P4 3GHz, 2GB RAM, Windows XP) in Matlab® version 7.0/R14 are summarized in Table I.

TABLE I

RUNNING TIMES (IN [S]) OF DIFFERENT SEPARATION ALGORITHMS

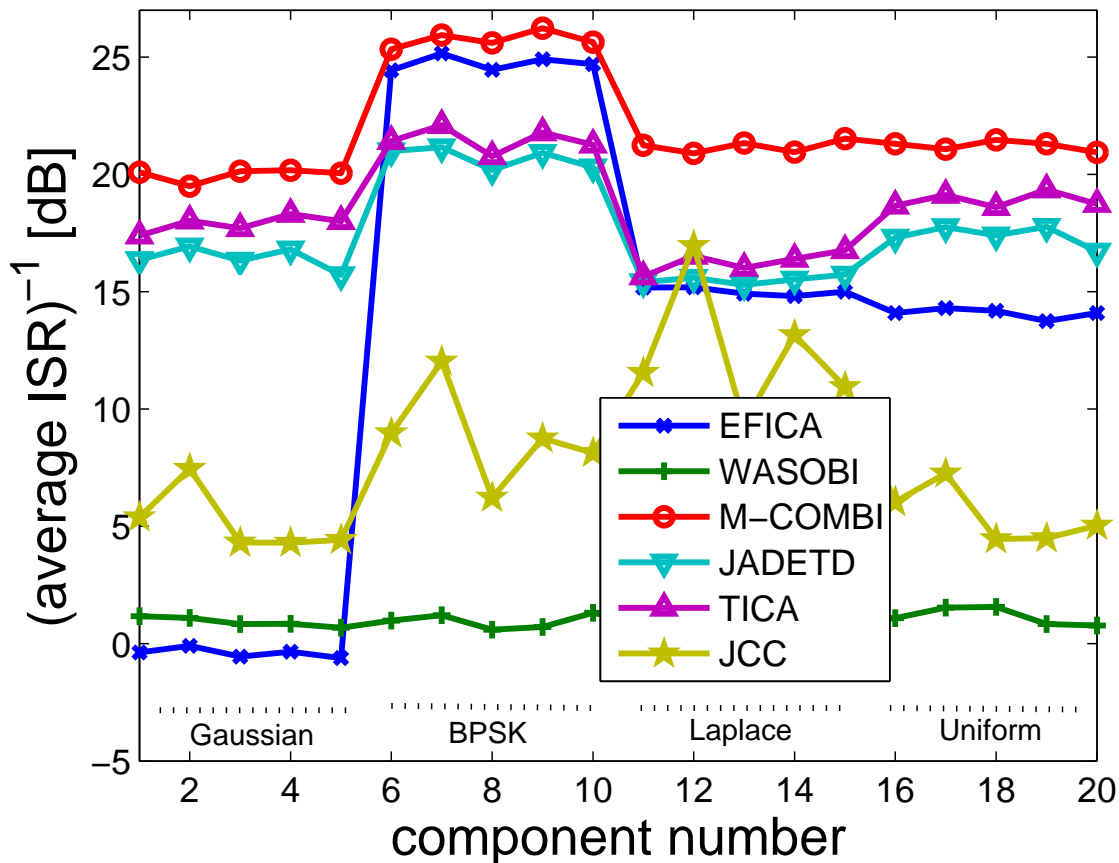| Algorithm	20 sources (ex. 5)	40 sources (ex. 7)
EFICA	2.5	6.7
WASOBI	0.5	2.1
MULTI-COMBI	5.2	22.6
JADE$_{\text{TD}}$	2.6	434.7
TICA	22.6	378.2
JCC	22.4	712.7

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 autoregression 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.

VI. CONCLUSIONS

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 ThinICA, both in terms of separation quality and in terms of computational efficiency.

REFERENCES

[1] F.R. Bach and M.I.Jordan, "Beyond independent components: trees and clusters. *Journal of Machine Learning Research*, vol. 4, pp. 1205-1233, 2003.

[5]The Matlab codes for EFICA, WASOBI, COMBI and MULTI-COMBI can be downloaded from the first author's homepage, http://si.utia.cas.cz/Tichavsky.html. A Matlab implementation of JADE$_{TD}$ is available at http://www.cs.tut.fi/~gomezher/software.htm.
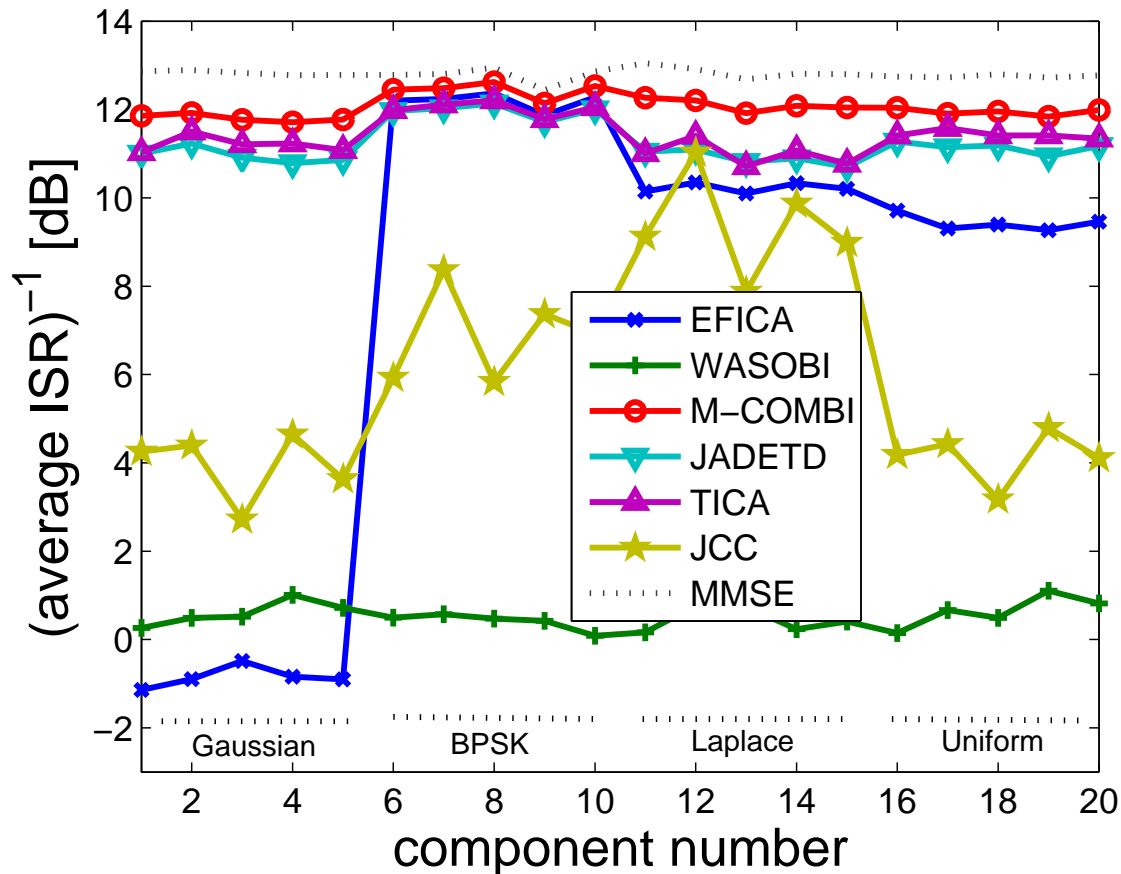
Fig. 6. Inverted average INSR achieved in separation of the same mixtures of 20 AR sources as in Figure 5, contaminated by AWGN at 0dB SNR. "MMSE" denotes the performance of a hypothetical "oracle" separator which uses the known mixing matrix and noises' variances.

[2] 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.

[3] J.-F. Cardoso, "Multidimensional Independent Component Analysis", *Proc. 1998 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP-1998)*, Seattle, USA, 1998.

[4] J.-F. Cardoso, "Blind signal separation: statistical principles," *Proceedings of the IEEE*, vol. 86, no. 10, pp. 2009-2025, 1998.

[5] J.-F. Cardoso, "The three easy routes to independent component analysis; contrasts and geometry," *Proc. ICA 2001, San Diego, 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. on Independent Component Analysis and Blind Signal Separation (ICA 2001)*, San Diego, pp. 740-745, December 2001.

[7] L. De Lathauwer, D. Callaerts, B. De Moor, J. Vandewalle, "Fetal Electrocardiogram Extraction by Source Subspace Separation", *Proc. IEEE Signal Processing / ATHOS Workshop on Higher-Order Statistics*, Girona, Spain, June 12-14, 1995, pp. 134–138.

[8] A. Delorme, S. Makeig et al., "EEGLAB: A software toolbox for electrophysiological data analysis" WWW release: http://sccn.ucsd.edu/eeglab/, The Salk Institute, La Jolla and Swartz Center for Computational Neuroscience, Institute for Neural Computation, University of California San Diego, September, 2002.

[9] E. Doron and A. Yeredor, "Asymptotically Optimal Blind Separation of Parametric Gaussian sources", *Proc. of ICA 2004*, Granada, Sep. 2004.
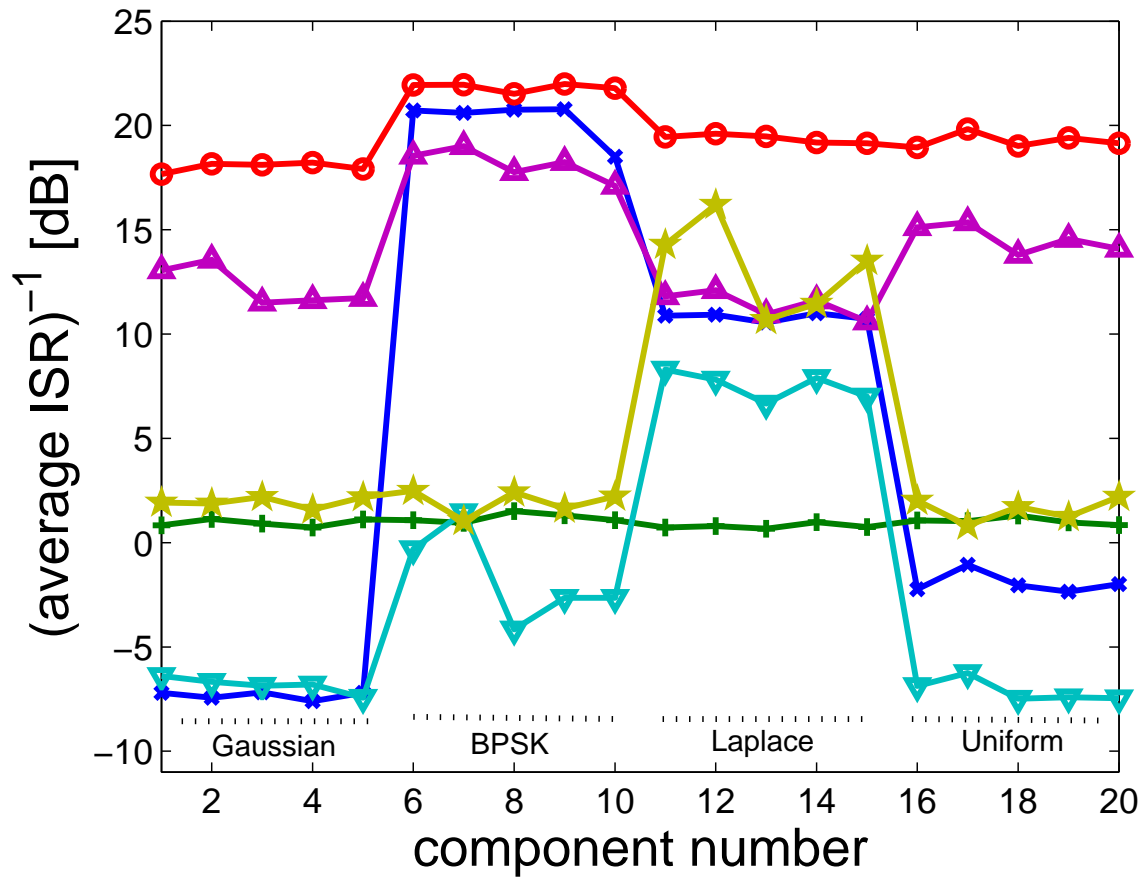
Fig. 7. Inverted average isr (for each component separately) achieved in separation of 40 sources; 20 AR sources of Figure 5 + 20 white Gaussian sources. Since the Gaussian sources are inseparable, they are excluded from the Figure. Each simulation point is an average of 100 trials.

[10] E. Doron, "Asymptotically Optimal Blind Separation of Parametric Gaussian Sources", M.Sc. Thesis, Tel-Aviv University, 2003.

[11] E. Doron, A. Yeredor and P. Tichavský, "Cramér-Rao-Induced Bound for Blind Separation of Stationary Parametric Gaussian Sources", *IEEE Signal Processing Letters*, vol. 14, no.6, pp. 417-420, June 2007.

[12] 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.

[13] A. Hyvärinen, "Fast and robust fixed-point algorithms for independent component analysis", *IEEE Tr. Neural Networks*, vol. 10, no. 3, pp. 626-634, May 1999.

[14] A. Hyvärinen, "Complexity Pursuit: Separating interesting components from time-series", *Neural Computation*, vol. 13, no. 4, pp. 883-898, 2001.

[15] A. Hyvärinen, "A unifying model for blind separation of independent sources", *Signal Processing*, vol. 85, no. 7, pp.1419-1427, 2005.

[16] C.J. James, C.V. Hesse, "Independent component analysis for biomedical signals", *Physiological Measurements*, vol. 26, pp. R15-R39, 2004.

[17] Z. Koldovský, P. Tichavský "Methods of Fair Comparison of Performance of Linear ICA Techniques in Presence of Additive Noise", *Proc. 2006 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP-2006)*, Vol. 5, pp. 873-876, Toulouse, France, May 14-19, 2006.

[18] Z. Koldovský, P. Tichavský and E. Oja, "Efficient Variant of Algorithm FastICA for Independent Component Analysis Attaining the

Cramer-Rao Lower Bound", *IEEE Tr. Neural Networks*, vol. 17, no. 5, pp. 1265- 1277, Sep. 2006.

[19] Mao Ye, "Global convergence analysis of a discrete time nonnegative ICA algorithm", *IEEE Trans. Neural Networks*, 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. on Biomedical Engineering*, vol. 49, no. 12, pp. 1514-1525, Dec. 2002.

[21] E. Oja and Z. Yuan, "The FastICA algorithm revisited: Convergence analysis", *IEEE Trans. on Neural Networks*, vol 17, no.6, pp. 1370-1381, November 2006.

[22] D.-T. Pham and J.-F. Cardoso, "Blind separation of instantaneous mixtures of non stationary sources," *IEEE Trans. on Signal Processing*, vol. 49, no. 9, pp. 1837-1848, 2001.

[23] 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, Jan. 1999.

[24] H. Späth, *Cluster Analysis Algorithms for Data Reduction and Classification of Objects*, Chichester, Ellis Horwood 1980.

[25] 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, 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 Tr. on Signal Processing*, vol. 54, no. 4, pp. 1189–1203, Apr. 2006.

[27] Tichavský P., Nielsen J., Krajča V., "Identification of epileptic activity of brain in EEG recordings using four ICA techniques", *Proc. Biosignal 2006*, Brno, Czech Republic, June 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", *Proc. EUSIPCO-2006*, 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", *Proc. EUSIPCO-2006*, Florence, Sep. 2006.

[30] P. Tichavsky, Z. Koldovsky, and E. Oja, "Speed and accuracy enhancement of linear ICA techniques using rational nonlinear functions", *Proc ICA'2007*, London, UK, September 9-12, 2007.

[31] F. Vrins, J.A. Lee and M. Verleysen, "A minimum-range approach to blind extraction of bounded sources", *IEEE Tr. Neural Networks*, vol. 18, no. 3, pp. 809-822, May 2007.

[32] Lu Wei and J.C. Rajapakse, "Approach and applications of constrained ICA", *IEEE Tr. Neural Networks*, 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 Processing Letters*, vol. 7, pp. 197-200, Jul. 2000.