CLOSED-FORM OPTIMIZED COMPOSITE-ORDER ESTIMATOR FOR BLIND SEPARATION OF INSTANTANEOUS LINEAR MIXTURES

Shafayat Abrar and Asoke K. Nandi

The University of Liverpool
Department of Electrical Engineering & Electronics
Brownlow Hill Street, Liverpool L69 3GJ, UK
phone: +44 (0)151 794 4525, fax: +44 (0)151 794 4540,
email: {shafayat, aknandi}@liverpool.ac.uk
web: http://www.liv.ac.uk/EEE/academicstaff/nandi.htm

ABSTRACT

In the field of blind source separation, formulation of closed-form estimators constitutes an important framework. In this work, we present an optimal composition of third- and fourth-order cumulants leading to a closed-form ICA estimator for a pair-wise blind source separation. We introduce a free weight parameter in combining the cumulants and evaluated its optimal value such that the mean-square estimation of Givens rotation is minimized. We have shown that the optimal value of weight parameter depends on the statistical knowledge of the mixing signals and additive Gaussian noise. Computer simulations have been performed to illustrate the behavior of the proposed optimized closed-form estimator for the maximization of the weighted contrast.

1. INTRODUCTION

1.1 Blind Source Separation: Problem

The problem of blind source separation (BSS) consists of recovering a set of unobserved signals, so-called sources, from another set of observed signals which are mixtures of the sources [1, 2]. The term "blind" signifies that (typically) very few assumptions are made about the sources and the mixing process. The problem of BSS arises in many signal processing applications like communications, array processing, speech analysis and speech recognition [2]. In all these instances, the underlying assumption is that several linear mixtures of unknown, random, zero mean, and statistically independent signals, called sources, are observed; the problem consists of recovering the original sources from their mixtures without a priori information of coefficients of the mixtures and knowledge of the sources. The principle involved in the solution to this problem is called *independent component analysis* (ICA), which can be viewed as an extension of the widely known *principal component analysis* (PCA). Nowadays, the independent component analysis (ICA) has become an active field of research that has attracted great interest because of its large number of applications in diverse fields.

1.2 Closed-Form Solutions

In the fundamental real-valued two-source scenario, the problem reduces to the identification of a single parameter, the unknown angle characterizing the Givens-rotation mixing matrix. A variety of closed-form methods for the estimation of this angle have been proposed in the literature [3, 4, 5, 6, 7, 8, 9, 10, 11]. These estimators consist of simple formulas involving straightforward operations on certain statistics of the whitened sensor output. Most of these share the common feature of being based on the fourth-order statistics of the whitened sensor output.

In an M-dimensional case, $M > 2$, ICA can be carried out by applying the two-signal estimators to each whitened signal pair over several sweeps until convergence [12]. This iterative approach is reminiscent of the Jacobi optimization technique for matrix diagonalization. It works through a sequence of sweeps on the whitened data until a given orthogonal contrast is optimized; sweep is defined to be a one complete pass through all the $M(M-1)/2$ possible pairs of distinct indices. In simple words, the Jacobi- iteration spans the whole set of rotation matrices in a sequential manner. The updating step on a pair, say (a,b), partially undoes the effect of previous optimizations on pairs containing either a or b [13]. For this reason, it is necessary to go through several sweeps before optimization is completed.

1.3 Motivation and Contribution

It is interesting to notice that all of the aforesaid closed-form estimators were based on fourth-order statistics. Notice that asymmetric sources arise in many practical scenarios, such as in sonar signal processing [14] or source separation of urban images [15] (see also [16]). In some cases, digitized speech signals have non-zero skewness; and separation of such signals gets benefit from third-order statistics [17]. Also, in biomedical applications, skewness is sometime more important to just non-Gaussianity for certain categories of signals, say, certain artifacts (like eye-blinking) and, some known components in electrocardiograms and electroencephalograms are not symmetric.

Due to the importance of asymmetry, we present a weighted form of third- and the fourth-order contrast (3) which is capable of handling the symmetric and asymmetric sources jointly in an *optimal* manner. Extending the results in our previous work [18], in this work, we obtain a closed-form ICA estimator in the presence of additive noise. The proposed weighted contrast, the closed-form estimator and the derivation of optimal weight parameter is described in Sections **2-3**, and the performance comparisons are provided in Section **4**. We conclude briefly in Section **4**. All simulations were done with MATLAB; analytical calculations in Section **2-3** were supported by Symbolic toolbox of MATLAB.

2. PROPOSED WORK

2.1 System Model

Consider an M-input M-output memoryless channel described by

$$\mathbf{x}(n) = \mathbf{A}\mathbf{s}(n)$$

where $n \in \mathbb{Z}$ is the discrete time, $\mathbf{x}(n)$ is an $M \times 1$ vector of the observed signals, $\mathbf{s}(n)$ is an $M \times 1$ vector of the (original) statistically independent sources, and $\mathbf{A} \in \mathbb{R}^{M\times M}$ is an unknown (invertible) mixing matrix. The goal of blind source separation (BSS) is to determine a separation matrix $\mathbf{B} \in \mathbb{R}^{M\times M}$ such that

$$\mathbf{y}(n) = \mathbf{B}\mathbf{x}(n) = \mathbf{B}\mathbf{A}\mathbf{s}(n) = \mathbf{C}\mathbf{s}(n)$$

recovers the source signal up to a permutation and scaling, where $\mathbf{C}$ ia a global matrix representing a mixing-nonmixing structure. Source separation is typically carried out in two-step. First, whitening or standardization projects the observed vector $\mathbf{x}(n)$ on the signal subspace and yields a set of second-order decorrelated, normalized signals $\mathbf{z}(n) = \mathbf{W}\mathbf{x}(n)$ such that $\mathsf{E}[\mathbf{z}\mathbf{z}^T] = \mathbf{I}_M$. As a result, the source and whitened vectors must be related through a unitary transformation

$$\mathbf{z}(n) = \mathbf{Q}\mathbf{s}(n).$$

The separation problem thus reduces to the computation of unitary matrix $\mathbf{Q}$, which is accomplished in a second step. The ICA approach to BSS consists of computing $\mathbf{Q}$ such that the entries of the separator output

$$\mathbf{y}(n) = \mathbf{C}\mathbf{s}(n) = \mathbf{Q}^T\mathbf{W}\mathbf{x}(n) = \mathbf{Q}^T\mathbf{z}(n) = \widetilde{\mathbf{Q}}\mathbf{z}(n)$$

are as independent as possible.

The separability condition for BSS problem has been studied in [12], and it was pointed out that for statistically independent non-Gaussian sources, the separation can be achieved by restoring the independence. In [12], the *mutual information* (MI) was suggested as a tool to measure the independence of the output signals, and the Edgeworth expansion was used to approximate the probability density function in the MI criterion. The Edgeworth expansion of the MI of a standardized (i.e. after whitening) real variable, up to an additive constant I_0 and as a function of standardized *cumulants*, is given as follows [12]:

$$-I[\mathbf{y}] \approx I_0 + \sum_i \big(4\kappa_{iii}^2(\mathbf{y}) + \kappa_{iiii}^2(\mathbf{y}) + 7\kappa_{iii}^4(\mathbf{y}) - 6\kappa_{iii}^2(\mathbf{y})\kappa_{iiii}(\mathbf{y})\big) \qquad (1)$$

where $\kappa_{iii}(\mathbf{y})$ and $\kappa_{iiii}(\mathbf{y})$ are the third-order and fourth-order marginal cumulants of each entry of $\mathbf{y}$, i.e., $\kappa_{iii}(\mathbf{y}) = \mathsf{E}[y_i^3]$ and $\kappa_{iiii}(\mathbf{y}) = \mathsf{E}[y_i^4] - 3\mathsf{E}^2[y_i^2]$. A cumulant of order higher than two qualifies as a *contrast* [12]. That is, the maximization of specific cumulants would result into a successful blind separation for particular type of sources. For example, if the sources are asymmetrical then the maximization of third-order cumulant $\kappa_{iii}^2(\mathbf{y})$ would be enough to ensure successful separation; similarly, for symmetric sources, the maximization of fourth-order $\kappa_{iiii}^2(\mathbf{y})$ would be sufficient [12], i.e.,

$$\mathscr{J}(\mathbf{y}) = \begin{cases} \sum_i \kappa_{iii}^2(\mathbf{y}), & \text{for asymm. sources} \\ \sum_i \kappa_{iiii}^2(\mathbf{y}), & \text{for symm. sources} \end{cases} \qquad (2)$$

Equation (2) is discriminating over the set of random vectors $\mathbf{y}$ provided there is at most one null third-order (resp. fourth-order) marginal cumulant for asymmetrical (resp. symmetrical) sources [12, 19]. There exist number of ways to find the nonmixing matrix $\widetilde{\mathbf{Q}}$ such that the contrast (2) is maximized; for example, [12, 19]. Just recently, Blaschke & Wiskott [20] showed that the *joint* use of third- and fourth-order cumulants is an *admissible* choice for a contrast:

$$\mathscr{J}(\mathbf{y}) = \sum_i \big(4\kappa_{iii}^2(\mathbf{y}) + \kappa_{iiii}^2(\mathbf{y})\big) \qquad (3)$$

The Blaschke-Wiskott's ICA algorithm is known as CuBICA. We are referring to [20] due to its simple presentation; otherwise, the joint use of third- and fourth-order cumulants for ICA has been investigated earlier by a number of researchers, like [21].

2.2 Weighted Contrast

In this section,a weighted contrast comprising third- and the fourth-order cumulants is presented as a generalization of (3) for the blind separation of mixture of symmetric and asymmetric sources, *viz*

$$\mathscr{J}(\mathbf{y}) = \sum_{i=1}^{M} \big(w_{S,i}\kappa_{iii}^2(\mathbf{y}) + w_{K,i}\kappa_{iiii}^2(\mathbf{y})\big) \qquad (4)$$

The contrast (4) does not arise directly from the MI criterion, but it is a weighted combination of two solutions (cf. (2)) that are not only contrast, but also provide a good approximation to MI under specific assumptions. However, in the absence of those assumptions, it is possible to obtain better results using (4) with appropriately selecting the values of free parameters. The algebraic nature of cumulants is tensorial (with symmetry) [22]; thanks to the multilinearity of cumulants $\kappa_{\dots}(\mathbf{y})$ in $\kappa_{\dots}(\mathbf{z})$, the criterion (4) becomes an implicit function of the elements of the unitary matrix $\widetilde{\mathbf{Q}}$, we obtain $\kappa_{iii}(\mathbf{y}) = \sum_{jkl} \widetilde{Q}_{ij}\widetilde{Q}_{ik}\widetilde{Q}_{il}\kappa_{jkl}(\mathbf{z})$ and $\kappa_{iiii}(\mathbf{y}) = \sum_{jklm} \widetilde{Q}_{ij}\widetilde{Q}_{ik}\widetilde{Q}_{il}\widetilde{Q}_{im}\kappa_{jklm}(\mathbf{z})$, where the unitary transformation matrix $\widetilde{\mathbf{Q}} = \mathbf{Q}^T$ is modeled as Givens rotation $\phi_{\mu\nu}$ which is a rotation around the origin within the plane of two selected components μ and ν. The estimation of the plane rotation $\phi_{\mu\nu}$ is obtained by an iterative Jacobi method over the set of orthonormal matrices. The orthonormal transforms are thus obtained as a sequence of plane rotations. Each plane rotation is applied to a pair of coordinates, such that, $y_\mu \leftarrow (y_\mu \cos\phi_{\mu\nu} + y_\nu \sin\phi_{\mu\nu})$ and $y_\nu \leftarrow (-y_\mu \sin\phi_{\mu\nu} + y_\nu \cos\phi_{\mu\nu})$, while leaving the other coordinates unchanged. Thus, the Jacobi approach considers a sequence of two-dimensional ICA problems. Considering the subspace of only two selected components, the Givens rotation matrix becomes:

$$\widetilde{\mathbf{Q}} = \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix}$$

That is, for $M = 2$, we get

$$\mathscr{J}(y) = w_{S,1}\kappa_{111}^2(\mathbf{y}) + w_{S,2}\kappa_{222}^2(\mathbf{y}) + w_{K,1}\kappa_{1111}^2(\mathbf{y}) + w_{K,2}\kappa_{2222}^2(\mathbf{y}) \qquad (5)$$

Now we have only four free parameters to determine in a pairwise fashion. In the blind scenario, where we usually have no a priori knowledge of mixing signals, tuning of these free parameters is not simple. Interestingly, the study of the *single* weight parameter for the optimized use of a fourth-order contrast function has been studied in [23, 24, 25]. Motivated by the convincing results reported in these works, we also limit our search to a single weight parameter. We select $w_{S,1} = w_{S,2} = 1$ and $w_{K,1} = w_{K,2} = \beta$, which lead to the following contrast:

$$\begin{aligned}\mathscr{J}(y) =& \kappa_{111}^2(\mathbf{y}) + \kappa_{222}^2(\mathbf{y}) \\ &+ \beta\left(\kappa_{1111}^2(\mathbf{y}) + \kappa_{2222}^2(\mathbf{y})\right)\end{aligned} \tag{6}$$

2.3 A Closed-Form Estimator

Owing to [20], it is possible to express the contrast (6) as the function of ϕ as follows:

$$\begin{aligned}\mathscr{J}(\phi) = \mathscr{A}_0 &+ \mathscr{A}_4\cos(4\phi + \phi_4) \\ &+ \mathscr{A}_8\cos(8\phi + \phi_8)\end{aligned} \tag{7}$$

where $\mathscr{A}_0$, $\mathscr{A}_4$ and $\mathscr{A}_8$ are positive constants. Note that the first term $\mathscr{A}_0$ plays no role in the estimation of ϕ. Similarly constants $\mathscr{A}_8$ and ϕ_8 do not comprise of third-order statistical information and contribute no significant role if the mixing sources are asymmetrical in nature (the detailed expressions for $\mathscr{A}_0$, $\mathscr{A}_8$ and ϕ_8 can be obtained in [20] for the specific case $\beta = 1$). The constant ϕ_4 in the middle term, however, comprise of third- and fourth-order statistics and can be used to obtain a closed form estimator for the separation of mixed symmetrical/asymmetrical sources. The angle ϕ that maximizes $\mathscr{A}_4\cos(4\phi + \phi_4)$ is (refer to [21] for theoretical details):

$$\hat{\phi} = -\frac{\phi_4}{4} = -\frac{1}{4}\arctan(\mathscr{S}_0 + \beta\mathscr{S}_1, \mathscr{C}_0 + \beta\mathscr{C}_1) \tag{8}$$

which exploits the relations $\mathscr{S}_0 + \beta\mathscr{S}_1 = \sin\phi_4$ and $\mathscr{C}_0 + \beta\mathscr{C}_1 = \cos\phi_4$. Also $\arctan(y,x)$ is the unique angle $\alpha \in (-\pi, \pi]$ for which $\cos(\alpha) = (x/\sqrt{x^2+y^2})$ and $\sin(\alpha) = (y/\sqrt{x^2+y^2})$. Constant $\mathscr{S}_0, \mathscr{S}_1, \mathscr{C}_0$ and $\mathscr{C}_1$ are computed as:

$$\begin{aligned}
\mathscr{S}_0 &= 24\big(\kappa_{111}(\mathbf{z})\kappa_{112}(\mathbf{z}) - \kappa_{122}(\mathbf{z})\kappa_{222}(\mathbf{z})\big)\\
\mathscr{S}_1 &= 4\Big(7\big(\kappa_{1111}(\mathbf{z})\kappa_{1112}(\mathbf{z}) - \kappa_{1222}(\mathbf{z})\kappa_{2222}(\mathbf{z})\big)\\
&\quad + 6\kappa_{1122}(\mathbf{z})\big(\kappa_{1112}(\mathbf{z}) - \kappa_{1222}(\mathbf{z})\big)\\
&\quad + \kappa_{1111}(\mathbf{z})\kappa_{1222}(\mathbf{z}) - \kappa_{1112}(\mathbf{z})\kappa_{2222}(\mathbf{z})\Big)\\
\mathscr{C}_0 &= 6\Big(\kappa_{111}^2(\mathbf{z}) + \kappa_{222}^2(\mathbf{z}) - 3\big(\kappa_{112}^2(\mathbf{z}) + \kappa_{122}^2(\mathbf{z})\big)\\
&\quad - 2\big(\kappa_{111}(\mathbf{z})\kappa_{122}(\mathbf{z}) + \kappa_{112}(\mathbf{z})\kappa_{222}(\mathbf{z})\big)\Big)\\
\mathscr{C}_1 &= 7\big(\kappa_{1111}^2(\mathbf{z}) + \kappa_{2222}^2(\mathbf{z})\big) - 36\kappa_{1122}^2(\mathbf{z})\\
&\quad - 2\kappa_{1111}(\mathbf{z})\kappa_{2222}(\mathbf{z}) - 32\kappa_{1112}(\mathbf{z})\kappa_{1222}(\mathbf{z})\\
&\quad - 12\big(\kappa_{1111}(\mathbf{z})\kappa_{1122}(\mathbf{z}) + \kappa_{1122}(\mathbf{z})\kappa_{2222}(\mathbf{z})\big)\\
&\quad - 16\big(\kappa_{1112}^2(\mathbf{z}) + \kappa_{1222}^2(\mathbf{z})\big)
\end{aligned}$$

If we consider $\beta = 1$, then the closed-form estimator (8) provides an equivalent formulation of ICA algorithm as CuBICA34a [20]; however, note that unlike CuBICA34a, the proposed expression (8) is a closed-form. In the next section, we study the optimal estimation of parameter β.

3. THE OPTIMUM WEIGHT PARAMETER

The optimum value of the weight parameter β in (8) can be obtained by performing *small error analysis*; i.e., the value of β is optimum if it minimizes the asymptotic (large-sample) mean-square error in the estimation of Givens rotation. First, we consider that the mixing matrix is orthonormal so that the prewhitening stage is not necessary. Further, the asymptotic analysis is carried out for the case of two real sources. In practice, the mixing model in Section **2** should also take into account a possible additive noise. This is considered hereafter because we want to take into account both the measurement noises and errors resulting from the first stage of whitening. Hence, now, the mixing model we consider reads $\mathbf{x} = \mathbf{A}\mathbf{s} + \mathbf{g}$, where $\mathbf{g}$ is the vector of additive noise. In a two-source scenario, each noise $g_i, i \in \{1,2\}$ is a zero-mean, independent and identically distributed Gaussian random signal with equal power, i.e., $\mathsf{E}[g_1^2] = \mathsf{E}[g_2^2] = \sigma^2$. Moreover, $g_i, i \in \{1,2\}$ are assumed statistically mutually independent and independent of the sources $s_i, i \in \{1,2\}$.

Now, we have to estimate an angle ϕ according to the maximization of $\mathscr{J}(\cdot)$, i.e., $\widehat{\phi} = \arg\max_\phi \mathscr{J}(\phi)$, where $\widehat{\phi}$ is an estimate of the true (separation) value $\widetilde{\phi}$. In practice, the maximization of contrast function does not provide the exact value of the parameter $\widetilde{\phi}$, since the true cumulants are actually approximated by the sample estimates. Replacing the expectations by sample averages leads to the empirical version of $\mathscr{J}(\mathbf{y})$, which is denoted $\widehat{\mathscr{J}}(\mathbf{y})$ and is given by

$$\widehat{\mathscr{J}}(\phi) = \widehat{\kappa}_{111}^2(\mathbf{y}) + \widehat{\kappa}_{222}^2(\mathbf{y}) + \beta\left(\widehat{\kappa}_{1111}^2(\mathbf{y}) + \widehat{\kappa}_{2222}^2(\mathbf{y})\right)$$

where $\widehat{\kappa}_{iii}(\mathbf{y}) = \frac{1}{N}\sum_{k=1}^N y_i^3(k)$, and $\widehat{\kappa}_{iiii}(\mathbf{y}) = -3 + \frac{1}{N}\sum_{k=1}^N y_i^4(k)$, $i = 1,2$. As a result, an estimation error is involved in the estimation of the true value $\widetilde{\phi}$. The estimated angle $\widehat{\phi}$ is actually the solution of the estimating equation $\widehat{\mathscr{J}}'(\widehat{\phi}) = \partial\widehat{\mathscr{J}}(\widehat{\phi})/\partial\phi|_{\phi=\widehat{\phi}} = 0$. Approximating this derivative around the true value $\widetilde{\phi}$ by means of its Taylor series expansion yields:

$$\widehat{\mathscr{J}}'(\widehat{\phi}) \approx \widehat{\mathscr{J}}'(\widetilde{\phi}) + \widehat{\mathscr{J}}''(\widetilde{\phi})(\widehat{\phi} - \widetilde{\phi})$$

where $\widehat{\mathscr{J}}''(\widetilde{\phi}) = \partial\widehat{\mathscr{J}}'(\phi)/\partial\phi|_{\phi=\widetilde{\phi}}$ and $\widehat{\mathscr{J}}'(\widetilde{\phi}) = \partial\widehat{\mathscr{J}}(\phi)/\partial\phi|_{\phi=\widetilde{\phi}}$. Assuming $\widehat{\phi}$ to be in the neighborhood of $\widetilde{\phi}$, we obtain $\widehat{\mathscr{J}}'(\widetilde{\phi}) \approx -\widehat{\mathscr{J}}''(\widetilde{\phi})(\widehat{\phi} - \widetilde{\phi})$. The *mean square error* (m.s.e.) is given by

$$\text{m.s.e.} = \frac{\mathsf{E}[(\widehat{\mathscr{J}}'(\widetilde{\phi}))^2]}{(\mathsf{E}[\widehat{\mathscr{J}}''(\widetilde{\phi})])^2}$$

When $\widehat{\phi} = \widetilde{\phi}$, $\mathbf{y} = \mathbf{s} + \mathbf{g}'$, where $\mathbf{g}' = \mathbf{C}\mathbf{g}$. The m.s.e. expression is generalized and is thus valid for any two-dimensional contrast for ICA problem. Further, the *strong law of large number* ensures that $\widehat{\mathscr{J}}''(\widetilde{\phi})$ converges with probability one to its expected value. As $N \to \infty$, we have

$$\mathsf{E}[\widehat{\mathscr{J}}''(\widetilde{\phi})] \to -2(B_0 + B_1\beta)$$

B_0 and B_1 are defined in (10). Next we obtain

$$\mathsf{E}\left[(\widehat{\mathscr{J}}'(\widetilde{\phi}))^2\right] \to \frac{4}{N}(A_0 - 2A_1\beta + A_2\beta^2)$$

where A_0, A_1 and A_2 are defined in (10). The m.s.e. depends on the statistics of the sources and on the parameter β. We now easily derive the optimum value of β, denoted β^*, such that the m.s.e. is minimum by solving the equation $\frac{\partial}{\partial\beta}$m.s.e. $= 0$, i.e.,

$$\beta^* = \frac{A_0 B_1 + A_1 B_0}{A_2 B_0 + A_1 B_1}. \tag{9}$$

where

$$\begin{aligned}
A_0 &= 9\Big(B_0\sigma^6 + 3B_0\sigma^4 + (3B_0 + d_0)\,\sigma^2 \\
&\quad + B_0 + d_0 - 2c_0^2\Big) \\
A_1 &= 12\Big(d_1\sigma^6 + (3d_1 - 10d_0)\,\sigma^4 \\
&\quad - (d_3 - d_2 - 3d_1 + 20d_0)\,\sigma^2 \\
&\quad - d_3 + d_2 + d_1 - 10d_0\Big) \\
A_2 &= 16\Big(\left(\tfrac{15}{4}B_1 - 18c_1\right)\sigma^8 + (15B_1 - 72c_1)\,\sigma^6 \\
&\quad + \left(d_5 - d_4 + \tfrac{90}{4}B_1 - 108c_1\right)\sigma^4 \\
&\quad + (d_7 + d_6 + 2d_5 - 2d_4 - 72c_1 + 15B_1)\,\sigma^2 \\
&\quad + d_7 + d_6 + d_5 - d_4 - 2c_1^2 - 18c_1 + \tfrac{15}{4}B_1\Big)
\end{aligned}$$

where auxiliary variables are defined as

$$\begin{aligned}
B_0 &= 3\left(\kappa_{111}^2(\mathrm{s}) + \kappa_{222}^2(\mathrm{s})\right) \\
B_1 &= 4\left(\kappa_{1111}^2(\mathrm{s}) + \kappa_{2222}^2(\mathrm{s})\right) \\
c_0 &= \kappa_{111}(\mathrm{s})\kappa_{222}(\mathrm{s}) \\
c_1 &= \kappa_{1111}(\mathrm{s})\kappa_{2222}(\mathrm{s}) \\
d_0 &= \kappa_{111}^2(\mathrm{s})\kappa_{1111}(\mathrm{s}) + \kappa_{222}^2(\mathrm{s})\kappa_{2222}(\mathrm{s}) \\
d_1 &= 3\left(\kappa_{111}^2(\mathrm{s})\kappa_{2222}(\mathrm{s}) + \kappa_{222}^2(\mathrm{s})\kappa_{1111}(\mathrm{s})\right) \\
d_2 &= \kappa_{111}^2(\mathrm{s})\kappa_{2222}^2(\mathrm{s}) + \kappa_{222}^2(\mathrm{s})\kappa_{1111}^2(\mathrm{s}) \\
d_3 &= \kappa_{111}(\mathrm{s})\kappa_{1111}(\mathrm{s})\kappa_{11111}(\mathrm{s}) \\
&\quad + \kappa_{222}(\mathrm{s})\kappa_{2222}(\mathrm{s})\kappa_{22222}(\mathrm{s}) \\
d_4 &= 6\left(\kappa_{1111}^2(\mathrm{s})\kappa_{2222}(\mathrm{s}) + \kappa_{2222}^2(\mathrm{s})\kappa_{1111}(\mathrm{s})\right) \\
d_5 &= 15\left(\kappa_{1111}^3(\mathrm{s}) + \kappa_{2222}^3(\mathrm{s})\right) \\
d_6 &= 10\left(\kappa_{111}^2(\mathrm{s})\kappa_{1111}^2(\mathrm{s}) + \kappa_{222}^2(\mathrm{s})\kappa_{2222}^2(\mathrm{s})\right) \\
d_7 &= \kappa_{1111}^2(\mathrm{s})\kappa_{111111}(\mathrm{s}) + \kappa_{2222}^2(\mathrm{s})\kappa_{222222}(\mathrm{s})
\end{aligned}$$

which indicates that the β^* depends on the statistics of mixing source and additive noise, and is independent of the coefficients of unknown mixing matrix. Hence, given the source and noise statistics, we can obtain a contrast with minimum asymptotic m.s.e. Finally, with the help of β^* (9), the optimum value of Givens rotation is estimated as:

$$\hat{\phi}^* = -\frac{1}{4}\arctan(\mathscr{S}_0 + \beta^*\mathscr{S}_1, \mathscr{C}_0 + \beta^*\mathscr{C}_1) \tag{10}$$

The closed-form optimized estimator (10) is named *Optimized Composite-Order ICA* **(OCOICA)**.

4. SIMULATION RESULTS

In order to illustrate the potential of the proposed estimator OCOICA, some computer simulations are now presented. We intend to compare the performance of the OCOICA with three existing ICA algorithms including the third-order cumulant based ICA (Com3) [19], the fourth-order cumulant based ICA (Com4) [12], and the CuBICA [20]. The performance measure, interference-to-signal ratio (ISR), has been used in our simulation to characterize quantitatively the restitution quality. The performance index reads

$$\mathrm{ISR} = \sum_{i=1}^{M}\left(\frac{\sum_{j=1}^{n}|c_{ij}|^2}{\max_j |c_{ij}|^2} - 1\right) \tag{11}$$

where c_{ij} represents the element (i, j) of the global mixing-nonmixing matrix $\mathbf{C}$. In the two-signal case, the ISR approximates the m.s.e. of the angle estimates around any valid separation solution [25]. Random sources with desired skewness and kurtosis are generated by *Fleishman's power distribution method* [26]. All sources are drawn with zero-mean and unit-variance.

We estimated the ISR performance as a function of sample size. The mixing matrix was taken to be of order 2×2. Matrices $(\mathbf{A})$ are generated from normal distribution with zero-mean and unit-variance. The condition number of $\mathbf{A}$ is constrained to be less than 50. Two cases are considered – no noise $\sigma = 0$ and with noise $\sigma = 0.0316$ [i.e., SNR=30dB]. The mixture is first whitened via PCA based on the singular value decomposition of the observed data matrix. The curves have been averaged over 2000 independent Monte Carlo runs. The weight parameter β has been computed from the statistical knowledge of whitened-sources. Results are depicted in Figure 1 comparing the performance of the OCOICA with those of Com3, Com4 and CuBICA.

In Figure 1(a), original sources are highly asymmetric in nature, that is why Com3 is performing better than Com4 and CuBICA, while in Figure 1(b), original sources are moderately skewed and CuBICA is performing better than Com4 and Com3. Notice that, the OCOICA is performing better than Com3, Com4 and CuBICA algorithms in both noise-free and noisy scenarios. Finally notice that the improvement in ISR reduction achieved by the proposed estimator is consistent even though the free parameter β has been computed from whitened-sources.

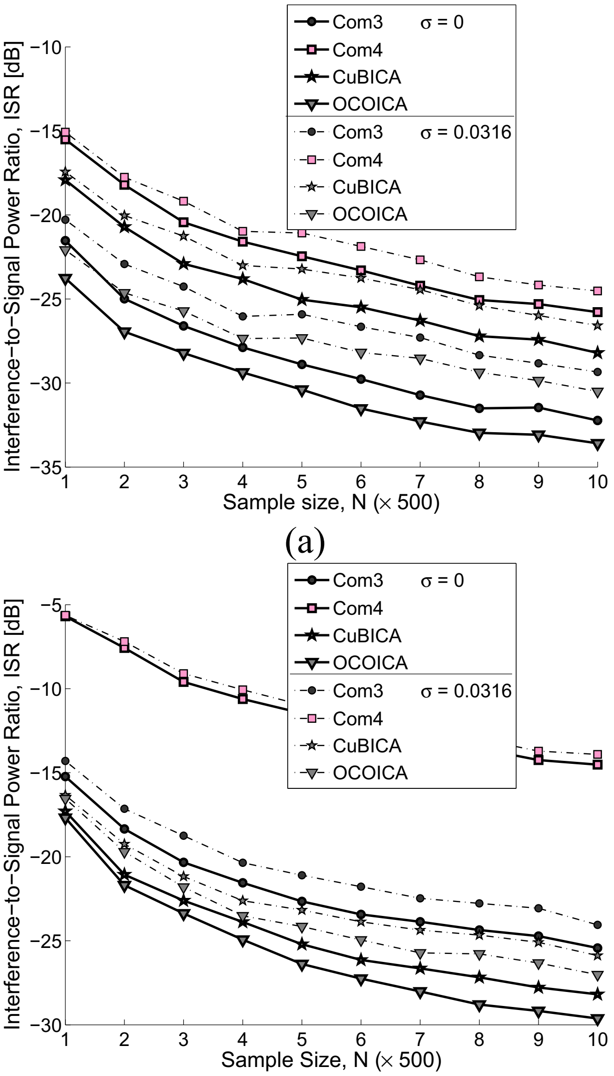
5. CONCLUSIONS

This work explores the combination of third- and fourth-order cumulant based tensor diagonalization in an optimal sense and comes up with a closed-form estimator. Computer simulation for the separation of real-valued sources is provided. The proposed estimator is shown to be outperforming three existing ICA algorithms. The proposed closed-form estimator can handle *symmetric and asymmetric* distributed sources, and exhibits *better* performance.

ACKNOWLEDGEMENT: One of the authors (Shafayat Abrar) would like to acknowledge the financial support of the Overseas Research Studentship Awards Scheme, UK, the University of Liverpool, UK, and COMSATS Institute of Information Technology (CIIT), Islamabad, Pakistan. SA is on leave of absence from the CIIT for higher studies.

REFERENCES

[1] V. Zarzoso and A.K. Nandi. *Blind Estimation Using Higher-Order Statistics*, chapter Blind source separation, page 167252. Boston, MA: Kluwer Academic, 1999.



(b)

Figure 1: ISR performance for 2-source mixing scenario. (a) $\kappa_{111}(\mathbf{s}) = -1.75$, $\kappa_{222}(\mathbf{s}) = 1.5$, $\kappa_{1111}(\mathbf{s}) = 3.75$ and $\kappa_{2222}(\mathbf{s}) = 3.5$; (b) $\kappa_{111}(\mathbf{s}) = -\kappa_{222}(\mathbf{s}) = 0.5$, $\kappa_{1111}(\mathbf{s}) = \kappa_{2222}(\mathbf{s}) = -0.25$. Solid lines, no noise; dashed lines, SNR=30dB.

[2] A. Hyvarinen, J. Karhunen, and E. Oja. *Independent Component Analysis*. New York: Wiley, 2001.

[3] P. Comon. Separation of stochastic processes. *Proc. Workshop Higher-Order Spectral Analysis*, page 174179, 1989.

[4] V. Zarzoso. *Closed-form higher-order estimators for blind separation of independent source signals in instantaneous linear mixtures*. PhD thesis, Univ. of Liverpool, Liverpool, UK, Oct. 1999.

[5] F. Harroy and J.-L. Lacoume. Maximum likelihood estimators and Cramer-Rao bounds in source separation. *Signal Process.*, 55(2):167–177, 1996.

[6] V. Zarzoso and A.K. Nandi. Blind separation of independent sources for virtually any source probability density function. *IEEE Trans. Signal Process.*, 47(9):2419–2432, Sep. 1999.

[7] M. Ghogho, A. Swami, and T. Durrani. Approximate maximum likelihood blind source separation with arbitrary source pdfs. *Proc. IEEE 10th Statistical Signal and Array Processing*, page 368372, Aug. 2000.

[8] V. Zarzoso, A.K. Nandi, F. Herrmann, and J. Millet-Roig. Combined estimation scheme for blind source separation with arbitrary source pdfs. *Electron. Lett.*, 37(2):132–133, Jan. 2001.

[9] F. Herrmann and A. Nandi. Blind separation of linear instantaneous mixture using closed-forms estimators. *Signal Process.*, 81(7):15371556, Jul. 2001.

[10] J. J. Murillo-Fuentes and F. J. Gonzlez-Serrano. A sinusoidal contrast function for the blind separation of statistically independent sources. *IEEE Trans. Signal Process.*, 52(12):34593463, 2004.

[11] V. Zarzoso, J.-J. Murillo-Fuentes, R. Boloix-Tortosa, and A.K. Nandi. Optimal pairwise fourth-order independent component analysis. *IEEE Trans. Signal Process.*, 54(8):3049–3063, 2006.

[12] P. Comon. Independent component analysis, a new concept? *Signal Process.*, 36(3):287314, 1994.

[13] J.-F. Cardoso. High-order contrasts for independent compoent analysis. *Neural Computation*, 11:157–192, 1999.

[14] C.L. Nikias and A.P. Petropulu. *Higher-Order Spectra Analysis*. Prentice-Hall, Englewood Cliffs, NJ, 1993.

[15] C. Ziegaus and E.W. Lang. Statistics of natural and urban images. *Lecture Notes in Computer Science*, 1327:219224, 1997.

[16] J. Karvanen and V. Koivunen. Independent component analysis via optimum combining of kurtosis and skewness-based criteria. *Journal of the Franklin Institute*, 341:401–418, 2004.

[17] S. Choi, R. Liu, and A. Cichocki. A spurious equilibria-free learning algorithm for the blind separation of non-zero skewness signals. *Neural Process. Lett.*, 7:61–68, 1998.

[18] S. Abrar and A.K. Nandi. Independent component analysis: Jacobi-like diagonalization of optimized composite-order cumulants. *Proc. R. Soc. A*, 465(2105):1393–1411, May 2009.

[19] P. Comon. Tensor diagonalization, a useful tool in signal processing. *Proc. IFAC-SYSID, 10th IFAC Symp. System Identification*, pages 77–82, July 1994.

[20] T. Blaschke and L. Wiskott. CuBICA: Independent component analysis by simultaneous third- and fourth-order cumulant diagonalization. *IEEE Trans. Sig. Process.*, 52(5):1250–1256, 2004.

[21] E. Moreau. A generalization of joint-diagonalization criteria for source separation. *IEEE Trans. Signal Processing*, 49(3):530–541, March 2001.

[22] P. McCullagh. *Tensor methods in statistics*. Chapman and Hall, London, 1987.

[23] E. Moreau and N. Thirion-Moreau. Nonsymmetrical contrasts for sources separation. *IEEE Trans. Sig. Process.*, 47(8):2241–2252, 1999.

[24] J.J. Murillo-Fuentes and F.J. Gonzalez-Serrano. A sinusoidal contrast function for the blind separation of statistically independent sources. *IEEE Trans. Sig. Process.*, 52(12):3459–3463, Dec. 2004.

[25] V. Zarzoso, J.J. Murillo-Fuentes, R. Boloix-Tortosa, and A.K. Nandi. Optimal pairwise fourth-order independent component analysis. *IEEE Trans. Sig. Process.*, 54(8):3049–3063, 2006.

[26] A.I. Fleishman. A method for simulating non-normal distributions. *Psychometrika*, 43(4):521–532, 1978.