Blind Separation of Independent Sources for Virtually Any Source Probability Density Function

Vicente Zarzoso, *Student Member, IEEE*, and Asoke K. Nandi, *Senior Member, IEEE*

***Abstract*—The blind source separation (BSS) problem consists of the recovery of a set of statistically independent source signals from a set of measurements that are mixtures of the sources when nothing is known about the sources and the mixture structure. In the BSS scenario, of two noiseless real-valued instantaneous linear mixtures of two sources, an approximate maximum-likelihood (ML) approach has been suggested in the literature, which is only valid under certain constraints on the probability density function (pdf) of the sources. In the present paper, the expression for this ML estimator is reviewed and generalized to include virtually any source distribution. An intuitive geometrical interpretation of the new estimator is also given in terms of the scatter plots of the signals involved. An asymptotic performance analysis is then carried out, yielding a closed-form expression for the estimator asymptotic pdf. Simulations illustrate the behavior of the suggested estimator and show the accuracy of the asymptotic analysis. In addition, an extension of the method to the general BSS scenario of more than two sources and two sensors is successfully implemented.**

***Index Terms*— Asymptotic performance, blind signal separation, closed-form estimators, higher order statistics, maximum likelihood, scatter diagram.**

I. Introduction

A NUMBER of applications in a variety of areas comprise the task of obtaining certain signals (so-called *sources*), which are not directly accessible but have to be extracted (*separated*) from another set of measurable signals regarded as mixtures of the sources. As neither the source signals nor the mixing structure are known, this is referred to as the *blind source separation* (BSS) problem. Major fields of application include array processing, communications, biomedical signal processing, image processing, and speech processing. The following are just a few examples where the BSS plays a relevant role: direction-of-arrival (DOA) estimation when waveforms are distorted and/or the array layout unknown or poorly calibrated, jammer rejection, multichannel blind equalization/deconvolution, mobile radio and regenerative satellite communications, data communications in the presence of cross-coupling effects, airport surveillance, speech recorded in the presence of background noise and/or competing speakers, blind image restoration, fetal electrocardiogram extraction, seismic exploration, etc. Several types of mixtures exist, such as convolutive or nonlinear, but in many of the above cases the signals observed at the sensor output can be considered to be instantaneous linear combinations of the sources so that the mixing structure is represented by an unknown *mixing matrix*. The BSS of instantaneous linear mixtures, which is schematically portrayed in Fig. 1, is the focus of this paper.

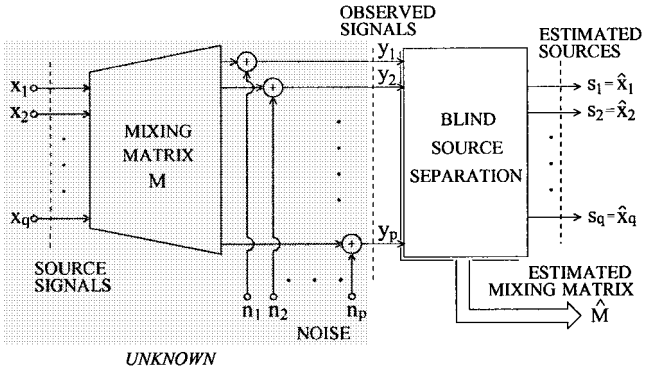


Fig. 1. Schematic diagram of the BSS of instantaneous linear mixtures.

The development of *higher order statistics* (HOS) in the last decade boosted the interest in this emerging research area, despite the fact that second-order (i.e., correlation-based) techniques can perform the source extraction under particular conditions. For instance, the separation is feasible resorting only to *second-order statistics* (SOS) for temporally colored (although spatially white) sources with different spectral content [1], [16]. However, if the sources are temporally white, the time information can no longer be considered, and then, the source waveforms cannot be identified undistorted by employing only SOS [10] unless the *transfer vectors* (the columns of the mixing matrix, also known as *source directions*) are orthogonal [18]. On the other hand, SOS-based procedures are also able to extract sources with an arbitrary spectral matrix by exploiting some prior knowledge on the mixing matrix structure, as done by the MUSIC method for DOA [12]. Strictly speaking, however, these eigenstructure-based or spatial-coherence exploitation techniques cannot be regarded as purely *blind* since by definition, a blind algorithm works under no assumptions on the mixing structure. Proceeding blindly often exhibits some advantages. For example, in the context of narrowband array processing, when the propagation conditions or the source and sensor locations can suffer variations that are difficult to

Manuscript received February 27, 1998; revised February 9, 1999. This work was supported by the University of Liverpool and by the Engineering and Physical Sciences Research Council (EPSRC) of the U.K. under Research Grant GR/L/07475. The associate editor coordinating the review of this paper and approving it for publication was Dr. Jonathon A. Chambers.

The authors are with the Signal Processing and Communications Group, Department of Electrical Engineering and Electronics, The University of Liverpool, Liverpool, U.K. (e-mail: aknandi@liverpool.ac.uk).

Publisher Item Identifier S 1053-587X(99)06766-5.

predict or model (e.g., multipaths in an urban environment), then it may be wiser to resort to a blind procedure for recovering the source signals [3]. Along the same lines, blind beamforming is shown in [2] to improve "informed" beamforming when multipath propagation takes place. The former is, by nature, insensitive to errors in the array manifold model.

In an actual BSS problem, only one major hypothesis is relied on to achieve the separation: the spatial statistical independence of the sources [2]–[7], [9]–[11]. In order for temporally white sources to be separated without waveform distortion under this sole assumption, the application of HOS is imperative. In statistical terms, a linear transformation on the observed variables is sought so that it results in a set of independent components. This is the so-called *independent component analysis* (ICA) of the observations [6]. The HOS (usually the higher order cumulants) can be used as a natural measure of the degree of independence. Equivalently, the problem of the source extraction becomes the task of separating the observed joint probability density function (jpdf) into the independent source pdf's that generate the former through a linear transformation. In signal terms, this leads to a successful source-waveform estimation. A necessary and sufficient condition for the waveform-preserving source estimation to be feasible is that there be, at most, one Gaussian signal among the sources [17]. This lack of ability of HOS in dealing with Gaussian sources stems from the fact that the cumulants of such distributions vanish for orders greater than two. An alternative explanation is that any linear combination of Gaussian random variables is Gaussian as well, and therefore, there is no way of discerning the original variables in the measured linear mixture.

A two-step approach is almost universally adopted when facing a source extraction problem [1], [2], [4]–[6], [9]. The first step is a *(pre-)whitening* operation, consisting of a projection of the observed processes on the signal subspace (i.e., the space spanned by the transfer vectors) plus power normalization, which can all be accomplished by means of second-order tools. This process is equivalent to the diagonalization of the sensor covariance matrix. After this initial step, only a unitary transformation remains to be identified in order to determine the actual source directions and, hence, the source waveforms and mixing matrix. It is the identification of this unitary matrix, carried out in a second step, that needs the intervention of higher order techniques when the temporal information cannot be exploited or is simply ignored. Algebraically, the transformation must diagonalize the higher order cumulant tensor of the whitened sensor output, thus providing a set of higher order independent signals: the wanted sources. In the basic scenario composed of two real-valued mixtures of two sources, the unitary transformation is reduced to a Givens rotation matrix so that only a rotation angle needs to be estimated.

Several methods have been proposed in the literature for the estimation of this angle. In [5], Comon finds an estimator that is a function of the fourth-order cross-cumulants of the whitened sensor signals. In [6], the angle estimation is accomplished by the maximization of contrast functions, giving rise to the first formal mathematical definition of ICA. The ML principle is considered by Harroy and Lacoume in [9], which is reminiscent of the seminal work on the ML approach to BSS carried out by the latter and Gaeta in [7]. The source pdf is approximated by its Gram–Charlier expansion truncated at the fourth order. Taking advantage of this expansion, an approximate ML estimator is found when no noise disturbs the measurements. However, the conditions for the Gram–Charlier expansion to be valid and certain simplifications and assumptions made in order to obtain a tractable analytical expression restrict the applicability of this estimator to the case of symmetric sources with the same distribution and kurtosis value lying in a certain positive range.

In the present paper, the expression of the Harroy–Lacoume's ML estimator is taken up, and a similar method is put forward and analyzed. The most appealing feature of the propounded estimator is that it is able to deal with almost any combination of source pdf's of arbitrary symmetry and kurtosis value. In this sense, the new estimator can be considered to be a generalization of the approximate ML approach, broadening its applicability domain to situations outside the validity region of the Gram–Charlier expansion and the other restrictions imposed in [9].

The paper is organized as follows. Section II summarizes the model, terminology, and basic assumptions of the problem in hand. The ML approach developed in [9] is recalled in Section III. Founded on this method, the new angle estimator is introduced in Section IV. Section V unveils an interesting geometrical rationale behind the new procedure in terms of the scatter diagrams of the signals involved. Later, Section VI is entirely devoted to studying the statistical properties of the estimator for both finite and large sample size. When considering this latter case, an analytical expression for the estimator asymptotic pdf is sought. A few experiments carried out to illustrate and validate the theoretical results are reported in Section VII. In that section, an extension to the general BSS setup composed of more than two sources and two measurements is also examined, and the effects of prewhitening on the estimator performance is evaluated. Finally, Section VIII sums up, provides some concluding remarks, and suggests potential avenues of further research.

II. Preliminaries: Problem Statement and Terminology

A. Problem Description

The goal of BSS can be briefly stated as recovering a set of q zero-mean statistically independent *source signals* $\mathbf{x}(k) = [x_1(k), \cdots, x_q(k)]^{\mathrm{t}}$ from a set of p instantaneous linear mixtures $\mathbf{y}(k) = [y_1(k), \cdots, y_p(k)]^{\mathrm{t}}$, which are the *observed signals* or *sensor output*. Symbol k represents a time index. In matrix form and in the noiseless case, this problem accepts the model

$$\mathbf{y}(k) = M\mathbf{x}(k), \qquad k = 1, 2, \cdots \tag{1}$$

where $M \in \mathbb{R}^{p\times q}$ is the *mixing* or *transfer matrix*, assumed full column rank. In the sequel, we will only consider real-valued sources and mixtures. The power of the sources is,

in principle, arbitrary since a scalar factor can be swapped between any source and its associated column in the mixing matrix without altering the measurements. The ordering of the sources is also irrelevant. These well-known facts constitute a basic indeterminacy in BSS. At best, matrix M can be identified up to a permutation and scaling of its columns. Therefore, nothing prevents us from further assuming that the sources are unit-power signals $\mathrm{E}[x_i^2(k)] = 1$ for $1 \leq i \leq q$ so that their dynamic range is accounted for by the magnitudes of the columns of M. It should be noted, however, that this is merely a convention. The source covariance matrix, hence, becomes

$$R_x \triangleq \mathrm{E}[\mathbf{x}\mathbf{x}^{\mathrm{t}}] = I_q \tag{2}$$

where the symbol t denotes the transpose operator and I_q the $q \times q$ identity matrix.

The first step to achieve the source extraction is to project the observations on the signal subspace (spanned by the columns of M) and to normalize them, which is a process involving conventional second-order analysis, as commented in the introduction. Comon refers to this preliminary second-order processing as *standardization* [6]. Consider the singular value decomposition (SVD) [8] of the mixing matrix $M = U\Sigma Q$. The covariance matrix of the observations may then be expressed as a function of M and its SVD, and in terms of its own eigenvalue decomposition (EVD)

$$R_y \triangleq \mathrm{E}[\mathbf{y}\mathbf{y}^{\mathrm{t}}] = \begin{cases} MR_xM^{\mathrm{t}} = U\Sigma^2U^{\mathrm{t}} \\ V\Delta V^{\mathrm{t}}. \end{cases} \tag{3}$$

From these equalities, the range space (i.e., the column space) of the transfer matrix can be identified as $U\Sigma = V\Delta^{1/2}$. The (pseudo)inverse of this matrix $(U\Sigma)^{\dagger}$ is referred to as the *whitening matrix* since when it is applied to the sensor output, it supplies a set of uncorrelated (second-order spatially white) unit-variance components, that is, with identity covariance matrix

$$\mathbf{z}(k) = (U\Sigma)^{\dagger}\mathbf{y}(k) = Q\mathbf{x}(k) \tag{4}$$

$$R_z \triangleq \mathrm{E}[\mathbf{z}\mathbf{z}^{\mathrm{t}}] = QR_xQ^{\mathrm{t}} = I_q. \tag{5}$$

Observe that second-order analysis leaves a unitary transformation undisclosed: the matrix Q in (4) that relates the whitened signals to the true sources. The estimation of this matrix requires the use of HOS-based techniques. In the case of batch processing, T signal samples are stored in a measurement matrix Y, with one vector sample per column. Then, the whitening process above does not require the explicit computation of the sensor-output covariance matrix, but it can be directly affected from the SVD of the observation matrix [6].

From now on, let us focus on the simplified two-source two-sensor BSS scenario, where the matrix Q becomes a Givens rotation of the form

$$Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}. \tag{6}$$

The problem, then, is reduced to determining the angle θ from the whitened observations. Once this has been achieved, a counter-rotation of this angle performed on the whitened observations provides the original source waveforms.

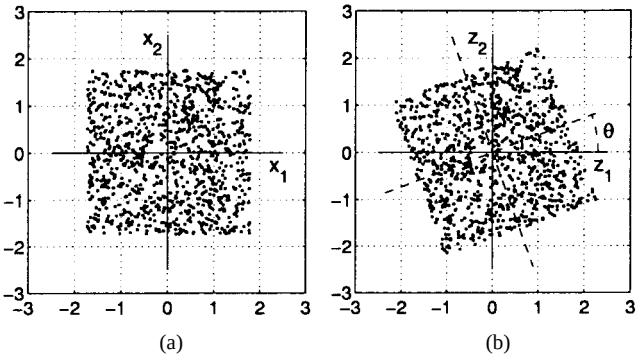


Fig. 2. Scatter diagram for 1000 samples of (a) original sources, two uniformly distributed independent signals, and (b) sensor signals after prewhitening. The latter corresponds to the former rotated by an unknown angle θ.

B. The Scatter Diagram

Consider the pair of signals $\mathbf{z}(k) = [z_1(k), z_2(k)]^{\mathrm{t}}$. The bidimensional plot of the amplitude points $(z_1(k), z_2(k))$, $k = 1, 2, \cdots$ is called the *scatter diagram* [9]. The scatter diagram is an approximation of the true jpdf of the corresponding pair of variables. Areas within the scatter diagram of high point density correspond to a high value for the associated jpdf around the same area. The left panel of Fig. 2 shows the scatter diagram of 1000 realizations of two uniformly distributed independent random variables (x_1, x_2), which are the source signals. It can be appreciated, in the first place, how the density is, roughly, uniform over the definition range, which means that the jpdf is flat. In addition, it corresponds to a jpdf of independent variables since it can be decomposed as the product of its two 1-D marginal pdf's. The effect of an orthogonal transformation is a scatter plot rotation. Effectively, the scatter diagram corresponding to the whitened pair of variables (z_1, z_2) connected to the sources according to (4) is represented in the right-hand panel of Fig. 2. It preserves the exact shape of the previous one, but it is rotated an angle θ with respect to the new axes. Observe that the source directions are the symmetry axes of the prewhitened jpdf, represented by the dashed lines in Fig. 2(b).

The scatter diagram points accept a polar as well as a complex form representation

$$\left.\begin{aligned} (x_1(k), x_2(k)) &= x_1(k) + jx_2(k) = \rho_k\, e^{j\phi'_k} = \rho_k \angle \phi'_k \\ (z_1(k), z_2(k)) &= z_1(k) + jz_2(k) = \rho_k\, e^{j\phi_k} = \rho_k \angle \phi_k \end{aligned}\right\} \quad k = 1, 2, \cdots \tag{7}$$

where, from (4) and (6), the angles ϕ'_k and ϕ_k are readily related by

$$\phi_k = \phi'_k + \theta. \tag{8}$$

These polar and complex form representations will prove very convenient at certain points of the development and will be helpful in providing a very interesting geometrical interpretation of the estimator introduced in this paper.

C. Some Statistical Relationships

Before proceeding, let us define some statistical terms. In the first place, $\mu^x_{mn} = \mathrm{E}[x_1^m x_2^n]$ represents the $(m+n)$th-order moment of the bivariate random variable $\mathbf{x} = [x_1, x_2]^{\mathrm{t}}$. Analogously, $\kappa^x_{mn} = \mathrm{Cum}_{mn}[x_1, x_2]$ denotes the $(m+n)$th-order cumulant of the same pair of variables. Similar notation can be employed for the whitened observations, just by substituting the super-index "x" with "z" in the moment and cumulant expressions. Moreover, it will be useful to recall the following relationships [15], which hold under the assumption of zero-mean unit-variance uncorrelated components:

$$\begin{aligned} &\kappa^z_{40} = \mu^z_{40} - 3; \qquad \kappa^z_{04} = \mu^z_{04} - 3 \\ &\kappa^z_{31} = \mu^z_{31}; \qquad \kappa^z_{13} = \mu^z_{13}; \qquad \kappa^z_{22} = \mu^z_{22} - 1 \\ &\kappa^x_{31} = \kappa^x_{13} = \kappa^x_{22} = 0. \end{aligned} \tag{9}$$

The first two lines express the fourth-order marginal cumulants, or *kurtosis*, and the fourth-order cross-cumulants of the whitened measurements as a function of their moments. Totally analogous expressions hold for the kurtosis and cross-cumulants of the sources. The final identity (zero source fourth-order cross-cumulants) comes from the source independence assumption. Owing to the unit-variance convention, all statistics become normalized or standardized [6].

III. An Approximate Maximum-Likelihood Estimator

By definition, the ML estimator of θ is the value of the rotation angle that maximizes the (log-)likelihood of the given whitened observations $\mathbf{z}(k)$. Mathematically, this idea can be expressed as

$$\hat{\theta}_{\mathrm{ML}} = \arg\max_{\theta} \sum_{k=1}^{T} \log p_z(\mathbf{z}(k)|\theta) \tag{10}$$

for T independent observations $\mathbf{z}(k)$, $k = 1, 2, \cdots, T$, where $p_{\mathbf{z}}(\cdot)$ denotes the pdf of the decorrelated measurements related to the source pdf through $p_{\mathbf{z}}(\mathbf{z}|\theta) = p_{\mathbf{x}}(Q^{-\mathbf{1}}\mathbf{z}) = \prod_{i=1}^{q} p_{x_i}((Q^{\mathrm{t}}\mathbf{z})_i)$, where $p_{x_i}(\cdot)$ is the marginal pdf of the ith source.

In [9], the source jpdf was approximated by its Gram–Charlier expansion truncated at the fourth-order term (hence, the adjective "approximate"). In order for this expansion to be a valid pdf approximation, the kurtosis of each source must lie in the range $[0, 4]$, which excludes all short-tailed (platykurtic, sub-Gaussian, with negative kurtosis) pdf's and even some long-tailed (leptokurtic, super-Gaussian, with positive kurtosis) ones, such as the exponential distribution. Furthermore, in a bid to simplify the development, two other conditions were assumed: The sources have the same distribution (at least up to the fourth-order in their Gram–Charlier expansion) and are symmetric, i.e., with zero skewness ($\kappa^x_{30} = \kappa^x_{03} = 0$), and therefore, third-order terms were also neglected. Then, developing the ML criterion (10) over the Gram–Charlier expansion of the source jpdf under these conditions, the ML estimator of θ was found to be

$$\hat{\theta}_{\mathrm{ML}} = \frac{1}{4} \arctan \frac{\sum_k \rho_k^4 \sin 4\phi_k}{\sum_k \rho_k^4 \cos 4\phi_k}. \tag{11}$$

The properties of the estimator were also studied in [9]. Under all the above assumptions, it was found to be unbiased for any sample size. In addition, the Cramér–Rao lower bound (CRLB) for the estimation of θ under the conditions of this ML development, i.e., when the source pdf's are approximated by their Gram–Charlier expansion, was deduced to be

$$\mathrm{Var}[\hat{\theta}] \geq \frac{6}{T[(\kappa^x_{40})^2 + (\kappa^x_{04})^2]} \tag{12}$$

where $\hat{\theta}$ is expressed in radians.

The same authors realized in [9] the lack of sensitivity of the estimator (11) to the data pdf, testing its robustness for sources out of the validity domain of its development, such as uniformly distributed signals, whose kurtosis is negative. Surprisingly, the results obtained were very favorable. An explanation for this unexpected phenomenon was given from a geometrical standpoint, based on the scatter plot and the source jpdf symmetry. Further experiments involving asymmetric sources with kurtosis values outside the interval $[0, 4]$ (e.g., exponential distributions) aroused the suspicion that the validity scope of (11) was actually less restricted than originally predicted. In fact, it is proved in the next section that the approximate ML estimator admits quite a straightforward generalization, whereby almost any source joint distribution can be treated.

IV. Generalization

A. The Extended ML Estimator

Dropping the time index k in the sequel for convenience, let us define the fourth-order complex statistic

$$\xi \triangleq \mathrm{E}[\rho^4 e^{j4\phi}] = \mathrm{E}[(z_1 + jz_2)^4]. \tag{13}$$

According to (7) and (8), (13) accepts an expansion as a function of the sources and the unknown rotation angle

$$\xi = e^{j4\theta}\mathrm{E}[\rho^4 e^{j4\phi'}] = e^{j4\theta}\mathrm{E}[(x_1 + jx_2)^4]. \tag{14}$$

However, from the expressions given in (9)

$$\mathrm{E}[(x_1 + jx_2)^4] = \kappa^x_{40} + \kappa^x_{04} \tag{15}$$

and therefore, ξ may be expressed as a function of the source statistics

$$\xi = e^{j4\theta}(\kappa^x_{40} + \kappa^x_{04}). \tag{16}$$

The term in brackets, the *source kurtosis sum* (sks), is not known because, by definition, the sources are not known either.

However, from (7) and (9), the sks can be obtained as

$$\gamma \triangleq \mathrm{E}[\rho^4] - 8 = \mathrm{E}[(x_1^2 + x_2^2)^2] - 8 = \kappa_{40}^x + \kappa_{04}^x \quad (17)$$

which is also available as a function of the whitened data

$$\gamma = \mathrm{E}[(z_1^2 + z_2^2)^2] - 8 \quad (18)$$

since from (7), $\rho^2 = x_1^2 + x_2^2 = z_1^2 + z_2^2$. Therefore

$$\angle\xi_\gamma \equiv 4\theta, \quad \xi_\gamma \triangleq \xi \cdot \mathrm{sign}(\gamma). \quad (19)$$

In practice, for single signal realizations composed of T samples, the ensemble or population averages in (13), (17), and (18) are estimated by their unbiased time or sample averages

$$\hat{\xi} = \frac{1}{T}\sum_{k=1}^{T} \rho_k^4 e^{j4\phi_k} = \frac{1}{T}\sum_{k=1}^{T} (z_1(k) + jz_2(k))^4 \quad (20)$$

$$\hat{\gamma} = \frac{1}{T}\sum_{k=1}^{T} \rho_k^4 - 8 = \frac{1}{T}\sum_{k=1}^{T} (z_1^2(k) + z_2^2(k))^2 - 8 \quad (21)$$

so that ξ_γ can be estimated from the above two

$$\hat{\xi}_\gamma = \hat{\xi} \cdot \mathrm{sign}(\hat{\gamma}). \quad (22)$$

As a result, the following angle estimator arises:

$$\hat{\theta}_{\mathrm{EML}} = \tfrac{1}{4}\mathrm{angle}(\hat{\xi}_\gamma) \quad (23)$$

in which function "angle($\hat{\xi}_\gamma$)" supplies the principal value of the argument of $\hat{\xi}_\gamma$. The acronym "EML" stands for *extended ML*. This name will find justification in the next section. Observe that no assumptions on the source pdf's have been made at all to arrive at this expression, which makes this estimator valid for *any* source distribution combination with *any* kurtosis value and *any* type of symmetry, as long as the sks is not zero. In the event of such a sum being zero, the magnitude of ξ would also be null, and θ could not be estimated from it any more. More specifically, we will see in Section VI when carrying out the asymptotic performance analysis that estimator (23) becomes inconsistent as $\gamma \to 0$.

Before closing this section, we remark that a $\pi/2$-rad estimation range for the unknown parameter θ suffices. If a rotation of arbitrary angle θ has been performed on the source jpdf, any counter-rotation of angle $\theta \pm (\pi/2)n$, for $n = 0, 1, 2, \ldots$, provides the source components, possibly with an axis interchange and sign variation, effects that are irrelevant as far as the separation is concerned. Hence, all possible values of θ can be reduced to the interval $[-\pi/4, \pi/4]$, yet still providing acceptable separation solutions. Since angle$(\cdot) \in [-\pi, \pi]$, the EML estimator (23) always yields angle estimates in the required range.

B. Connection with the Approximate ML Estimator

Now, let us relate (11) and (23). Equation (11) can be written as

$$\hat{\theta}_{\mathrm{ML}} = \tfrac{1}{4}\arctan\left(\frac{\mathrm{Im}(\hat{\xi})}{\mathrm{Re}(\hat{\xi})}\right) \quad (24)$$

where Re$(\cdot)$ and Im$(\cdot)$ represent, respectively, the real and imaginary part of their complex argument. First, let us assume that $\hat{\theta}_{\mathrm{ML}}$ is calculated without regard to the sign of the numerator and denominator of the arctangent argument. Contrasting the latter expression with (23), it turns out that both provide the same estimates when $\angle\hat{\xi}_\gamma \in [-\pi/2, \pi/2]$ or, assuming that no error occurs in the estimation of ξ_γ, i.e., $\hat{\xi}_\gamma \equiv \xi_\gamma$, and by virtue of (19), when $\theta \in [-\pi/8, \pi/8] = \Omega_{\pi/4}$. Estimator $\hat{\theta}_{\mathrm{ML}}$ is a particular case of $\hat{\theta}_{\mathrm{EML}}$ when $\theta \in \Omega_{\pi/4}$, and hence, $\hat{\theta}_{\mathrm{ML}}$ is indeed applicable under the same source pdf conditions as $\hat{\theta}_{\mathrm{EML}}$. When $\theta \notin \Omega_{\pi/4}$, estimator (11) is biased. This was expected since if the sign of the arctangent arguments is not taken into account, $\arctan(\cdot) \in [-\pi/2, \pi/2]$, and then, (11) only provides estimates in the range $[-\pi/8, \pi/8]$. Effectively, the arctangent function in (24) obtains $4\theta - \pi$ when $4\theta > \pi/2$ and $4\theta + \pi$ when $4\theta < -\pi/2$. Consequently, if $\theta > \pi/8$, the expected bias is $-\pi/4$ rad, whereas if $\theta < -\pi/8$, such bias becomes $\pi/4$. For $\theta \notin \Omega_{\pi/4}$, we then need to resort to (23).

In the second place, even if the sign of the arctangent arguments are considered, estimator (24) only provides valid separation solutions when the sks is positive. Else, a bias arises in $\hat{\theta}_{\mathrm{ML}}$ relative to $\hat{\theta}_{\mathrm{EML}}$, which is $\pi/4$ rad when $\theta < 0$ and $-\pi/4$ rad when $\theta > 0$. In this respect, the effect of the sign function in (22) and (23) can alternatively be seen as applying an extra 45° rotation to the estimate obtained for positive sks so that the potential bias is avoided (this notion will be revisited when giving a geometrical interpretation of the method in Section V). In essence, by virtue of the sign function, the method is made transparent to the source pdf tailness. To emphasize the key role played by the sks, notice that such a factor also appears itself in Harroy–Lacoume's development but vanishes when nulling the score function (see the term "Lgc$_4(\rho)$" in [9, Sec. 3, p. 171]). In conclusion, it can be claimed that $\hat{\theta}_{\mathrm{EML}}$ extends $\hat{\theta}_{\mathrm{ML}}$ in the source pdf's with which it is able to deal. The constraints (symmetric sources with the same distribution and positive kurtosis in the range $[0, 4]$) under which the approximate ML estimator was originally found to work are alleviated in the proposed method and reduced to a much weaker single condition on the sum of source kurtosis. Although $\hat{\theta}_{\mathrm{EML}}$ has not been derived from the ML principle, the term "ML" has some significance. The condition for $\hat{\theta}$ to be the ML estimator of θ under the assumptions of [9] may be restated as $\tan(4\hat{\theta}) = (\mathrm{Im}(\hat{\xi})/\mathrm{Re}(\hat{\xi}))$, which is fulfilled by (23), and hence, the latter is also the ML estimator of θ under the same assumptions.

C. Other Links

It is worth computing ξ as a function of the statistical properties of the whitened observations. On the one hand,

from (9) and (13)

$$\xi = \mathrm{E}[(z_1+jz_2)^4] = (\kappa_{40}^z - 6\kappa_{22}^z + \kappa_{04}^z) + j4(\kappa_{31}^z - \kappa_{13}^z) \quad (25)$$

and, on the other hand, from (9) and (16)–(18)

$$\xi = c_\gamma^{j4\theta} = e^{j4\theta}(\kappa_{40}^z + 2\kappa_{22}^z + \kappa_{04}^z). \quad (26)$$

The last two equations state that θ can be determined from the fourth-order cumulants of the decorrelated measurements. In particular, the modulus of (25) and (26) must be equal, which leads to the following relationship among the fourth-order cumulants of the whitened sensor outputs:

$$(\kappa_{31}^z - \kappa_{13}^z)^2 - \kappa_{22}^z(\kappa_{40}^z + \kappa_{04}^z) + 2(\kappa_{22}^z)^2 = 0. \quad (27)$$

This relationship was originally deduced by Comon in [5] following different and more algebraic arguments.

V. Geometrical Interpretation

There exists a fairly enlightening geometrical rationale behind the apparently obscure estimator expression (23) when the two sources are symmetrically distributed. Analogous geometric ideas are exposed in [4]. First, consider a mixture of two leptokurtic pdf's, such as, for instance, two Laplacian distributions. Since the kurtosis of both sources is positive, it is guaranteed that the sum is. The jpdf of the unit-power sources exhibits highest values along the lines defined by the angles $\phi' = n(\pi/2)$, $n = 0, 1, 2, 3$. Accordingly, the source scatter diagram has a maximum point concentration along the same lines. On the other hand, as explained in Section II-B, the jpdf of the whitened measurements looks like that of the sources, but rotated by θ degrees, θ being unknown. As a result, the new scatter diagram displays the highest density along $\phi = \theta + n(\pi/2)$, $n = 0, 1, 2, 3$. Now, assume that all the scatter diagram points are transformed according to

$$\rho_k\, e^{j\phi_k} \mapsto \rho_k^4 e^{j4\phi_k}. \quad (28)$$

In particular, all the previous points cluttered around $\phi = \theta + n(\pi/2)$ will now be gathered together along $\phi = 4\theta$, as graphically depicted in Fig. 3(a). In this way, it can be argued that a *centroid* calculated as the resultant mean point after the transformation, that is

$$\xi = \mathrm{E}[\rho_k^4 e^{j4\phi_k}] \quad (29)$$

will show that orientation. Therefore, (23) will provide the required angle with which the whitened jpdf is rotated with respect to the true source jpdf.

In the second place, let us now consider two platykurtic distributions playing the role of sources. Hence, their kurtosis sum is negative, as for two sinusoidal signals. In this case, the maximum concentration in the source scatter diagram occurs along the lines $\phi' = (\pi/4) + n(\pi/2)$, $n = 0, 1, 2, 3$. Therefore, in the scatter diagram of the signals after decorrelation, this high density will be shown along $\phi = \theta + (\pi/4) + n(\pi/2)$, $n = 0, 1, 2, 3$. These points will all clutter round $\phi = 4\theta + \pi$ when transformed as indicated by (28), and therefore, the centroid (29) will also have this orientation [Fig. 3(b)]. In conclusion, the centroid projected across the origin, i.e., $-\xi$, forms an angle 4θ with respect to the x axis. That is, the purpose of the "sign$(\cdot)$" term in the estimator expression (22) and (23) when $\kappa_{40}^x + \kappa_{04}^x < 0$ is to rotate the sample centroid $\hat{\xi}$ (20) by π rad so that its argument becomes an estimate of 4θ.

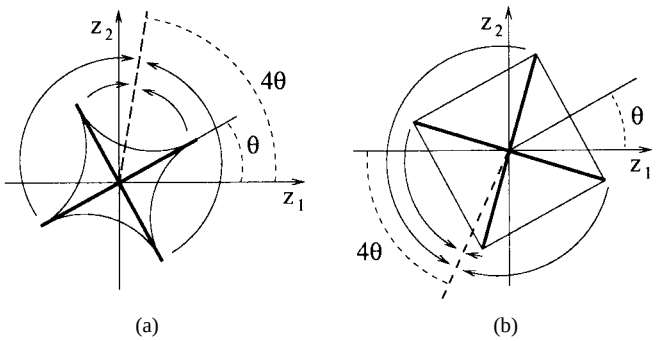


Fig. 3. Centroid location and angle estimation for a mixture of (a) two leptokurtic distributions and (b) two platykurtic distributions.

For combinations of platykurtic and leptokurtic distributions, the geometrical interpretation becomes more involved and less intuitive, but the results obtained in Section IV indicate that those cases can also be explained in terms of the above two instances, comprising sources with positive and negative kurtosis sum.

VI. Statistical Characterization

Having introduced the expression of the EML estimator and examined its behavior from a geometrical point of view, a study of its statistical properties is in order. The present section is devoted to analyzing its asymptotic performance, that is, its estimation characteristics when large samples are processed. Specifically, our main objective is to obtain an analytical expression for the estimator asymptotic pdf.

A. Bias

Before entering the asymptotic domain, observe first that the EML estimator (23) is *unbiased for any sample size* when at least one of the sources has a symmetric distribution. The proof of this claim flows along the same lines as that offered in [9, Sec. 4.1] for the approximate ML estimator (11). If there is an asymmetric source, let its corresponding axis form an angle θ with respect to the whitened observations frame of reference. Consider a set of observations (ρ_k, ϕ_k). From this set, parameters $\hat{\xi}^{(1)}$ and $\hat{\gamma}^{(1)}$ are computed by means of (20) and (21), respectively, yielding an angle estimator $\hat{\theta}_{\mathrm{EML}}^{(1)}$. Now, consider the set of symmetric (with respect to the true angle θ) observations $(\rho_k, 2\theta - \phi_k)$. It turns out that the new parameters are related to the previous ones via $\hat{\gamma}^{(2)} = \hat{\gamma}^{(1)}$ and $\hat{\xi}^{(2)} = \hat{\xi}^{(1)*} e^{j8\theta}$, symbol * denoting complex conjugation. Then, from (22) and (23), $\hat{\theta}_{\mathrm{EML}}^{(2)} = (2\theta - \hat{\theta}_{\mathrm{EML}}^{(1)})$.[1] By symmetry of the source(s), $\mathrm{E}[\hat{\theta}_{\mathrm{EML}}^{(1)}] = \mathrm{E}[\hat{\theta}_{\mathrm{EML}}^{(2)}]$. It follows that $\mathrm{E}[\hat{\theta}_{\mathrm{EML}}] = \theta$ for any sample size. We remark that for this result to hold, we do not need to assume that both sources possess equal symmetric distribution as in [9] but just that one source be symmetric.

[1] Plus a potential, irrelevant $\pm pi/2$ bias term.

B. Basic Asymptotic Results

If the sources are both asymmetric, the above result is not generally true. However, under basic regularity conditions, other interesting conclusions about estimator (23) can be drawn, which hold for all kind of source pdf. Effectively, assuming independent and identically distributed (i.i.d.) observations, the strong law of large numbers [13] ensures that $\hat{\xi} \xrightarrow[T\to\infty]{\text{wp1}} \xi$ and $\hat{\gamma} \xrightarrow[T\to\infty]{\text{wp1}} \gamma$, and in compliance with (19)–(23)

$$\hat{\theta}_{\text{EML}} \xrightarrow[T\to\infty]{\text{wp1}} \theta \tag{30}$$

where $\xrightarrow{\text{wp1}}$ denotes convergence with probability one. Hence, the estimator is *strongly consistent* and *asymptotically unbiased*, provided $\gamma \neq 0$.

C. Asymptotic pdf

In order to arrive at a more complete characterization of the estimator, consider first that from (8)

$$\hat{\xi} = \frac{1}{T}\sum_{k=1}^{T} \rho_k^4 e^{j4\phi_k} = \left\{ \frac{1}{T}\sum_{k=1}^{T} \rho_k^4 e^{j4\phi_k'} \right\} e^{j4\theta} = \hat{\xi}_x\, e^{j4\theta} \tag{31}$$

in which

$$\hat{\xi}_x = \frac{1}{T}\sum_{k=1}^{T} \rho_k^4 e^{j4\phi_k'} = \frac{1}{T}\sum_{k=1}^{T} (x_1(k) + jx_2(k))^4. \tag{32}$$

Then, the estimator becomes

$$\hat{\theta} = \tfrac{1}{4}\,\text{angle}(\hat{\xi}\cdot\text{sign}(\hat{\gamma})) = \theta + \tfrac{1}{4}\,\text{angle}(\hat{\xi}_x\cdot\text{sign}(\hat{\gamma})) \tag{33}$$

so that

$$\Delta\hat{\theta} = \hat{\theta} - \theta = \tfrac{1}{4}\,\text{angle}(\hat{\xi}_x\cdot\text{sign}(\hat{\gamma})). \tag{34}$$

That is, the bias of the estimator $\Delta\hat{\theta} = \hat{\theta} - \theta$ depends exclusively on the source signals (note that $\hat{\gamma}$ only depends on them too) and is independent of the specific value of the rotation angle θ. Our objective, then, is reduced to obtaining the pdf of the bias $\Delta\hat{\theta}$.

In the second place, let us assume that the sign of the sks is properly estimated, which is a reasonable assumption for a high enough sample size T and/or when $|\gamma| = |\kappa_{40}^x + \kappa_{04}^x|$ is not too small. Let us also assume that this sign is positive. Calling

$$\delta = 4\Delta\hat{\theta} = \text{angle}(\hat{\xi}_x) \tag{35}$$

we have

$$p_{\Delta\hat{\theta}}(\Delta\hat{\theta}) = 4p_\delta(4\Delta\hat{\theta}), \quad |\delta| \le \pi,\ |\Delta\hat{\theta}| \le \frac{\pi}{4} \tag{36}$$

so that it is sufficient to obtain $p_\delta(\delta)$. If the sks is negative, the actual pdf will be a π-radian circularly-shifted version of $p_\delta(\delta)$ above, due to the sign function in (34). At this point, it is convenient to decompose $\hat{\xi}_x$ in its real and imaginary parts

$$\hat{\xi}_x = \omega_1 + j\omega_2 \tag{37}$$

where ω_1 and ω_2 are obtained by developing the fourth power in (32)

$$\begin{aligned}\omega_1 &= \frac{1}{T}\sum_{k=1}^{T}(x_1^4(k) - 6x_1^2(k)x_2^2(k) + x_2^4(k)) \\ &= \frac{1}{T}\sum_{k=1}^{T}\tilde{\omega}_1(k)\end{aligned} \tag{38}$$

and

$$\omega_2 = \frac{1}{T}\sum_{k=1}^{T} 4(x_1^3(k)x_2(k) - x_1(k)x_2^3(k)) = \frac{1}{T}\sum_{k=1}^{T}\tilde{\omega}_2(k). \tag{39}$$

With the help of (9) and [15], the statistical properties of $\tilde{\omega}_1$ and $\tilde{\omega}_2$ are obtained as follows. In the first place, the mean and variance of $\tilde{\omega}_1$ are given by

$$\text{E}[\tilde{\omega}_1] = \mu_{40}^x + \mu_{04}^x - 6 = \kappa_{40}^x + \kappa_{04}^x = \gamma \tag{40}$$

$$\begin{aligned}\tilde{\sigma}_1^2 &\triangleq \text{Var}[\tilde{\omega}_1] \\ &= (\mu_{80}^x - (\mu_{40}^x)^2) - 12(\mu_{60}^x - \mu_{40}^x) + 36(\mu_{40}^x\mu_{04}^x - 1) \\ &\quad - 12(\mu_{06}^x - \mu_{04}^x) + (\mu_{08}^x - (\mu_{04}^x)^2) \\ &= (\kappa_{80}^x + \kappa_{08}^x) + 16(\kappa_{60}^x + \kappa_{06}^x) + 56(\kappa_{50}^x\kappa_{30}^x + \kappa_{05}^x\kappa_{03}^x) \\ &\quad + 34((\kappa_{40}^x)^2 + (\kappa_{04}^x)^2) + 160((\kappa_{30}^x)^2 + (\kappa_{03}^x)^2) \\ &\quad + 144(\kappa_{40}^x + \kappa_{04}^x) + 36\kappa_{40}^x\kappa_{04}^x + 192.\end{aligned} \tag{41}$$

Analogously, the first- and second-order statistical parameters of $\tilde{\omega}_2$ are

$$\text{E}[\tilde{\omega}_2] = 0, \tag{42}$$

$$\begin{aligned}\tilde{\sigma}_2^2 &\triangleq \text{Var}[\tilde{\omega}_2] \\ &= 16(\mu_{60}^x - 2\mu_{40}^x\mu_{04}^x + \mu_{06}^x) \\ &= 16((\kappa_{60}^x + \kappa_{06}^x) + 9(\kappa_{40}^x + \kappa_{04}^x) \\ &\quad + 10((\kappa_{30}^x)^2 + (\kappa_{03}^x)^2) - 2\kappa_{40}^x\kappa_{04}^x + 12).\end{aligned} \tag{43}$$

For two Gaussian sources, these statistical parameters become

$$\gamma = 0, \qquad \tilde{\sigma}_1^2 = \tilde{\sigma}_2^2 = 192. \tag{44}$$

The covariance of $\tilde{\omega}_1(k)$ and $\tilde{\omega}_2(k)$ is

$$\begin{aligned}\tilde{c}_{12} &\triangleq \text{E}[\tilde{\omega}_1\tilde{\omega}_2] - \text{E}[\tilde{\omega}_1]\text{E}[\tilde{\omega}_2] = -28(\mu_{50}^x\mu_{03}^x - \mu_{30}^x\mu_{05}^x) \\ &= -28(\kappa_{50}^x\kappa_{03}^x - \kappa_{30}^x\kappa_{05}^x)\end{aligned} \tag{45}$$

and therefore, these variables are not necessarily uncorrelated. However, the sole existence of a symmetric source makes $\tilde{c}_{12}$ vanish, and the same applies when both sources have identical distribution, for example.

Next, from (38) and (39), and assuming that each source signal is temporally white [a property that is inherited by statistics $\tilde{\omega}_1(k)$ and $\tilde{\omega}_2(k)$], the central limit theorem [13] states that for a high enough sample size T, ω_1 can be approximated by a Gaussian distribution of mean γ and variance $\sigma_1^2 = \tilde{\sigma}_1^2/T$, and ω_2 can be approximated by another Gaussian variable of zero mean and variance $\sigma_2^2 = \tilde{\sigma}_2^2/T$, i.e.,

$$\begin{aligned}\omega_1 &\xrightarrow[T\to\infty]{\text{d}} \text{N}(\gamma,\, \sigma_1^2) = \text{N}(\gamma,\, \tilde{\sigma}_1^2/T) \\ \omega_2 &\xrightarrow[T\to\infty]{\text{d}} \text{N}(0,\, \sigma_2^2) = \text{N}(0,\, \tilde{\sigma}_2^2/T)\end{aligned} \tag{46}$$

where $\xrightarrow{\text{d}}$ indicates convergence in distribution. On the other hand, for temporally white sources, $\text{E}[\tilde{\omega}_1(k)\tilde{\omega}_2(n)] =$

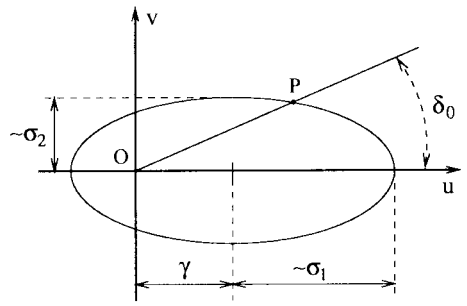


Fig. 4. Level curves of the asymptotic jpdf of ω_1 and ω_2, the source centroid coordinates in the complex plane.

$\mathrm{E}[\tilde{\omega}_1(k)]\mathrm{E}[\tilde{\omega}_2(n)] = 0$ for all $k \neq n$, and thus, the covariance of ω_1 and ω_2 is given by

$$c_{12} \triangleq \mathrm{E}[\omega_1\omega_2] - \mathrm{E}[\omega_1]\mathrm{E}[\omega_2] = \frac{\tilde{c}_{12}}{T}. \tag{47}$$

If $\tilde{c}_{12} = 0$, then ω_1 and ω_2 are uncorrelated. However, even if the coefficient $\tilde{c}_{12}$ is far from zero, c_{12} cancels as T tends to infinity, and hence, ω_1 and ω_2 are at least asymptotically uncorrelated in any case. In conclusion, the asymptotic jpdf of the real and imaginary parts of $\hat{\xi}_\gamma$ is given by the product of their marginal pdf's

$$p_{\omega_1\omega_2}(u, v) = \frac{1}{2\pi\sigma_1\sigma_2} \exp\left\{-\frac{(u-\gamma)^2}{2\sigma_1^2} - \frac{v^2}{2\sigma_2^2}\right\}. \tag{48}$$

The level curves of this pdf are ellipses of semi-axes proportional to σ_1 and σ_2 centered at $(\gamma, 0)$, as depicted in Fig. 4.

Having obtained the jpdf of ω_1 and ω_2, (35) and (37) suggest that we try the variable change

$$(u, v) \longrightarrow (r, \delta): \begin{cases} u = r\cos\delta \\ v = r\sin\delta \end{cases} \tag{49}$$

which yields for the marginal pdf of δ:

$$p_\delta(\delta) = \int_0^\infty \frac{r}{2\pi\sigma_1\sigma_2} \cdot \exp\left\{-\frac{(r\cos\delta - \gamma)^2}{2\sigma_1^2} - \frac{(r\sin\delta)^2}{2\sigma_2^2}\right\} dr. \tag{50}$$

Although this expression admits an analytical solution, as will be seen later on, let us first derive further simplifications to gain a deeper insight into its behavior as a function of parameters γ, σ_1, and σ_2. Fig. 4 can still be viewed as an approximation to a level curve of the integrand in (50). Roughly speaking, the value of p_δ at a particular angle δ_0 corresponds to the integral with respect to r of a slice of the function (48) along the angle $\delta = \delta_0$. Therefore, we could think of this value as being related to the distance from the origin to the point on the level curve located at $\delta = \delta_0$, i.e., distance OP in Fig. 4. As a result of this, it is easy to see now that for $\sigma_1 > \sigma_2$ and $\gamma > 0$, the distribution of δ will concentrate more closely around its mode $\delta = 0$ as (γ/σ_1) increases. The same outcome is obtained if the ratio (σ_1/σ_2) is increased. On the other hand, if $\sigma_1 < \sigma_2$, the distribution of δ will show two modes, neither at $\delta = 0$, though symmetric with respect to that point. Analogous results are obtained for $\gamma < 0$, but in that case, the pdf of δ will exhibit one mode at $\delta = \pi$ for $\sigma_1 > \sigma_2$ [this is the statistical counterpart of the geometric justification for the term "sign(·)" given in Section V] or two modes symmetrically located for $\sigma_1 < \sigma_2$. For $\sigma_1 < \sigma_2$, the two modes merge into only one as $(|\gamma|/\sigma_1)$ increases. As an example to illustrate this intuitive reasoning, consider two Gaussian sources, for which, according to (44), $\gamma = 0$ and $\sigma_1 = \sigma_2$; thus, the level curves becoming circumferences centered at the origin. This gives a $p_\delta(\delta)$ uniformly distributed between $-\pi$ and π.

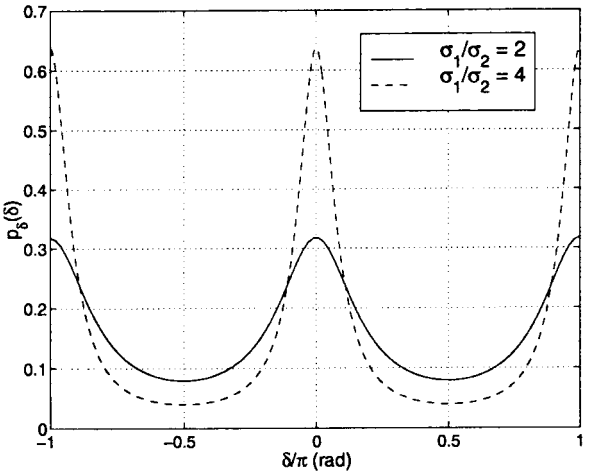
The main result derived from these geometrical considerations is that the distribution of δ (and, hence, of $\Delta\hat{\theta}$) becomes more "peaky" around 0 or π (depending on whether γ is positive or negative, respectively) as the ratios $(|\gamma|/\sigma_1)$ and (σ_1/σ_2) become larger. In order to give a first mathematical support to this claim, we can readily work out the analytical expression of p_δ when $\gamma = 0$:

$$p_\delta(\delta) = \frac{\frac{\sigma_1}{\sigma_2}(1 + \tan^2\delta)}{2\pi\left[1 + \left(\frac{\sigma_1}{\sigma_2}\tan\delta\right)^2\right]}. \tag{51}$$

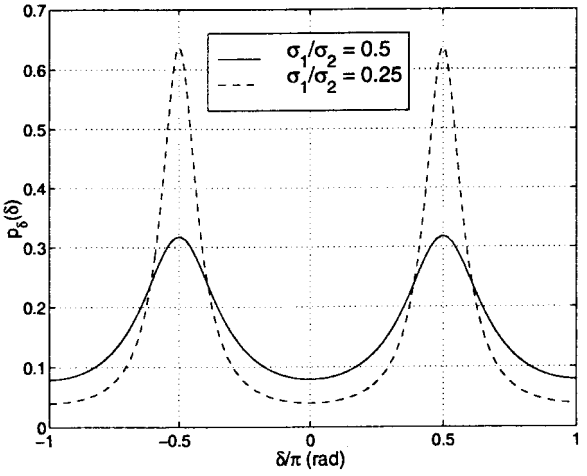
Then, for $\sigma_1 > \sigma_2$, the maximum value of p_δ is $(\sigma_1/\sigma_2)/(2\pi)$ and occurs at $\delta = 0, \pm\pi$, whereas the minimum is $1/(2\pi(\sigma_1/\sigma_2))$ at $\delta = \pm\pi/2$. The ratio of the maximum to the minimum value is thus $(\sigma_1/\sigma_2)^2$, confirming that the distribution becomes more "peaky" as (σ_1/σ_2) increases, as seen in Fig. 5(a). If, on the contrary, $\sigma_1 < \sigma_2$, the maximum value becomes $1/(2\pi(\sigma_1/\sigma_2))$ at $\delta = \pm\pi/2$ and the minimum $(\sigma_1/\sigma_2)/(2\pi)$ at $\delta = 0, \pm\pi$, as graphically described in Fig. 5(b). The pdf is flat if and only if $\sigma_1 = \sigma_2$, as occurs, for instance, for Gaussian sources [see (44)].

All this preceding analysis establishes approximately the behavior of p_δ as a function of (γ/σ_1) and (σ_1/σ_2). Now, it will be seen how these trends are actually exhibited by the analytical expression of the pdf. With the help of the symbolic mathematics package MAPLE™ and after some tedious algebraic manipulations, we can arrive from (50) at the following general closed-form analytical expression for the pdf of δ:

$$\begin{aligned} p_\delta(\delta) = {} & \frac{\frac{\sigma_1}{\sigma_2}(1 + \tan^2\delta)}{2\pi\left[1 + \left(\frac{\sigma_1}{\sigma_2}\tan\delta\right)^2\right]} \exp\left(\frac{-\gamma^2}{2\sigma_1^2}\right) \\ & + \frac{\gamma(1 + \tan^2\delta)\,\mathrm{sign}(\cos\delta)}{\sqrt{8\pi}\,\sigma_2\left[1 + \left(\frac{\sigma_1}{\sigma_2}\tan\delta\right)^2\right]^{3/2}} \\ & \cdot \exp\left(\frac{-\gamma^2\tan^2\delta}{2\sigma_2^2\left[1 + \left(\frac{\sigma_1}{\sigma_2}\tan\delta\right)^2\right]}\right) \\ & \cdot \left\{1 + \mathrm{erf}\left(\frac{\gamma\,\mathrm{sign}(\cos\delta)}{\sqrt{2}\,\sigma_1\sqrt{1 + \left(\frac{\sigma_1}{\sigma_2}\tan\delta\right)^2}}\right)\right\}. \end{aligned} \tag{52}$$



(a)



(b)

Fig. 5. Pdf of δ for $\gamma = 0$. (a) $\sigma_1/\sigma_2 > 1$. (b) $\sigma_1/\sigma_2 < 1$.

Observe that for $\gamma = 0$, the above formula simplifies into (51). The asymptotic results derived in Section VI-B may also be deduced from (52). Effectively, this pdf is, in all cases, symmetric with respect to $\delta = 0$, and hence, $\mathrm{E}[\delta] = 0$, that is, the EML estimator is unbiased for large enough sample size. In addition, as T increases, both σ_1 and σ_2 tend to zero, and therefore, the distribution of ω_1 and ω_2 tends to a point of probability mass 1 located at $(\gamma, 0)$ in the (ω_1, ω_2) plane. As a consequence, the EML estimator shows strong consistency, as long as $\gamma \neq 0$, i.e., $\kappa_{40}^x + \kappa_{04}^x \neq 0$. For $\gamma = 0$, since $(\sigma_1/\sigma_2) = (\tilde{\sigma}_1/\tilde{\sigma}_2)$, which is a constant for a given pair of source distributions, according to (51), the estimator pdf does not vary as the sample size increases. Therefore, the estimator loses its consistency property when the sks is null.

Finally, in order to circularly shift the pdf expression by π rad, which is necessary when $\kappa_{40}^x + \kappa_{04}^x < 0$, as discussed at the beginning of this section, it suffices to use $|\gamma|$ instead of γ in (52).

VII. Simulation Examples

The purpose of this section is to illustrate the properties of the suggested estimator and to test the quality of the theoretical asymptotic analysis. The effects of a whitening process prior to the application of the estimator are also examined. Finally, an uncomplicated extension to the general BSS scenario composed of more than two signals is tested. This extension is based upon the ideas of [6].

TABLE I
Results of the Simulations with the EML estimator. Signals are Composed of 5000 i.i.d. Samples. Each Four-Element Column Displays, from Top to Bottom, the Mean, the Standard Deviation, the Maximum, and the Minimum Value Obtained for the Corresponding Parameter and Signal Combination Over 100 Monte Carlo Runs. Angle Values $(\hat{\theta} - \theta)$ Are Expressed in Degrees

DISTR.	BOTH SYMMETRIC			SYMMET. - ASYMMET.			BOTH ASYMMETRIC		
TYPE	PAIR	$\hat{\gamma}$	$\hat{\theta} - \theta$	PAIR	$\hat{\gamma}$	$\hat{\theta} - \theta$	PAIR	$\hat{\gamma}$	$\hat{\theta} - \theta$
SHORT-	Uni-	−2.397	0.000	Uni-	−1.790	−0.051	Shsk-	−1.199	−0.004
-SHORT	-Uni	0.032	0.005	-Shsk	0.052	0.262	-Shsk	0.066	0.061
		−2.283	0.017		−1.619	0.668		−1.068	0.158
		−2.462	−0.014		−1.933	−0.888		−1.421	−0.237
SHORT-	Uni-	1.795	−0.595	Uni-	4.869	−0.186	Shsk-	5.353	−0.036
-LONG	-Lap	0.677	1.760	-Exp	1.202	0.601	-Exp	1.126	0.328
		6.593	3.487		8.869	1.083		10.807	0.824
		0.886	−4.801		2.605	−1.813		3.391	−0.719
LONG-	Lap-	5.900	−0.006	Exp-	8.915	−0.117	Exp-	6.087	−0.003
-LONG	-Lap	0.595	0.195	-Lap	1.599	0.339	-Ray	1.181	0.402
		7.089	0.548		16.622	0.619		11.570	0.899
		4.684	−1.168		6.031	−1.212		4.205	−1.405

A. Two-Source Two-Sensor Scenario

In the first place, several Monte Carlo simulations have been run to test and compare the new estimator performance. Three different combinations of source pdf's, with regard to their tail or kurtosis sign, have been considered (both short, long-short, and both long), together with three different symmetry combinations (both symmetric, symmetric-asymmetric, and both asymmetric). That makes a total of nine distinct distribution pairs for the source signals. The actual source pdf's employed are (in parentheses the abbreviations used in the results table): uniform ("Uni"), exponential ("Exp"), Laplacian ("Lap"), Rayleigh ("Ray"), and a short-skewed distribution ("Shsk"), which is simply an asymmetric triangular pdf.

For each distribution pair, zero-mean unit-variance source signal realizations are created from 5000 i.i.d. samples, and possible remains of statistical dependence up to fourth-order are removed by using the ICA procedure developed in [6] (denoted herein as ICA-HOEVD method). Then, a rotation matrix of fixed angle $\theta = 15°$ is applied to the set of source signals, giving a hypothetical set of whitened sensor outputs. The estimation of the rotation angle is carried out through (22) and (23), where the sample estimates $\hat{\xi}$ and $\hat{\gamma}$ are obtained from the whitened sensor data by means of (20) and (21). From each realization, two parameters are computed: the bias $(\hat{\theta} - \theta)$ expressed in degrees and the values of $\hat{\gamma}$ to check how they approximate the sks $(\kappa_{40}^x + \kappa_{04}^x)$. The mean, standard deviation, maximum, and minimum value for these two parameters computed over 100 Monte Carlo runs are summarized in Table I.

In light of these results, the new estimator can be judged as successful in estimating the unknown rotation angle θ. This

TABLE II
COMPARISON OF SEVERAL ESTIMATORS. EML, ML WITHOUT REGARD TO ARGUMENT SIGNS (ML-a), ML CONSIDERING ARGUMENT SIGNS (ML-b), AND COMON'S FORMULA [5] (CF). SOURCES ARE COMPOSED OF 1000 INDEPENDENT UNIFORMLY DISTRIBUTED SAMPLES. EACH FOUR-ELEMENT COLUMN SHOWS THE MEAN, THE STANDARD DEVIATION, THE MAXIMUM, AND THE MINIMUM VALUE OBTAINED FOR THE BIAS $(\hat{\theta} - \theta)$ OVER 100 MONTE CARLO RUNS. ALL VALUES ARE EXPRESSED IN DEGREES. EXPERIMENTS FOR DIFFERENT ANGLE VALUES ARE INDEPENDENT

METHOD / ANGLE	EML	ML-a	ML-b	CF
$\theta = 20°$	0.006	0.006	−44.994	1.382
	0.033	0.033	0.033	1.145
	0.123	0.123	−44.877	5.117
	−0.056	−0.056	−45.056	−1.217
$\theta = -30°$	0.002	45.002	45.002	−0.698
	0.034	0.034	0.034	0.614
	0.088	45.088	45.088	0.994
	−0.109	44.891	44.891	−2.527

is manifested in the low bias and variance of $\hat{\theta}$. In addition, $\hat{\gamma}$ approximates the sks very accurately in all cases (e.g., for uniform-exponential sources, its mean is approximately equal to $-1.2 + 6 = 4.8$).

It is interesting to compare the variance obtained here with the bound given by (12). For example, for two Laplacian distributions ($\kappa_{40}^x = \kappa_{04}^x = 3$) and $T = 5000$ samples, (12) predicts a lower bound for the standard deviation of 0.468°. However, from Table I, the empirical standard deviation obtained for this case turns out to be just 0.195°, which is clearly below that limit. This outcome confirms the initial suspicion that the conditions of the Gram–Charlier expansion are too restrictive and, indeed, not necessary to be met if estimator (11) is to be applied.

For the sake of comparison with the ML estimator (11), and in order to illustrate the comments made in Sections IV-B and V, Table II shows the angle bias results obtained by both estimators in a further simulation. This time, the two independent sources are composed of 1000 samples drawn from a uniform distribution. The simulation is run under the conditions cited earlier, generating independent source-realizations over 100 Monte Carlo iterations but feeding the same data into both methods at each iteration. For the ML method, two cases are considered: when the sign of the arctangent arguments are not taken into account ("ML-a") and when they are ("ML-b"). The first row corresponds to a true rotation angle of $\theta = 20°$. As predicted in Section IV-B, the outcome is identical for EML and ML-a because $|\theta| < (\pi/8)$, but a $-45°$ bias appears in the results of ML-b since the sources have negative kurtosis and $\theta > 0$. The second row of Table II shows the results for a new independent set of source realizations and a new angle, $\theta = -30° < -22.5°$. Note that now, an exact 45° bias comes up in the results of both ML estimators, just as anticipated in Section IV-B: in ML-a because $|\theta| > (\pi/8)$ and in ML-b because the sks is negative and $\theta < 0$. The final column summarizes the performance under the same source realizations of the closed-form formula for the estimation of θ found by Comon in [5]. EML results are consistently more reliable than those offered by this latter analytical expression.

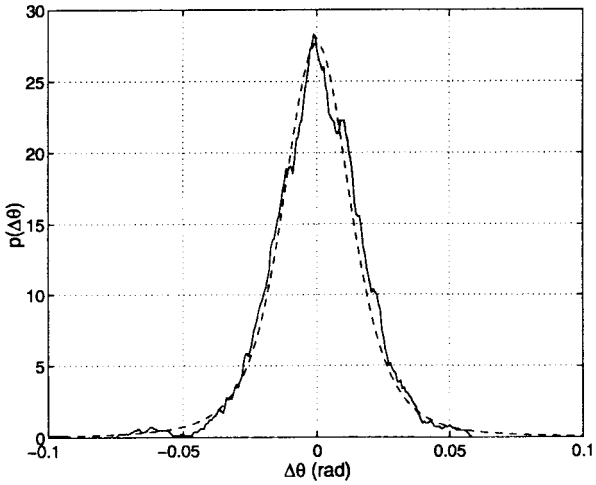


Fig. 6. Pdf of the bias for exponential and Rayleigh sources, $T = 1000$ samples per signal, 1000 signal realizations. Solid line: naive density estimate, window width = $8.3 \cdot 10^{-3}$ rad. Dashed line: pdf estimate from analytical expression (52).

B. Asymptotic Behavior

In the second place, some other experiments have been carried out to test the validity and accuracy of the analytical expression obtained for the pdf of $\Delta\hat{\theta}$. In the first simulation, the source distributions are chosen to be the exponential and Rayleigh distributions, respectively, yielding a positive sum of kurtosis. Each signal realization is composed of 1000 samples, and from each of them, another realization of the bias $\Delta\hat{\theta}$ is directly calculated from (21), (32), and (34). From 1000 realizations of $\Delta\hat{\theta}$, its pdf is estimated by using a standard method (the naive pdf estimator [14]) with window width $h = 8.3 \cdot 10^{-3}$ rad. This pdf estimate corresponds to the solid curve in Fig. 6. From the realizations of the real and imaginary parts of $\hat{\xi}_x$, parameters γ, σ_1, and σ_2 are approximated by the sample estimates of $\mathrm{E}[\omega_1]$, $\mathrm{Var}[\omega_1]$, and $\mathrm{Var}[\omega_2]$, which are denoted by $\hat{\gamma}'$, $\hat{\sigma}_1^2$, and $\hat{\sigma}_2^2$, respectively.[2] The exact values obtained for the parameters are $\hat{\gamma}' = 6.299$, $\hat{\sigma}_1 = 2.563$, and $\hat{\sigma}_2 = 0.363$; therefore, $(\hat{\gamma}'/\hat{\sigma}_1) = 2.458$, and $(\hat{\sigma}_1/\hat{\sigma}_2) = 7.063$. Substituting these coefficients in (52) and bearing in mind (36), we obtain the analytic pdf estimate represented by the dashed line in Fig. 6. The similarity between the two curves is remarkable.

In the next experiment, a negative kurtosis sum is tested by means of a uniform and a binary source. Under exactly the same conditions as above, the solid line in Fig. 7(a) shows the naive pdf estimate directly obtained from the angle estimates. Now, $\hat{\gamma}' = -3.192$, $\hat{\sigma}_1 = 0.040$, and $\hat{\sigma}_2 = 0.142$, giving $(|\hat{\gamma}'|/\hat{\sigma}_1) = 80.211$, and $(\hat{\sigma}_1/\hat{\sigma}_2) = 0.280$, for which the analytical expression of the pdf produces the dashed-line curve in Fig. 7(a). Again, there is an outstanding resemblance between the two plots. Since $\hat{\sigma}_1 < \hat{\sigma}_2$, two modes were expected, but $(|\hat{\gamma}'|/\hat{\sigma}_1)$ is so high as to prevent them from turning up. Fig. 7(b) displays the results for the same source

[2] We remark that the sample estimate $\hat{\gamma}'$ of $\mathrm{E}[\omega_1]$ defined in (40) is different from the sample estimate $\hat{\gamma}$ (21) of γ defined in (17) and (18), despite the expected value of both estimates being $\gamma = \kappa_{40}^x + \kappa_{04}^x$. The former involves several signal realizations, whereas the latter comprises a single signal realization.

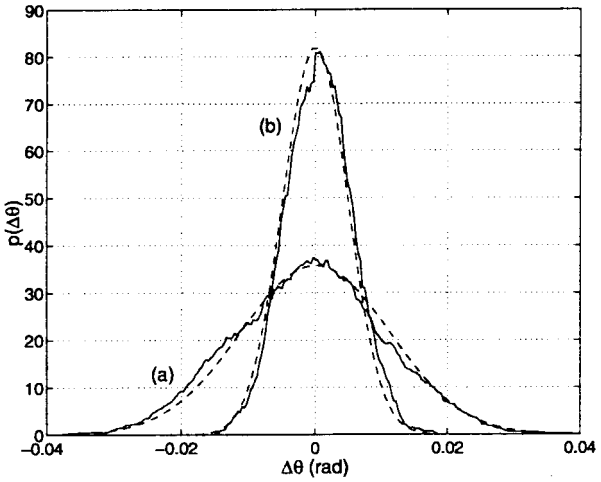


Fig. 7. Pdf of the bias for uniform and binary sources. (a) $T = 1000$ samples per signal, 1000 signal realizations. Solid line: naive density estimate, window width $= 8.3 \cdot 10^{-3}$ rad. Dashed line: pdf estimate using analytical expression (52). (b) $T = 5000$ samples per signal, 1000 signal realizations. Solid line: naive density estimate, window width $= 4.2 \cdot 10^{-3}$ rad. Dashed line: pdf estimate using analytical expression (52).

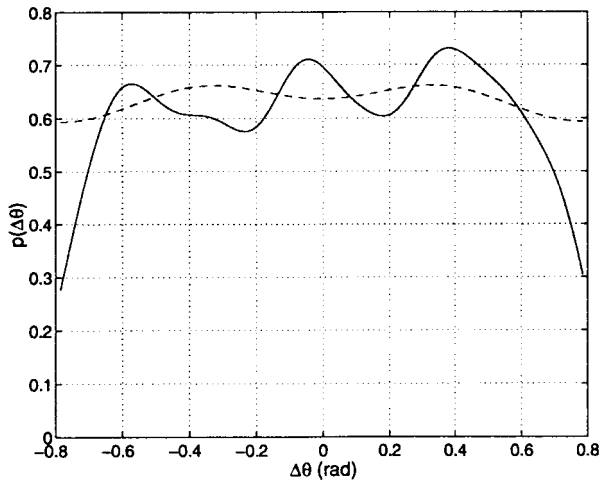


Fig. 8. Pdf of the bias for two Gaussian sources, $T = 5000$ samples per signal, 1000 signal realizations. Solid line: Gaussian kernel density estimate, function width $= 78.5 \cdot 10^{-3}$ rad. Dashed line: pdf estimate from analytical expression (52).

distributions but for a sample size of $T = 5000$. The new values for the pdf parameters are $\hat{\gamma}' = -3.199$, $\hat{\sigma}_1 = 0.016$, and $\hat{\sigma}_2 = 0.062$, i.e., $(|\hat{\gamma}'|/\hat{\sigma}_1) = 197.438$, and $(\hat{\sigma}_1/\hat{\sigma}_2) = 0.261$. The values of $\hat{\sigma}_1$ and $\hat{\sigma}_2$ decrease by a factor of $\sqrt{5}$, as (46) predicts and, hence, is the variance reduction in the obtained pdf.

Finally, two 5000-sample Gaussian signals are chosen to play the role of sources. For 1000 signal realizations, both a standard estimate and the analytical solution of the bias pdf are plotted in Fig. 8. They approximate a uniform distribution, just as anticipated in Section VI. The standard pdf estimation procedure employed in this example is the kernel method [14], with Gaussian kernel of width $78.5 \cdot 10^{-3}$ rad. For this source combination, $\hat{\gamma}' = -0.005$, $\hat{\sigma}_1 = 0.194$, and $\hat{\sigma}_2 = 0.201$, i.e., $(|\hat{\gamma}'|/\hat{\sigma}_1) = 0.028$, and $(\hat{\sigma}_1/\hat{\sigma}_2) = 0.966$. Since $\hat{\gamma}' < 0$, it was assumed that $\gamma < 0$ to get the analytical solution. Note that $\hat{\sigma}_1 \approx \hat{\sigma}_2 \approx \sqrt{192/(5 \cdot 10^3)} = 0.196$, as deduced from (44).

C. Influence of Prewhitening

All the experiments reported up to now assess the estimator performance when only the higher order processing part of the separation is considered. In a practical separation system, however, the EML estimator (23) would have to be applied after the sensor output has been whitened, as explained in Section II-A. Intuitively, the estimator performance when second-order processing is included would differ from the results when it is not since the uncertainties generated by the prewhitening would add to those introduced by the EML estimator on its own. In order to empirically check this claim, we can compare the variance of the angle estimates obtained with and without prewhitening for fixed mixing matrices. Fig. 9 shows how these variances behave as the sample size increases for two uniformly distributed sources. The solid line represents the variance of the EML angle estimates for a 30° rotation on the sources. The dash-dotted line (with "×" markers) also includes the prewhitening applied on the same sensor-output realizations followed by EML estimation. The dashed line corresponds to a particular mixing matrix with nonorthogonal transfer vectors with second plus fourth-order separation stages. The prewhitening process is accomplished by means of the SVD of the sensor-output data matrix, as done, e.g., in [6]. The variances at each sample-size value are obtained by sample averaging over ν Monte Carlo runs with $\nu T = 5 \cdot 10^5$. The same source realizations are used in all three cases. As expected, for i.i.d. signals, the variances decrease as T^{-1}. This is endorsed by the dotted lines of Fig. 9, which represent least-squares approximations to the curves of the form "constant$/T$." The results for the first and second curves are identical, whereas for the third, the results are slightly higher. Second-order processing introduces errors in the estimation of the whitening matrix that translate into inaccuracies in the resulting orthogonal rotation operating on the sources. Consequently, the variance of the angle estimates increases with regard to prewhitening. Results of Fig. 9 also reveal that the variance increase seems to depend on the mixing matrix conditioning. Nonetheless, further simulations show that this dependence is not too acute, but the variance trajectory remains essentially the same and very close to that without prewhitening over a wide range of mixing matrix conditioning. In particular, when the mixing matrix is orthogonal, the effects of the second-order processing stage on the performance of the EML estimator are totally negligible.

D. Extension to More than Two Sources and Sensors

A final simulation demonstrates how the method propounded here, which is, in principle, only valid for the simplified two-source two-sensor BSS scenario, is also useful in separating more than two source signals from more than two measurements. The idea is inspired by [6] and basically consists of operating pairwise and applying (23) to every whitened-signal pair in turn over several sweeps until convergence. Geometrically, this corresponds to performing

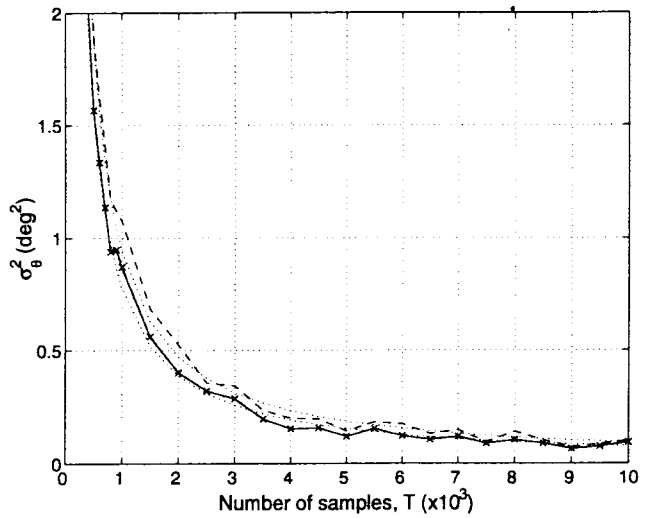


Fig. 9. Influence of prewhitening on the variance of the EML estimator for two i.i.d. sources with uniform distribution. Solid line: EML estimation only, orthogonal mixing matrix. Dash-dotted line (with "×" markers): prewhitening plus EML estimation, orthogonal mixing matrix. Dashed line: prewhitening plus EML estimation, nonorthogonal mixing matrix. Dotted lines: $O(T^{-1})$ LS approximations.

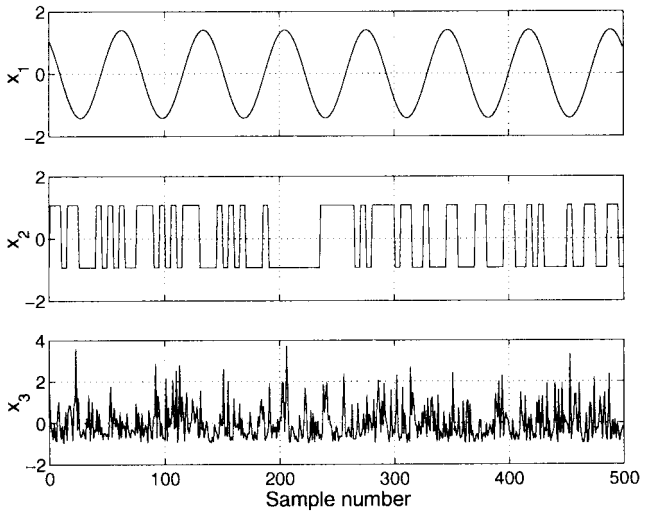


Fig. 10. Realization of three sources: a sinusoid, a binary sequence, and an exponentially distributed random process.

at each iteration an elementary planar rotation on the two-dimensional scatter plot of the current signal pair. Fig. 10 displays three sources to be separated by means of this straightforward extension:

1) a sinusoid;
2) a binary sequence,
3) an exponentially distributed signal;

all composed of 1000 samples of which only the first 500 are shown. The three-sensor output resulting from a particular instantaneous linear mixture of those three sources appear in Fig. 11. Again, the SVD is used for prewhitening. The pairwise extension of the EML procedure supplies the sources shown in Fig. 12, whereas Fig. 13 displays the sources obtained by the ICA-HOEVD method of [6]. Both set of estimated source waveforms are nearly identical, apart from

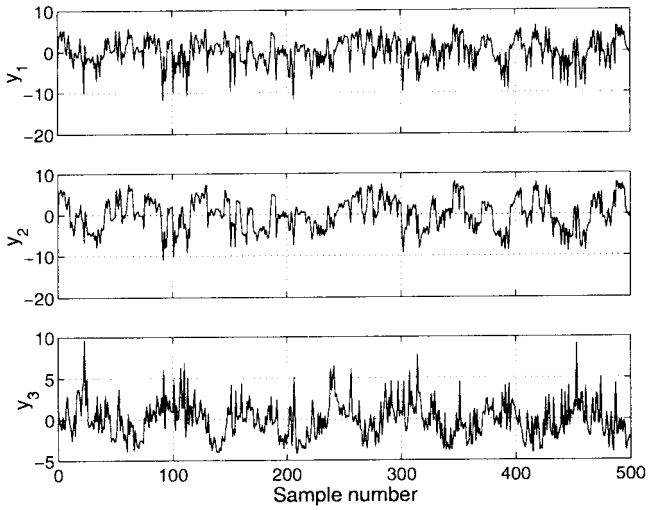


Fig. 11. Instantaneous linear mixture of the source signals shown in Fig. 10.

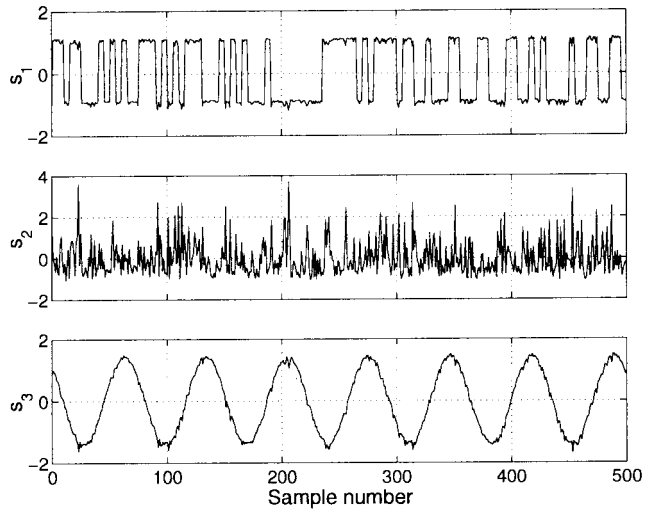


Fig. 12. Sources retrieved by the EML method from the mixture shown in Fig. 11.

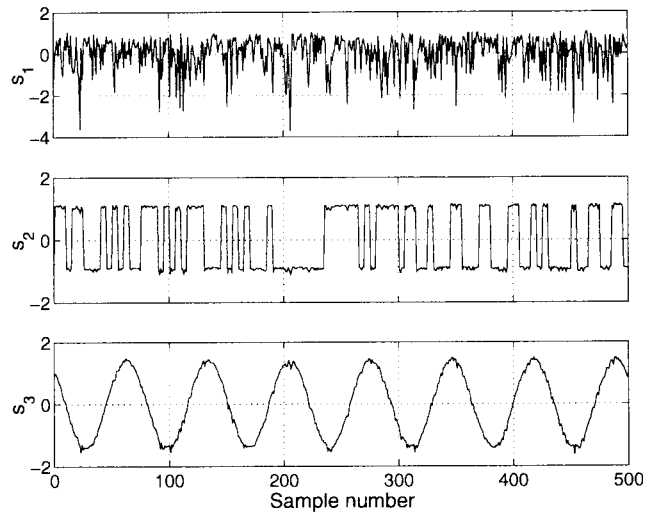


Fig. 13. Sources extracted by the ICA-HOEVD method from the mixture shown in Fig. 11.

the different signal arrangements and unitary scale factors. Additional experiments on a number of sources and measurements higher than two validate the implemented extension.

VIII. Summary, Conclusions, and Outlook

The approximate ML approach proposed in [9] has been considered. Due to the validity domain of the Gram–Charlier expansion and certain assumptions made during its development, the applicability of this ML estimator is restricted to symmetric sources with the same distribution and normalized kurtosis lying in the range [0, 4].

Along the same lines, a new expression for the estimation of the rotation angle is found, also based on the fourth-order statistics of the data. Interestingly, to arrive at the new expression, neither pdf expansions nor assumptions on the source pdf's are necessary. Consequently, the proposed estimator is valid for any source pdf combination with arbitrary kurtosis value and symmetry. The only condition is that the sks must be different from zero. This estimator can be regarded as an extension of the ML solution of [9], and hence, the name *extended ML* (EML) estimator.

The *modus operandi* of the EML estimator has been explained in terms of the scatter plots of the signals involved. This geometrical rationale behind the method becomes especially apparent when the sources have the same distribution with symmetric pdf. From this standpoint, it can be argued that the estimator exploits the specific symmetries existing in the jpdf of independent variables: the sources.

A closed-form analytical expression for the asymptotic pdf of the estimator bias when i.i.d. observations are processed has been obtained. From this formula, it has also been seen how the estimator behaves as some parameters (γ, σ_1, and σ_2) vary. These parameters depend, basically, on the higher order statistics (up to the eighth order) of the source signals. It is shown that the larger the $(|\gamma|/\sigma_1)$ and (σ_1/σ_2), the more precise becomes the estimation. In the i.i.d. case, the estimator exhibits strong consistency (except when the sks is null) and asymptotic unbiasedness. When there is a symmetric source, the estimator is unbiased for any sample size.

Several simulations prove the validity of the theoretical results, offering satisfactory angle estimation in a wide variety of source pdf combinations, with different tails (i.e., kurtosis value and sign) and symmetries. Additional experiments endorse the validity of the analytical solution found for the estimator asymptotic pdf. The potential of the presented method is substantiated through an extension to the general BSS setup composed of more than two mixtures of more than two sources. Performance is also very satisfactory for this extension.

To complete the statistical characterization and as a first point of further research, the dependence of the pdf behavior on the above parameters needs to be translated into particular conditions on the source signals in order to find out which distributions are more easily treated by the proposed estimator. In an initial step to achieve this goal, it would be desirable to obtain the estimator variance as a function of the mentioned parameters.

It has also been proved that spurious modes can appear in the asymptotic pdf of the EML estimator when $\sigma_1 < \sigma_2$ if $(|\gamma|/\sigma_1)$ is not large enough, thereby degrading the estimation quality. In which situations this occurs deserves to be investigated as well. The estimator has been proven to turn inefficient when the sks approaches zero. In order to enhance the performance in such scenarios, an additional expression may be utilized instead. This alternative fourth-order estimator is currently under study, together with an appropriate decision rule to make an optimal choice between both expressions given a signal set. Preliminary results for this combined estimation strategy are very encouraging. Alternative generalizations to the case of more than two signals are being considered as well.

Noise has been disregarded in this paper, yet the noise impact on the estimator performance is another important issue to be explored.

References

[1] 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.

[2] J.-F. Cardoso and A. Souloumiac, "Blind beamforming for non-Gaussian signals," *Proc. Inst. Elect. Eng. F*, vol. 140, no. 6, pp. 362–370, Dec. 1993.

[3] J.-F. Cardoso and B. H. Laheld, "Equivariant adaptive source separation," *IEEE Trans. Signal Processing*, vol. 44, pp. 3017–3030, Dec. 1996.

[4] I. J. Clarke, "Direct exploitation of non-Gaussianity as a discriminant," in *Proc. EUSIPCO*, Rhodes, Greece, Sept. 1998, vol. IV, pp. 2057–2060.

[5] P. Comon, "Separation of sources using higher-order cumulants," in *Proc. SPIE Vol. 1152 Adv. Algorithms Architectures Signal Process. IV*, 1989, pp. 170–181.

[6] ——, "Independent component analysis, A new concept?," *Signal Process.*, vol. 36, no. 3, pp. 287–314, Apr. 1994.

[7] M. Gaeta and J.-L. Lacoume, "Sources separation without *a priori* knowledge: The maximum likelihood solution," in *Proc. EUSIPCO*, Barcelona, Spain, 1990, pp. 621–624.

[8] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 2nd ed. Baltimore, MD: Johns Hopkins Univ. Press, 1989.

[9] F. Harroy and J.-L. Lacoume, "Maximum likelihood estimators and Cramer–Rao bounds in source separation," *Signal Process.*, vol. 55, pp. 167–177, Dec. 1996.

[10] J.-L. Lacoume and P. Ruiz, "Separation of independent sources from correlated inputs," *IEEE Trans. Signal Processing*, vol. 40, pp. 3074–3078, Dec. 1992.

[11] D. T. Pham, "Blind separation of instantaneous mixture of sources via an independent component analysis," *IEEE Trans. Signal Processing*, vol. 44, pp. 2768–2779, Nov. 1996.

[12] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propagat.*, vol. AP-34, pp. 276–280, Mar. 1986.

[13] R. J. Serfling, *Approximation Theorems of Mathematical Statistics*. New York: Wiley, 1980.

[14] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*. London, U.K.: Chapman & Hall, 1986, Monographs on Statistics and Applied Probability.

[15] A. Stuart and J. K. Ord, *Kendall's Advanced Theory of Statistics*, 6th ed. London, U.K.: Edward Arnold, 1994, vol. I.

[16] L. Tong, R. Liu, V. C. Soon, and Y.-F. Huang, "Indeterminacy and identifiability of blind identification," *IEEE Trans. Signal Processing*, vol. 38, pp. 499–509, May 1991.

[17] L. Tong, Y. Inouye, and R. Liu, "Waveform-preserving blind estimation of multiple independent sources," *IEEE Trans. Signal Processing*, vol. 41, pp. 2461–2470, July 1993.

[18] J. Vanderschoot *et al.*, "Two methods for optimal MECG elimination and FECG detection from skin electrode signals," *IEEE Trans. Biomed. Eng.*, vol. BME-34, pp. 233–243, Mar. 1987.

Vicente Zarzoso (S'93) was born in Valencia, Spain, on September 12, 1973. He attended the Universidad Politécnica de Valencia for the first four years of his degree and the University of Strathclyde, Glasgow, U.K., on an Erasmus exchange program for the final year of his degree. In July 1996, he received the B.Eng. degree with the highest distinction (Premio Extraordinario de Terminación de Estudios) in telecommunications engineering from the Universidad Politécnica de Valencia. He was awarded a scholarship by the University of Strathclyde to study toward the Ph.D. degree, the first year of which was also partly funded by a grant from the Defence Evaluation and Research Agency (DERA) of the U.K. Since March 1999, he has been with the University of Liverpool, Liverpool, U.K., continuing his doctoral studies.

His research interests include blind signal separation, higher order statistics, and statistical signal and array processing and their application to biomedical problems and communications.

Asoke K. Nandi (SM'96) received the Ph.D. degree from Trinity College, University of Cambridge, Cambridge, U.K., in 1979.

He held a scholarship at Trinity College during his doctoral studies at the Cavendish Laboratory, Cambridge. Since then he has held several research positions: Associateship with Rutherford Appleton Laboratory, Oxfordshire, U.K. and in the European Organization for Nuclear Research (CERN), Geneva, Switzerland, Advanced Fellowship in the Department of Physics, Queen Mary College, London, U.K., and in the Department of Nuclear Physics, University of Oxford, U.K. In 1987, he joined the Imperial College, London, as the Solartron Lecturer with the Signal Processing Section, Electrical Engineering Department. In 1991, he joined the Signal Processing Division of the Electronic and Electrical Engineering Department, University of Strathclyde, as a Senior Lecturer. In 1995, he was appointed Reader, and he has been a Professor since 1998. Since March 1999, he has held the David Jardine Chair of Electrical Engineering at the University of Liverpool, Liverpool, U.K. In 1983, he was part of the UAI team at CERN that discovered the three fundamental particles known as W^+, W^-, and Z^0, providing the evidence for the unification of the electromagnetic and weak forces. Currently, he leads a strong group of doctoral and post-doctoral researchers with interests in the areas of nonlinear systems and non-Gaussian signal processing as well as communications research. With his group, he has been carrying out research in machine condition monitoring, signal modeling, system identification, communications signal processing, time delay estimation, ECG signals, underwater sonar, applications of artificial neural networks, ultrasonics, blind source separation, and blind deconvolution. He has authored or coauthored over 200 technical publications, including two books and more than 90 journal papers.

Prof. Nandi is a Fellow of the Cambridge Philosophical Society, the Institution of Electrical Engineers, the Institute of Mathematics and its Applications, and the Institute of Physics. He is also a Member of the British Computer Society and the European Association for Signal Processing.