ICA IN SIGNALS WITH MULTIPLICATIVE NOISE USING FOURTH-ORDER STATISTIC

D. Blanco, B. Mulgrew, D. P. Ruiz and M. C. Carrión

Dpt. of Applied Physiscs, University of Granada
Avenida Severo Ochoa s/n, Granada, Spain
(dblanco druiz mcarrion)@ugr.es

Institute for Digital Communication, Univ. of Edinburgh
The King's Buildings, EH9 3JL, Edinburgh, UK
Bernie.Mulgrew@ee.ed.ac.uk

ABSTRACT

An extension of Independent Component Analysis (ICA) to the situation when the mixture of signals is contaminated by multiplicative noise is proposed in this paper. The ICA methods search for the most independent output after a linear transformation of the data vector. If the ICA model is followed by these data, the result of this search is the inverse of the unknown mixture. On the other hand, if there is multiplicative noise the model is not followed and the previous search does not obtain the wanted matrix. However, when the inverse of the mixture is applied to the noisy data, the output possesses a specific statistical structure that can be used to solve the problem. This paper exploits this structure up to fourth order in the statistic to design a method that is able to find the mixture in presence of multiplicative noise, improving greatly the results of the standard ICA methods in this situation, without any limitation in the nature of the sources or the noise.

1. INTRODUCTION

The aim of Independent Component Analysis (ICA) is to find the linear transformation that produces the most independent output of some data. If these data are linear mixture of independent sources, which is called *ICA model* and the mixture the *mixing matrix*, the ICA methods can recover the inverse of the mixing matrix and the original sources. In order to do that, the ICA methods do not need any information about the mixture or the sources, so the obtaining is "blind". ICA was first formalized in [?], and since then it has been applied to many fields, e.g. speech enhancement, medical signal processing, image analysis, telecommunications and financial series, among others. ICA has been extended to other situations where the ICA model is not exactly satisfied but the assumption of independent sources allows the generalization of the analysis. The existence of a bias term, exclusive classes, non-linearities in the mixture, non-instantaneous mixtures, additive noise, etc. are some of these situations. In this paper, the goal is to extend the ICA ideas to the situation when the linear mixture of independent sources is contaminated by multiplicative noise, which will be called *multiplicative ICA model* or *MICA model*.

Multiplicative noise appears in problems as ultrasound images, astrophysics, laser imagery, synthetic aperture radar (SAR) images, etc. Therefore, in these signals, although the data without noise can follow the ICA model, the recorded data will never do, so the applicability of the ICA methods is reduced. Such applications have been done mainly in coherent images, where the multiplicative noise appears due to the coherent interference between reflected waves in the formation of the image [?]. In [?], [?] and [?], ICA is applied to SAR images, and in [?] to ultrasound images. In all these works the ICA methods are used as if there is no multiplicative noise, so their applicability is reduced. On the other hand, a method for denoising images with multiplicative noise using non-linear ICA is developed in [?] where both, the multiplicative noise and the noise-free image, are considered as the independent sources and separete.

This work was partially supported by the "Ministerio de Ciencia y Tecnología" of Spain under Project TEC2004-06036-C03-02/TCM.

The multiplicative noise is taken into account in the separation of the linear mixture for the first time in [?] and [?], where a method that exploits the structure of the second- and third-order statistic of the noisy data is designed. The method is able to extract the mixing matrix in the MICA model and it gives much better results than the standard ICA method in these signals. However, the method needs more than seven sources, with at most one symmetrical (symmetric probability density function (PDF) respect to the mean), to converge to the correct solution [?]. Although the results of this method are promising, these restrictions make its application to real problems difficult, especially the limitation in the PDF of the sources, since the problem is blind and generally there is not control over them. In [?], a general approach to extend the ideas of the method to use fourth-order statistic is presented, although without developing an algorithm and therefore without studying if the inclusion of this fourth-order statistic can actually overcome the previous limitations. In this paper, an ICA algorithm to extract the mixing matrix in signals following the MICA model is designed. This algorithm uses fourth-order statistic, which allows to overcome the limitations that the method designed in [?] possesses. Specifically, as the new method uses fourth-order cumulants it can work with symmetrical signals, while the method in [?] does not because it uses only third-order statistic and this is null for symmetrical processes. On the other side, the inclusion of fourth-order cumulants adds new unknowns to the problem but also new equations, which are more, such that the minimum number of sources necessary to solve decreases from eight in [?] to only three in this paper.

The paper is organised as follow. The MICA model is presented in Section II, and the second-, third- and fourth-order statistics of the noisy data are also shown. In Section III, the structure of these statistics is used to develop the fourth-order multiplicative ICA (FMICA) method. The convergence of the method is briefly treated in Sections IV. In Section V, the FMICA method is compared with standards ICA methods, and finally the principal conclusions of the paper are summarized in Section VI.

2. MULTIPLICATIVE ICA MODEL

The MICA model [?] assumes that the signals are a linear mixture of independent sources contaminated by multiplicative noise. This can be expressed as:

$$z_i = v_i x_i,\ i = 1, \ldots, N \text{ with } \mathbf{x} = \mathbf{A}\mathbf{s} \tag{1}$$

where $\mathbf{s} = [s_i, \ldots, s_N]^T$ is the vector of independent sources, $\mathbf{v} = [v_i, \ldots, v_N]^T$ is the multiplicative noise, which is formed by mutually independent random variables with mean one and the noise-free data $\mathbf{x}$ and this noise are also independent [?]. $\mathbf{A}$ is the $N \times N$ mixing matrix and N is the number of signals. For simplicity, real signals and the same number of sources and signals are assumed in the paper, but these assumptions can be relaxed without loss of generality. The covariances, third- and fourth-order cumulants of the noisy observations $\mathbf{z}$ can be related to the covariances and third- and fourth-order cumulants of the noise-free data $\mathbf{x}$, the noise $\mathbf{v}$ and the independent sources $\mathbf{s}$.

The covariance between two elements z_i and z_j of the vector $\mathbf{z}$ is defined as $\sigma^z_{ij} = \mathscr{E}\{(z_i - \mu^z_i)(z_j - \mu^z_j)\}$, where μ^z_i is the mean of the signal z_i and $\mathscr{E}\{\cdot\}$ is the expectation operator. The third-order cumulant of the signals z_i, z_j and z_k is defined as $\gamma^z_{ijk} = \mathscr{E}\{(z_i-\mu^z_i)(z_j-\mu^z_j)(z_k-\mu^z_k)\}$ and the fourth-order cumulant of the signals z_i, z_j, z_k and z_l is defined as $\kappa^z_{ijkl} = \mathscr{E}\{(z_i-\mu^z_i)(z_j-\mu^z_j)(z_k-\mu^z_k)(z_l-\mu^z_l)\} - \sigma^z_{ij}\sigma^z_{kl} - \sigma^z_{ik}\sigma^z_{jl} - \sigma^z_{il}\sigma^z_{jk}$. All the indices go from 1 to N, as all the indices will do in the rest of the paper, unless otherwise stated. Substituting the expression (1) in the three previous definitions of σ^z_{ij}, γ^z_{ijk} and κ^z_{ijkl} and taking into account the independence between the elements of $\mathbf{v}$, between the elements of $\mathbf{s}$ and between $\mathbf{x}$ and $\mathbf{v}$, these three functions are computed as:

$$\begin{aligned}\sigma^z_{ij} &= \sigma^x_{ij} + \lambda_i\delta_{ij}\\ \gamma^z_{ijk} &= \gamma^x_{ijk} + \lfloor \delta_{jk}\sum_r A_{ir}\alpha_{jr}\rfloor_{ijk} + \delta_{ijk}\beta_i\\ \kappa^z_{ijkl} &= \kappa^x_{ijkl} + \lfloor \delta_{jkl}\sum_r A_{ir}\psi_{jr}\rfloor_{ijkl}\\ &\quad + \lceil \delta_{ij}(\sum_{rs} A_{kr}A_{ls}\chi_{irs} + \delta_{kl}\xi_{ik}/2)\rceil_{ijkl} + \phi_i\delta_{ijkl}\end{aligned} \tag{2}$$

where μ^x_i, σ^x_{ij}, γ^x_{ijk} and κ^x_{ijkl} are the mean, covariance, third- and fourth-order cumulants, respectively, of the components of $\mathbf{x}$ specified in the subindices; all the sums go from 1 to N, as will do in the rest of the paper unless otherwise stated; for any arbitrary functions f_{ijk}, dependent on the indices i, j and k, and g_{ijkl}, dependent on the indices i, j, k and l, it is defined $\lfloor f_{ijk}\rfloor_{ijk} \triangleq f_{ijk} + f_{jki} + f_{kij}$, $\lfloor g_{ijkl}\rfloor_{ijkl} \triangleq g_{ijkl} + g_{jkli} + g_{klij} + g_{lijk}$ and $\lceil g_{ijkl}\rceil_{ijkl} \triangleq g_{ijkl} + g_{ikjl} + g_{iljk} + g_{jkil} + g_{jlik} + g_{klij}$; and the rest of the parameters in (2) are:

$$\begin{aligned}\lambda_i &= \sigma^v_i(\sigma^x_{ii} + (\mu^x_i)^2);\;\; \beta_i = \gamma^v_i(\gamma^x_{iii} + 3\mu^x_i\sigma^x_{ii} + (\mu^x_i)^3)\\ \alpha_{ij} &= \sigma^v_i\left(A_{ij}{}^2\gamma^s_j + 2A_{ij}\mu^x_i\right)\\ \phi_i &= \kappa^v_i(\kappa^x_{iiii} + 4\mu^x_i\gamma^x_{iii} + 6(\mu^x_i)^2\sigma^x_{ii} + 3(\sigma^x_{ii})^2 + (\mu^x_i)^4)\\ \psi_{ij} &= \gamma^v_i A_{ij}(\kappa^s_j A_{ij}{}^2 + 3\mu^x_i(\gamma^s_j A_{ij} + \mu^x_i) + 3\sigma^x_{ii})\\ \chi_{ijk} &= \sigma^v_i A_{ij}(\delta_{jk}(A_{ij}\kappa^s_j + 2\gamma^s_j\mu^x_i) + 2A_{ik})\\ \xi_{ij} &= \sigma^v_i\sigma^v_j(\kappa^x_{iijj} + 2(\lfloor\mu^x_i\gamma^x_{ijj}\rfloor_{ij} + (2\mu^x_i\mu^x_j + \sigma^x_{ij})\sigma^x_{ij}))\end{aligned} \tag{3}$$

where κ^v_i, γ^v_i and σ^v_i are the kurtosis, skewness and variance of v_i; κ^s_i and γ^s_i are the kurtosis and skewness of s_i; and for an arbitrary function f_{ij}, dependent on the indices i and j, it is defined $\lfloor f_{ij}\rfloor_{ij} \triangleq f_{ij} + f_{ji}$. These parameters are unknown in the problem, since they depend on the mixing matrix and the sources and the problem is blind.

3. MULTIPLICATIVE ICA METHOD (MICA)

ICA searches for the linear transformation whose output is as independent as possible. If the data follow the ICA model, the result is the inverse of the mixing matrix, which is called the *unmixing matrix* and is the goal of all ICA methods. The most usual measures of the independence in the different ICA methods are the mutual information (MI), the negentropy and the cumulants. In the case of MICA model, none linear transformation of the noisy data can produce independent output [**?**], but the output after the application of the unmixing matrix possesses a statistical structure that can be used to find this matrix. Specifically, the cumulants of the output $\mathbf{y} = \mathbf{Bz}$, with $\mathbf{B}$ the unmixing matrix, can be obtained from (2), using that $\mathbf{BA} = \mathbf{I}$. The result is:

$$\begin{aligned}\sigma^y_{ij} &= \delta_{ij} + \sum_r B_{ir}B_{jr}\lambda_r\\ \gamma^y_{ijk} &= \gamma^s_i\delta_{ijk} + \sum_r B_{ir}B_{jr}B_{kr}\beta_r + \lfloor\sum_r B_{jr}B_{kr}\alpha_{ri}\rfloor_{ijk}\\ \kappa^y_{ijkl} &= \kappa^s_i\delta_{ijkl} + \lfloor\sum_r B_{ir}B_{jr}B_{kr}(B_{lr}\phi_r/4 + \psi_{ri})\rfloor_{ijkl}\\ &\quad + \lceil\sum_r B_{ir}B_{jr}(\chi_{rkl} + \sum_s B_{ks}B_{ls}\xi_{rs}/2)\rceil_{ijkl}\end{aligned} \tag{4}$$

where σ^y_{ij}, γ^y_{ijk} and κ^y_{ijkl} are the covariance, third- and fourth-order cumulant, respectively, of the correspondent components of $\mathbf{y}$. The arbitrary scaling factor associated with ICA problems appears also in the MICA model, due to the part of the signal modelled as $\mathbf{x} = \mathbf{As}$, and the normalization $\sigma^s_i = 1$ is set to eliminate it.

Then, the cumulants up to fourth-order of the output of a linear transformation $\mathbf{y} = \mathbf{Bz}$ of the noisy data possess the special structure (4) in the case $\mathbf{B} = \mathbf{A}^{-1}$, so the unmixing matrix can be found as the linear transformation which reproduces this structure. In order to do this, the covariance, third- and fourth-order cumulants of $\mathbf{y}$ can be estimated from $\hat\sigma^z_{ij}$, $\hat\gamma^z_{ijk}$ and $\hat\kappa^z_{ijkl}$, that are the covariance, third- and fourth-order cumulants estimated directly from the data $\mathbf{z}$ through their definitions. This is done as:

$$\begin{aligned}\hat\sigma^y_{ij} &= \sum_{mn} B_{im}B_{jn}\hat\sigma^z_{mn}\;;\; \hat\gamma^y_{ijk} = \sum_{mnp} B_{im}B_{jn}B_{kp}\hat\gamma^z_{mnp}\\ \hat\kappa^y_{ijkl} &= \sum_{mnpq} B_{im}B_{jn}B_{kp}B_{lq}\hat\kappa^z_{mnpq}\end{aligned} \tag{5}$$

These estimated functions $\hat\sigma^y_{ij}$, $\hat\gamma^y_{ijk}$ and $\hat\kappa^y_{ijkl}$ depend only on the unmixing matrix $\mathbf{B}$, while the functions σ^y_{ij}, γ^y_{ijk} and κ^y_{ijkl} in (4) depend on the unmixing matrix and also on the set of parameters in (3), that, as the problem is blind, are unknown. This set of parameters is highly dependent one on each other, so it would be interesting to find a new set of independent parameters, reducing in this way the dimension of the problem. This is possible taking into account the unknown information. Respect to the sources, the mean, the variance, the third- and the fourth-order cumulants of their components are unknown, but the variance of the sources is normalized to the unity to avoid the scale indetermination. Respect to the noise, its mean should to be one, so variance, third- and fourth-order cumulants are unknown. Apart from data, the mixing and the unmixing matrix are unknown, although they are related by the inversion operation. Then, a set of independent parameters are:

$$\{\phi_i, \eta_i, \omega_{ij}, \rho_i, \gamma^s_i, \kappa^s_i, B_{ij}\}_{i,j=1,\ldots,N} \tag{6}$$

where $\eta_i = \sqrt{\sigma^v_i}\mu^x_i$, $\omega_{ij} = \sqrt{\sigma^v_i}A_{ij}$ and $\rho_i = \gamma^v_i/(\sigma^v_i)^{3/2}$. The mean of the components of $\mathbf{x}$ can be estimated by the mean of the noisy data $\mathbf{z}$, since the noisy is one mean and independent of $\mathbf{x}$, so $\mu^x_i = \mu^z_i$. On the other side, the parameters ω_{ij}, B_{ij} and η_i are not independent, since $\omega_{ij} = (\eta_i/\mu^x_i)(B_{ij})^{-1}$. In other to avoid to include the inversion in the method, which is really problematic in posterior minimization, these parameters are considered independent and the relation between them will be taken account later in the cost function. So the total number of independent parameters is $N_1 = N(2N+5)$, and the parameters (3) (and with them the functions in (4)), can be expressed as function of (6). The explicit relation is omitted for space reason.

Then, the estimated functions (5) will be equal to the functions (4) when $\mathbf{B} = \mathbf{A}^{-1}$ and the rest of the parameters in (6) take their theoretical values, which will be call the correct solution. To measure how well the structure is reproduced for a set of parameters, a cost function J can be built as:

$$\begin{aligned}J &= \sum_{ij}(\mu^x_i\sum_k\omega_{ik}B_{kj} - \eta_i\delta_{ij})^2 + \sum_{i\geq j}\left(\sigma^y_{ij} - \hat\sigma^y_{ij}\right)^2\\ &\quad + \sum_{i\geq j\geq k}\left(\gamma^y_{ijk} - \hat\gamma^y_{ijk}\right)^2 + \sum_{i\geq j\geq k\geq l}\left(\kappa^y_{ijkl} - \hat\kappa^y_{ijkl}\right)^2\end{aligned} \tag{7}$$

with the definitions in (4) and (5). The first sum in the cost function is included to take into account the relation between the parameters B_{ij}, ω_{ij} and η_i. This function is formed by $N_2 = N(N+1)/2((N+2)/3(1+(N+3)/4)) + N^2$ terms, is function on the N_1 parameters (6) and will be zero in the correct solution. This kind of cost function has been widely used in cumulant matching in blind equalization and channel estimation and is similar to the one used in [**?**].

Then, the problem is reduced to find the value of the parameters $\{\phi_i, \eta_i, \omega_{ij}, \rho_i, \gamma_i^s, \kappa_i^s, B_{ij}\}_{i,j=1,\ldots,N}$ that minimizes the cost function (7), which means a problem of non-linear minimization of J. This is usually solved in the ICA bibliography by the steepest descendant method, such that, if the parameters are grouped in a $N_1 \times 1$ vector $\mathbf{b}$ with components equal to the different parameters, the vector of unknowns is obtained in the step k as $\mathbf{b}(k) = \mathbf{b}(k-1) - \mu \nabla_{\mathbf{b}} J$, where $\nabla_{\mathbf{b}} J$ is the gradient or the natural gradient (depending on the method) of the function J respect to the parameters, and μ is the learning ration that control the size of each step. Only the gradient of the cost function and the starting points are necessary to finish the method. The gradient is obtained taking the derivative in the definition (7), using (4), (5) and the relation between the parameters (3) and (6). The result of this process is the expression of $\nabla_{\mathbf{b}} J$, but it has to be omitted for space reasons. With the gradient, the method only needs the starting point to be finished.

However, the structure of J is much complicate that in the standard ICA case or in [?], and the steepest descendant method is too slow in the convergence. The use of other non-linear optimization methods have been tested and the one with better results is the quasi-Newton method of BFGS (Broyden-Fletcher-Goldfard-Shanno), where the parameters are obtained in the step $k+1$ as

$$\mathbf{b}(k+1) = \mathbf{b}(k) - \mu(k)\mathbf{H}(k)\nabla_{\mathbf{b}} J \tag{8}$$

where the matrix $\mathbf{H}(k)$ is an estimation of the inverse of the Hessian in the step k, obtained using the value of the parameters and the gradient in the steps k and $k-1$, and $\mu(k)$ is the learning ratio in the step k. The readers are referred to the bibliography for details about BFGS method [?].

At this stage, only the initialization is needed to finish the method. If the cost function is unimodal, any initial values will lead to the correct solutions. However, in practice if the method is applied with random initial values, it seldom reaches the correct solution. Most of the methods of minimizing non-linear functions only converge to the correct solution if the starting point is close enough to it. If this is not so, the method can diverge, converge to a spurious minimum or wander without reaching any solution. So, suitable starting points close to the solution need to be specified in order the method to convergence properly. To determine them, the same approach than in [?] is used, i.e. the noise is supposed small, so it is neglected in first approximation, except in the case of the variance, that it is supposed known or estimated. So, the starting point for the parameters $\{\phi_i, \rho_i\}_{i,j=1,\ldots,N}$ is zero. On the other hand, there is no previous information about the value of $\mathbf{B}$, and it is necessary to resort to a standard ICA method to find the starting point. This is not a problem, because the intention of this paper is not find an independent way to solve the problem, rather try to improve the results of the standard ICA methods in the MICA model. Then, the starting point for the unmixing matrix, $\mathbf{B}^0$, is the solution of a standard ICA method. With this value, the starting points of γ_i^s and κ_i^s are chosen as $(\hat{\gamma}_{iii}^y)^0$ and $(\hat{\kappa}_{iiii}^y)^0$, respectively, where they are the estimated skewness and kurtosis of the components of $\mathbf{y} = \mathbf{B}^0\mathbf{z}$ obtained using (5). With an estimation of the variance of the multiplicative noise $\hat{\sigma}_i^v$, the starting point of η_i is the one given by its definition and the fact that $\mu_i^x = \mu_i^z$, so $\eta_i^0 = \sqrt{\hat{\sigma}_i^v}\mu_i^z$. Finally, the starting point of ω_{ij} is $\omega_{ij}^0 = \sqrt{\hat{\sigma}_i^v}(\mathbf{B}^0)_{ij}^{-1}$. The estimation of the variance of the noise can be obtained from a uniform region in the noisy data, as a uniform patch in an image, or from theoretical considerations, but only a rough approximation is necessary. If this previous estimation is not possible, the variance of the noise is considered zero, and the starting point of the parameters η_i and ω_{ij} is set zero. Both elections in the starting point lead to the convergence of the minimization, but the first provides better results.

In summary, the FMICA method consists of the equations (??), with the gradients of the cost function and the specified starting points.

4. CONVERGENCE

The FMICA uses the gradient of the cost function in a point to find the smallest value of the cost function in the neighbourhood of the point, so the minimum is found in successive steps. This process leads to the correct solution if it is a minimum of the cost function, there are not zero gradient directions passing through it and the starting point is in the "attraction zone" of this correct solution. The correct solution is, at least asymptotically, a zero of a positive cost function, by construction, so it is a minimum. The problem of the zero-gradient direction is a problem of local convergence, while the determination of the attraction zone of the solution could be seen as a problem of global convergence. The study of the global convergence is a very difficult problem for practical cost functions and is out of the scope of this paper. However, the local convergence is possible to be studied. The process is the same that the follow in [?], where it is studied through the Hessian of the cost function. Specifically, a positive definite Hessian matrix in the correct solution insures that the method converges to it if the starting point is close enough. It can be shown how the computation of the Hessian matrix in the correct solution does not need of second derivatives, so it is sufficient first derivatives, what makes it considerably easier. As the first derivatives have been computed in the gradients, the Hessian matrix in the correct solution can be built and its sign studied. The result of this process is the condition when the local convergence is guaranteed, so it will find the correct solution with adequate close starting points. It can be shown that the method converge for any set of three or more independent sources, mixed with a full rank matrix and contaminated by whatever multiplicative noise. This means that there is no restriction in the PDF of the sources and the noise, i.e. no limitation in their statistical nature. In the practice, however, there is a limitation and it is that at most one of the sources can be Gaussian. This limitation is due to the standard ICA methods can not provide starting point close enough to the correct solution in this situation. The explicit study has to be omitted because of lack of space.

Then, in the practice, the FMICA method converges to the correct solution for three or more sources, if the starting points are close enough of this solution. If a starting point is or not "close enough" should be study based on the results.

5. RESULTS

In this section, the behaviour of the FMICA method is investigated by simulations. It will be compared with a standard ICA method, specifically the FastICA method [?], for different levels of multiplicative noise. The sources are generated following some PDF, with their variance equal to one, in order to satisfy the normalisation. These sources are mixed with a full rank square matrix, whose elements are selected randomly in the interval [0 1). To complete the multiplicative ICA model, the components of the multiplicative noise are generated distributed following other PDF, always with mean one. The standard deviation of the noise is varied in order to test the behaviour of the FMICA method in different levels of noise. The starting points are the ones specified in Section III. The election of the PDFs of the sources and the noise is arbitrary, since the method is independent on it.

The result of the FMICA is the matrix $\mathbf{B}$, which should be the unmixing matrix. The product of $\mathbf{B}$ and the mixing matrix is called the global transformation. Although in the development of the FMICA method the arbitrary scale factor has been eliminated with a condition on the variances of the sources, there is, as in the standard ICA model, also a sign and a permutation indetermination in the model. So, if the MICA model is an exact representation of the data and the functions in (5) are perfectly estimated, the inverse of the mixing matrix is not the only one that reproduces the structure (4) in the second-, third- and fourth-order statistics, but also all permuted sign-switched versions of this matrix does. Thus, in a perfect estimation, the global transformation should be $\mathbf{BA} = \mathbf{PS}$, where $\mathbf{P}$ is a permutation matrix and $\mathbf{S}$ is a diagonal matrix with its diagonal elements equal to one or minus one. A parameter d is de-

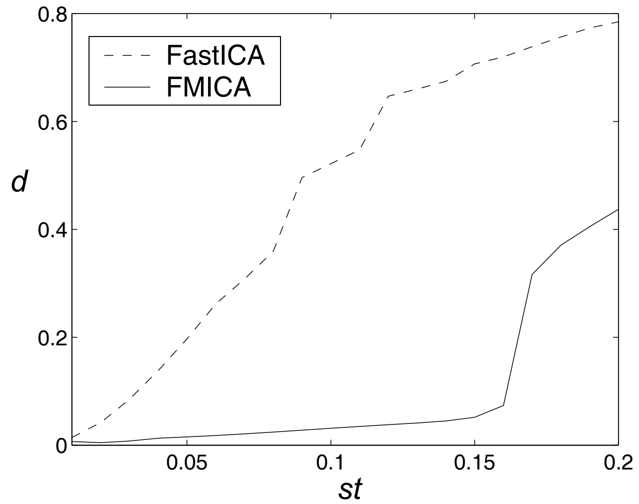

Figure 1: Mean of d as function of st for a linear mixture of three uniform sources, contaminated by multiplicative Gaussian noise.

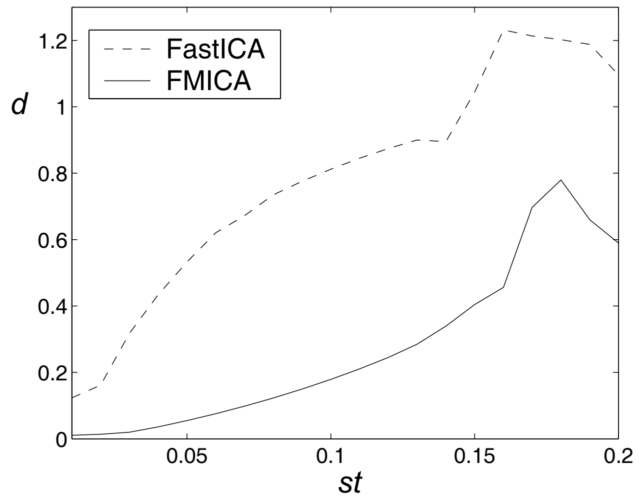

Figure 2: Mean of d as function of st for a linear mixture of four truncated rational sources, contaminated by multiplicative uniform noise.

fined as the minimum distance from the global transformation $\mathbf{BA}$ to the identity or any permuted sign-switched version of the identity [**?**]. The distance is measured as the Frobenius norm of the difference and this parameter will be used to measure the performance of the methods. Specifically:

$$d = \min_{\mathbf{PS}} \{ \sum_{ij} \left((\mathbf{BA} - \mathbf{PS})_{ij} \right)^2 \} \tag{9}$$

where the minimum is taken over all possible permutations and sign changes. The result of the method is the unmixing matrix $\mathbf{B}$, and a parameter d is defined to measure how close it is to the inverse of the mixing matrix.

In Figures 1 and 2, the values of the parameter d for the FastICA method and the MICA method as function of the standard deviation of the speckle noise, st, are presented. The Figure 1 corresponds with three sources, uniformily distributed in the interval $[0\ a_i)$, and contaminated by multiplicative Gaussian noise. The Figure 2 corresponds with four signals generated generated by taking the exponential of N uniform distributed signals in the interval $[0\ a_i)$ and contaminated with multiplicative uniform noise. The value of a_i is taken different in each one of the N sources. The number of data in the signals is 100,000. The shown d are the mean of 10 realizations. The MICA method is stopped when a fixed number of iterations have been completed or when the difference in the value of the cost function between ten consecutive steps is smaller than a threshold.

It can be seen that the results obtained with the different MICA method are always better than the FastICA method for both the number of signals and the range of noise studied. This comparison has been done for different mixtures, various number of sources (always greater than three) and for a wide range of st. The results obtained always have this behaviour.

6. CONCLUSIONS

In this paper a new method, the FMICA method, that obtains the unmixing matrix from a linear mixture of independent sources in the presence of multiplicative noise has been developed. The method tries to overcome the limitations of the ICA methods for this kind of signal. In order to do this, the FMICA method does not find a linear transformation whose outputs possess null cross-correlation and null higher-order cross-cumulants, as ICA does, but searches for a specific structure in the second-, third- and fourth-order statistic of the outputs that takes into account the existence of multiplicative noise in the data. The FMICA method has been tested for different sources and noises and its results improve clearly with respect to the ones obtained by standard ICA methods for linear mixture of independent sources contaminated with multiplicative noise. The convergence of the method has been also studied, showing that there is not restriction in the statistical nature of the sources or of the noise.

REFERENCES

[1] P. Comon, "Independent component analysis—a new concept?," *Signal Processing*, vol. 36, pp. 287–314, 1994.

[2] S. Chitroub and B. Sansal, "Statistical characterisation and modelling of SAR images," *Signal Processing*, vol. 82, pp. 69–92, 2002.

[3] X. Zhang and C. H. Chen, "A new independent component analyis (ICA) method and its application to SAR images," in *Proc. of NNSP*, 2001, pp. 283–292.

[4] S. Chitroub and B. Sansal, "Unsupervised learning rules for POLISAR images analysis," in *Proc. of NNSP*, 2002, pp. 567–576.

[5] J. Karvonen and M. Similä, "Independent component analysis for ice SAR image classification," in *Proc. of IGARSS*, 2001, pp. 1255–1257.

[6] H. Neemuchwala, A. Hero, and P. Carson, "Image registration using entropic graph-mathching criteria," in *Proc. of Asilomar Conf. on Signal and Sistem*, 2002.

[7] M. Haritopoulos, H. Yin, and N. M. Allison, "Multiplicative noise removal using self-organizing maps," in *Proc. of ICA 2001*, 2001, pp. 206–211.

[8] D. Blanco, B. Mulgrew, and S. McLaughlin, "ICA method for speckle signals," in *Proc. of ICASSP*, 2004.

[9] D. Blanco and B. Mulgrew, "ICA in signals with multiplicative noise," *acepted in IEEE Trans. Signal Processing*, 2004.

[10] D. Blanco, B. Mulgrew, S. McLaughlin, D. P. Ruiz, and M. C. Carrion, "The use of ICA in speckle signals," in *Proc. of ICA*, 2004.

[11] S. S. Rao, *Engineering optimization: theory and practice*, Wiley-Interscience, 1996.

[12] A. Hayvärinen and E. Oja, "A fast fixed-point algorithm for independent component analysis," *Neural Computation*, vol. 9, pp. 1483–1492, 1997.