2012 IEEE INTERNATIONAL WORKSHOP ON MACHINE LEARNING FOR SIGNAL PROCESSING, SEPT. 23–26, 2012, SANTANDER, SPAIN

GLRT FOR TESTING SEPARABILITY OF A COMPLEX-VALUED MIXTURE BASED ON THE STRONG UNCORRELATING TRANSFORM

David Ramírez[1], *Peter J. Schreier*[1], *Javier Vía*[2] *and Ignacio Santamaría*[2]

[1] Signal and System Theory Group, Universität Paderborn, Germany
e-mail: {david.ramirez, peter.schreier}@sst.upb.de
[2] Communications Engineering Dept., University of Cantabria, Santander, Spain
e-mail: {jvia, nacho}@gtas.dicom.unican.es

ABSTRACT

The Strong Uncorrelating Transform (SUT) allows blind separation of a mixture of complex independent sources if and only if all sources have distinct circularity coefficients. In practice, the circularity coefficients need to be estimated from observed data. We propose a generalized likelihood ratio test (GLRT) for separability of a complex mixture using the SUT, based on estimated circularity coefficients. For distinct circularity coefficients (separable case), the maximum likelihood (ML) estimates, required for the GLRT, are straightforward. However, for circularity coefficients with multiplicity larger than one (non-separable case), the ML estimates are much more difficult to find. Numerical simulations show the good performance of the proposed detector.

Index Terms— Complex independent component analysis (ICA), circularity coefficients, generalized likelihood ratio test (GLRT), hypothesis test, maximum likelihood (ML) estimation.

1. INTRODUCTION

The Strong Uncorrelating Transform (SUT) allows blind separation of a mixture of complex independent sources based on second-order statistics only [1–3]. This is achieved by exploiting the invariance property of the circularity coefficients under linear transformations. A circularity coefficient measures how noncircular or improper a random variable is. The density contours of univariate complex Gaussian random variables are ellipses, and the shape of these ellipses is controlled by the circularity coefficients [3]. If a Gaussian random variable has circularity coefficient 0, then its probability density contours are circular [4–6]; if it has circularity coefficient 1, then its probability density contours degenerate into a line in the complex plane.

A necessary and sufficient condition for separability using the SUT is that all circularity coefficients of the independent sources are distinct. It thus makes intuitive sense that separation of the mixture should be easier if the circularity coefficients are more clearly separated, and it should become more difficult if the circularity coefficients are more clustered. This intuition is supported theoretically by [7].

In practice, the circularity coefficients are not known a priori and must be estimated from the observed data. We are thus confronted with the question whether or not a mixture is separable, based on a given set of observations. This paper deals with this problem by proposing a generalized likelihood ratio test (GLRT) for separability of a mixture of complex sources. This test boils down to testing whether all circularity coefficients are distinct or whether there are circularity coefficients with multiplicity larger than one.

2. REVIEW OF ICA BASED ON SUT

We first briefly review the SUT, which is a technique for independent component analysis (ICA) of complex sources based on second-order statistics [1–3]. Consider the instantaneous noiseless linear complex ICA model

$$\mathbf{x} = \mathbf{A}\mathbf{s}, \tag{1}$$

where $\mathbf{x} \in \mathbb{C}^P$ are the zero-mean measurements, $\mathbf{A} \in \mathbb{C}^{P\times P}$ is the *unknown* mixing matrix, assumed to have full rank, and $\mathbf{s} \in \mathbb{C}^P$ are the zero-mean sources, which are assumed to be independent. The assumption of equal number of measurements and sources can be made without loss of generality.

The objective of ICA is to blindly recover the sources $\mathbf{s}$ from the measurements $\mathbf{x}$ as $\hat{\mathbf{s}} = \mathbf{B}\mathbf{x}$, using a separating matrix $\mathbf{B}$. ICA seeks to determine independent components. Scaling of $\mathbf{s}$, i.e., multiplication with a diagonal matrix, and reordering the components of $\mathbf{s}$, i.e., multiplication with a permutation matrix, preserves independence. Hence, we can determine $\mathbf{B}$ only up to multiplication with a monomial matrix, which is the product of a diagonal and a permutation matrix.

The assumption of independent sources means that the covariance matrix $\mathbf{R}_{ss} = E\left[\mathbf{s}\mathbf{s}^H\right]$ and the complementary covariance matrix $\tilde{\mathbf{R}}_{ss} = E\left[\mathbf{s}\mathbf{s}^T\right]$ are both diagonal. Using the monomial matrix ambiguity, we may even make the stronger assumption that $\mathbf{R}_{ss} = \mathbf{I}$ and $\tilde{\mathbf{R}}_{ss} = \mathbf{K}$, where $\mathbf{K} = \text{diag}(k_1, \ldots, k_P)$ contains the circularity coefficients $1 \geq k_1 \geq \ldots \geq k_P \geq 0$ on the diagonal [2].

The mixture $\mathbf{x}$ has covariance matrix $\mathbf{R}_{xx} = E\left[\mathbf{x}\mathbf{x}^H\right] = \mathbf{A}\mathbf{R}_{ss}\mathbf{A}^H = \mathbf{A}\mathbf{A}^H$ and complementary covariance matrix $\tilde{\mathbf{R}}_{xx} = E\left[\mathbf{x}\mathbf{x}^T\right] = \mathbf{A}\tilde{\mathbf{R}}_{ss}\mathbf{A}^T = \mathbf{A}\mathbf{K}\mathbf{A}^T$. Let $\mathbf{C} = \mathbf{F}\mathbf{\Lambda}\mathbf{F}^T$ be Takagi's factorization [8] of the coherence matrix $\mathbf{C} = \mathbf{R}_{xx}^{-1/2}\tilde{\mathbf{R}}_{xx}\mathbf{R}_{xx}^{-T/2}$. Then, a separating matrix for the complex ICA problem is given by $\mathbf{B} = \mathbf{F}^H\mathbf{R}_{xx}^{-1/2}$, provided that all circularity coefficients are distinct. A key difference between ICA of real and ICA of complex sources is the following: If there are two real Gaussian sources in the mixture, separation is not possible. However, complex Gaussians are separable if and only if all circularity coefficients are distinct [2].

3. PROBLEM STATEMENT

In practice, we do not know the true circularity coefficients and we have to work with estimates instead. Based on these estimates, how can we tell whether a complex mixture can be separated using the SUT? To that end, we now develop a test for separability of complex ICA based on the SUT. Our hypothesis test can be stated as follows. Given a finite set of observations, namely $\{\mathbf{x}_n\}_{n=0}^{N-1}$, we must decide which one of the following hypotheses holds:

$$\begin{aligned} \mathcal{H}_1 &: \text{The model is separable,} \\ \mathcal{H}_0 &: \text{The model is not separable.} \end{aligned} \tag{2}$$

We have seen that the mixture is separable if and only if all circularity coefficients are distinct. Define $\mathbb{D}_+$ as the set of $P \times P$ diagonal matrices with elements in $[0, 1]$, and define $\mathbb{D}_{2+}$ as that subset of $\mathbb{D}_+$ where at least two of the diagonal entries are identical. Then the hypothesis test can be written as

$$\begin{aligned} \mathcal{H}_1 &: \tilde{\mathbf{R}}_{ss} \in \mathbb{D}_+, \\ \mathcal{H}_0 &: \tilde{\mathbf{R}}_{ss} \in \mathbb{D}_{2+}. \end{aligned} \tag{3}$$

That is, under the alternative hypothesis, the complementary covariance matrix cannot have repeated values, whereas under the null hypothesis it has at least two repeated entries (but it might have more). Both hypotheses are composite because the circularity coefficients are unknown. Moreover, even applying invariance techniques [9–11] to reduce the number of unknown parameters, it is easily shown that both hypotheses remain composite.

We shall assume that the sources are zero-mean (circular or noncircular) complex Gaussian. Their augmented covariance matrix [3] depends on the hypothesis:

$$\begin{aligned} \mathcal{H}_1 &: \mathbf{x} \sim \mathcal{CN}\left(\mathbf{0}, \underline{\mathbf{R}}_{xx}^{(1)}\right), \\ \mathcal{H}_0 &: \mathbf{x} \sim \mathcal{CN}\left(\mathbf{0}, \underline{\mathbf{R}}_{xx}^{(0)}\right), \end{aligned} \tag{4}$$

where the augmented covariance matrices are

$$\underline{\mathbf{R}}_{xx}^{(i)} = E\left[\underline{\mathbf{x}}\,\underline{\mathbf{x}}^H\right] = \begin{bmatrix} \mathbf{R}_{xx}^{(i)} & \tilde{\mathbf{R}}_{xx}^{(i)} \\ \tilde{\mathbf{R}}_{xx}^{(i)*} & \mathbf{R}_{xx}^{(i)*} \end{bmatrix}, \tag{5}$$

$\underline{\mathbf{x}} = \left[\mathbf{x}^T \ \mathbf{x}^H\right]^T$ is an augmented vector, and the covariance and complementary covariance matrices are

$$\mathbf{R}_{xx}^{(i)} = \mathbf{A}\mathbf{A}^H, \qquad \tilde{\mathbf{R}}_{xx}^{(i)} = \mathbf{A}\mathbf{K}_i\mathbf{A}^T,$$

with diagonal

$$\mathbf{K}_1 \in \mathbb{D}_+, \qquad \mathbf{K}_0 \in \mathbb{D}_{2+}. \tag{6}$$

Our hypothesis test is therefore a test for the structure of the augmented covariance matrix.

4. DERIVATION OF THE GLRT

In order to solve this hypothesis testing problem, we propose a generalized likelihood ratio test (GLRT). The generalized likelihood ratio is [12]

$$\mathscr{G} = \frac{\max\limits_{\mathbf{A},\mathbf{K}_0 \in \mathbb{D}_{2+}} p\left(\mathbf{X}; \mathbf{A}, \mathbf{K}_0\right)}{\max\limits_{\mathbf{A},\mathbf{K}_1 \in \mathbb{D}_{+}} p\left(\mathbf{X}; \mathbf{A}, \mathbf{K}_1\right)}, \tag{7}$$

where we have defined the data matrix $\mathbf{X} = [\mathbf{x}_0, \ldots, \mathbf{x}_{N-1}]$, and $p(\cdot)$ denotes the probability density function of the observations. We must first find the maximum likelihood (ML) estimates of the unknown parameters under each of the hypothesis.

The log-likelihood is[1]

$$\log p\left(\mathbf{X}; \underline{\mathbf{R}}_{xx}\right) = -\log\det\left(\underline{\mathbf{R}}_{xx}\right) - \text{tr}\left(\underline{\mathbf{R}}_{xx}^{-1}\hat{\underline{\mathbf{R}}}\right), \tag{8}$$

where the augmented sample covariance matrix is

$$\hat{\underline{\mathbf{R}}}_{xx} = \frac{1}{N}\underline{\mathbf{X}}\,\underline{\mathbf{X}}^H = \begin{bmatrix} \hat{\mathbf{R}}_{xx} & \hat{\tilde{\mathbf{R}}}_{xx} \\ \hat{\tilde{\mathbf{R}}}_{xx}^* & \hat{\mathbf{R}}_{xx}^* \end{bmatrix} \in \mathcal{R}.$$

Here, the augmented data matrix is $\underline{\mathbf{X}} = [\mathbf{X}^T, \mathbf{X}^H]^T$, and $\mathcal{R}$ denotes the set of augmented covariance matrices without any imposed structure.

4.1. ML estimates under $\mathcal{H}_1$

Under $\mathcal{H}_1$, the only constraint on $\underline{\mathbf{R}}_{xx}^{(1)}$ is that it be an augmented covariance matrix, and the maximum likelihood is therefore obtained for

$$\max_{\mathbf{A} \in \mathbb{C}^{P \times P}, \mathbf{K}_1 \in \mathbb{D}_+} p\left(\mathbf{X}; \mathbf{A}, \mathbf{K}_1\right) = \max_{\underline{\mathbf{R}}_{xx}^{(1)} \in \mathcal{R}} p\left(\mathbf{X}; \underline{\mathbf{R}}_{xx}^{(1)}\right).$$

[1] For the sake of notational simplicity, we drop additive and multiplicative constant terms that do not depend on the data.

That is, rather than estimating $\mathbf{A}$ and $\mathbf{K}_1$, we can estimate $\underline{\mathbf{R}}_{xx}^{(1)}$, which is much easier [3]: The ML estimate is simply the sample covariance matrix, i.e., $\hat{\underline{\mathbf{R}}}_{xx}^{(1)} = \hat{\underline{\mathbf{R}}}_{xx}$. The compressed log-likelihood then becomes

$$\log p\left(\mathbf{X}; \hat{\underline{\mathbf{R}}}_{xx}^{(1)}\right) = -\log\det\left(\hat{\mathbf{R}}_{xx}\right) - \frac{1}{2}\sum_{i=1}^{P}\log\left(1-\psi_i^2\right). \quad (9)$$

In this expression, ψ_i are the estimated circularity coefficients, that is, the singular values of the estimated coherence matrix $\hat{\mathbf{C}} = \hat{\mathbf{R}}_{xx}^{-1/2}\hat{\tilde{\mathbf{R}}}_{xx}\hat{\mathbf{R}}_{xx}^{-T/2}$.

4.2. ML estimates under $\mathcal{H}_0$

Under $\mathcal{H}_0$, obtaining the ML estimates is more difficult because we do not know a priori the number and "order" of the circularity coefficients with multiplicity greater than one. However, since fewer constraints will lead to greater likelihood, the maximum of the likelihood will be achieved when *exactly two* circularity coefficients are enforced to be identical. With this reasoning, we have to solve the optimization problem

$$\max_{\mathbf{A},\mathbf{K}_0\in\mathbb{D}_{2+}} p(\mathbf{X};\mathbf{A},\mathbf{K}_0) = \max_{\substack{\mathbf{A}\in\mathbb{C}^{P\times P}\\ \mathbf{K}_0\in\cup_i\mathbb{D}_2^i}} p(\mathbf{X};\mathbf{A},\mathbf{K}_0) = \max_i \max_{\substack{\mathbf{A}\in\mathbb{C}^{P\times P}\\ \mathbf{K}_0\in\mathbb{D}_2^i}} p(\mathbf{X};\mathbf{A},\mathbf{K}_0),$$

where $i = 1,\ldots,P-1$. In this expression, $\mathbb{D}_2^i$ denotes the subset of $\mathbb{D}_{2+}$ where only the ith and $(i+1)$th entries are identical, and $\cup_i\mathbb{D}_2^i$ denotes the union of those sets. Enforcing this structure in the optimization is rather involved. In the following, we shall describe the ML estimates of $\mathbf{A}$ and $\mathbf{K}_0$ and a sketch of the proof. The full derivation will be presented in a forthcoming journal paper.

First, we rewrite the log-likelihood. Taking into account the structure of the augmented covariance matrices given in (5), the log-likelihood becomes

$$\log p(\mathbf{X};\mathbf{A},\mathbf{K}_0) = \log\det\left(\underline{\mathbf{A}}^{-1}\underline{\mathbf{A}}^{-H}\right) - \log\det\left(\underline{\mathbf{K}}_0\right) - \operatorname{tr}\left(\underline{\mathbf{K}}_0^{-1}\underline{\mathbf{S}}\right), \quad (10)$$

where we have defined

$$\underline{\mathbf{S}} = \begin{bmatrix}\mathbf{S} & \tilde{\mathbf{S}}\\ \tilde{\mathbf{S}}^* & \mathbf{S}^*\end{bmatrix} = \begin{bmatrix}\mathbf{A}^{-1}\hat{\mathbf{R}}_{xx}\mathbf{A}^{-H} & \mathbf{A}^{-1}\hat{\tilde{\mathbf{R}}}_{xx}\mathbf{A}^{-T}\\ \mathbf{A}^{-*}\hat{\tilde{\mathbf{R}}}_{xx}^*\mathbf{A}^{-H} & \mathbf{A}^{-*}\hat{\mathbf{R}}_{xx}^*\mathbf{A}^{-T}\end{bmatrix},$$

the augmented covariance matrix of the sources is

$$\underline{\mathbf{K}}_0 = \begin{bmatrix}\mathbf{I} & \mathbf{K}_0\\ \mathbf{K}_0 & \mathbf{I}\end{bmatrix},$$

and $\underline{\mathbf{A}}$ denotes the augmented matrix of the linear transformation $\mathbf{A}$, i.e.,

$$\underline{\mathbf{A}} = \begin{bmatrix}\mathbf{A} & \mathbf{0}_{P\times P}\\ \mathbf{0}_{P\times P} & \mathbf{A}^*\end{bmatrix},$$

with $\mathbf{0}_{P\times Q}$ denoting the $P\times Q$ zero matrix.

Let $\hat{\mathbf{C}} = \mathbf{G}\mathbf{\Psi}\mathbf{G}^T$ denote Takagi's factorization of the estimated coherence matrix. Then we define $\mathbf{W} = \mathbf{A}^{-1}\hat{\mathbf{R}}_{xx}^{1/2}\mathbf{G}$, which may be seen as the inverse of the residual matrix, after applying the separation matrix given by the *estimated* SUT. With this, the likelihood in (10) is

$$\log p(\mathbf{X};\mathbf{W},\mathbf{K}_0) = -\log\det\left(\hat{\mathbf{R}}_{xx}\right) + \log\det\left(\mathbf{W}\mathbf{W}^H\right) - \frac{1}{2}\log\det\left(\underline{\mathbf{K}}_0\right) - \frac{1}{2}\operatorname{tr}\left(\underline{\mathbf{K}}_0^{-1}\underline{\mathbf{S}}\right), \quad (11)$$

with $\underline{\mathbf{S}}$ now rewritten as

$$\underline{\mathbf{S}} = \begin{bmatrix}\mathbf{W}\mathbf{W}^H & \mathbf{W}\mathbf{\Psi}\mathbf{W}^T\\ \mathbf{W}^*\mathbf{\Psi}\mathbf{W}^H & \mathbf{W}^*\mathbf{W}^T\end{bmatrix}.$$

Hence, in order to maximize (11) we may obtain the ML estimates of the *transformed* parameters, which are presented in the following lemmas.

First, we are looking for an ML estimate of $\underline{\mathbf{K}}_0$, for which the repeated entry occurs at positions i and $i+1$. However, taking into account the problem invariances, we can relax the restrictions that $\underline{\mathbf{K}}_0$ must have 1s on the main diagonal, and that the diagonal elements of $\mathbf{K}_0$ must be in $[0,1]$, since they can be imposed in $\mathbf{A}$ (or $\mathbf{W}$). Such an ML estimate is presented next.

Lemma 1 *The ML estimate of* $\underline{\mathbf{K}}_0$ *is given by*

$$\hat{\underline{\mathbf{K}}}_0 = \begin{bmatrix}\hat{\mathbf{I}} & \hat{\mathbf{K}}_0\\ \hat{\mathbf{K}}_0^* & \hat{\mathbf{I}}^*\end{bmatrix},$$

where $\hat{\mathbf{I}}$ *and* $\hat{\mathbf{K}}_0$ *are diagonal matrices composed by the diagonal elements of* $\mathbf{S}$ *and* $\tilde{\mathbf{S}}$*, respectively, with the exception of the ith and $(i+1)$th entries of both matrices, which are replaced by their respective average. That is,*

$$\hat{\mathbf{I}} = \operatorname{diag}(s_{11},\ldots,s_{i-1,i-1},\rho,\rho,s_{i+2,i+2},\ldots,s_{P,P}),$$

and

$$\hat{\mathbf{K}}_0 = \operatorname{diag}(\tilde{s}_{11},\ldots,\tilde{s}_{i-1,i-1},\tilde{\rho},\tilde{\rho},\tilde{s}_{i+2,i+2},\ldots,\tilde{s}_{P,P}),$$

with the averages given by

$$\rho = \frac{1}{2}(s_{i,i}+s_{i+1,i+1}), \qquad \tilde{\rho} = \frac{1}{2}(\tilde{s}_{i,i}+\tilde{s}_{i+1,i+1}).$$

Since we have relaxed the energy constraints in the estimation of $\underline{\mathbf{K}}_0$, they have to be imposed when estimating $\mathbf{W}$, yielding

$$s_{l,l} = \|\mathbf{w}_l\|^2 = 1, \qquad \tilde{s}_{l,l} = \sum_{m=1}^{P}\psi_m w_{l,m}^2 \in [0,1],$$

with $l \neq i, i+1$, and

$$\rho = \frac{1}{2}(\|\mathbf{w}_i\|^2 + \|\mathbf{w}_{i+1}\|^2) = 1,$$

$$\tilde{\rho} = \frac{1}{2}\sum_{l=i}^{i+1}\sum_{m=1}^{P} \psi_m w_{l,m}^2 \in [0,1],$$

where $\mathbf{w}_l$ denotes the lth row of $\mathbf{W}$. The proof continues by introducing a permutation such that $\mathbf{W} = [\mathbf{W}_A^T \ \mathbf{W}_B^T]^T$, with $\mathbf{W}_A$ containing the rows corresponding to distinct circularity coefficients and $\mathbf{W}_B$ containing the two rows corresponding to the circularity coefficients with multiplicity 2.

Lemma 2 *The ML estimate of* $\mathbf{W}_A$ *and* $\mathbf{W}_B$ *are*

$$\hat{\mathbf{W}}_A = \begin{bmatrix}\mathbf{I}_{P-2} & \mathbf{0}_{(P-2)\times 2}\end{bmatrix},$$
$$\hat{\mathbf{W}}_B = \begin{bmatrix}\mathbf{0}_{2\times(P-2)} & \operatorname{diag}(\sqrt{1+\chi}, \sqrt{1-\chi})\end{bmatrix},$$

respectively, with

$$\chi = \frac{2 - \left(\psi_i^2 + \psi_{i+1}^2\right) - 2\sqrt{\left(1-\psi_i^2\right)\left(1-\psi_{i+1}^2\right)}}{\left(\psi_i^2 - \psi_{i+1}^2\right)}.$$

Taking into account the ML estimates $\mathbf{W}_A$ and $\mathbf{W}_B$, it is easy to show that the ML estimate of the (inverse) mixing matrix becomes

$$\hat{\mathbf{A}}^{-1} = \hat{\mathbf{W}}\mathbf{G}^H\hat{\mathbf{R}}^{-1/2},$$

where $\hat{\mathbf{W}}$ is a diagonal matrix, whose elements are equal to one, with the exception of the ith and $(i+1)$th entries, which are $\sqrt{1+\chi}$ and $\sqrt{1-\chi}$, respectively. Hence, the ML estimate of the mixing matrix is just the estimated strong uncorrelating transform, with its ith and $(i+1)$th rows scaled.

Plugging the ML estimates in the likelihood, after some algebra, the final expression for the compressed log-likelihood is

$$\log p\left(\mathbf{X}; \hat{\mathbf{W}}, \hat{\mathbf{K}}_0\right) = \max_i \left\{ \log\left(\frac{2}{\tilde{\psi}_i\tilde{\psi}_{i+1} - \psi_i\psi_{i+1} + 1}\right) - \log\det\left(\hat{\mathbf{R}}_{xx}\right) - \sum_{\substack{l=1\\ l\neq i,i+1}}^{P} \log\tilde{\psi}_l \right\}, \quad (12)$$

where we have introduced the following transformation of the circularity coefficients $\tilde{\psi}_i^2 = 1 - \psi_i^2$.

4.3. A closed-form expression for the GLRT

To obtain the GLRT, we insert the expressions for the compressed likelihoods, given by (9) and (12), into (7). This yields the log-GLRT

$$\log \mathscr{G} = \max_i \left\{ f(\psi_i, \psi_{i+1}) - f(\psi_i^{-1}, \psi_{i+1}^{-1}) \right\} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \eta, \quad (13)$$

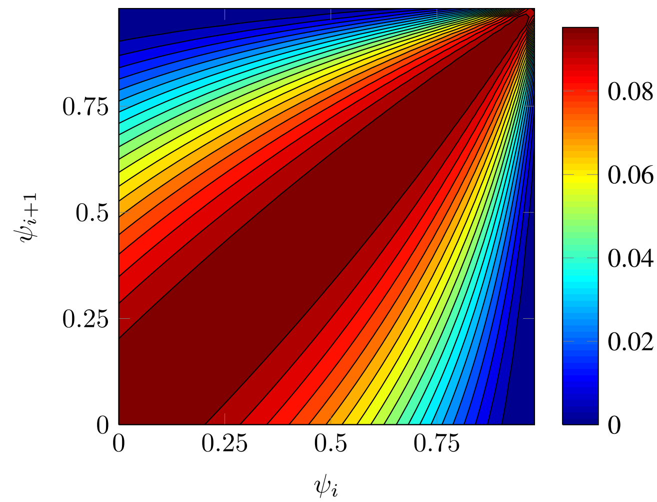

Fig. 1. Plot of the contours of $\mathscr{G}$ (in linear scale) as a function of ψ_i and ψ_{i+1}.

where

$$f(a,b) = \left[\left(1-a^{-2}\right)\left(1-b^{-2}\right)\right]^{-1/2}.$$

Let us define $\psi_i = \cos(\alpha_i)$, where α_i are the estimated principal angles [13], that is, the angles between s_i and s_i^*. Such angles play an important role for the capacity, information rate and inference of the Gaussian channel [13]. Using the principal angles, the log-GLRT may be rewritten as

$$\log \mathscr{G} = \max_i \left\{ \frac{\cos(\alpha_i - \alpha_{i+1}) + \cos(\alpha_i + \alpha_{i+1}) - 2}{\cos(\alpha_i - \alpha_{i+1}) - \cos(\alpha_i + \alpha_{i+1})} \right\} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \eta,$$

which shows that the GLRT is just a function of the sum and difference of principal angles. Finally, Figure 1 shows a plot of $\mathscr{G}$ as a function of ψ_i and ψ_{i+1} for some fixed i. Different colors denote the critical regions for different values of the threshold.

5. NUMERICAL RESULTS

We now numerically evaluate the performance of the proposed GLRT. As there is no other existing test for this problem, we cannot compare our results with any competitor. We consider $P = 4$ sources. In all cases, the circularity coefficients under $\mathcal{H}_1$ are $0.9, 0.6, 0.35$ and 0.1. Under $\mathcal{H}_0$, we will consider four examples:

1. $\{0.9, 0.35, 0.35, 0.1\}$ (two identical circularity coefficients)
2. $\{0.6, 0.6, 0.35, 0.35\}$ (two pairs of circularity coefficients)
3. $\{0.9, 0.35, 0.35, 0.35\}$ (three identical circularity coefficients)

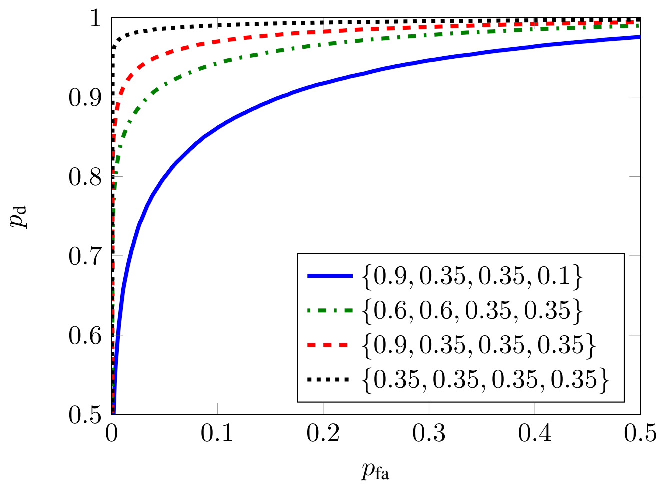

Fig. 2. ROC curves for four different examples with $P = 4$ sources and $N = 250$ samples. The circularity coefficients for each example are shown in the legend of the figure.

4. $\{0.35, 0.35, 0.35, 0.35\}$ (four identical circularity coefficients)

Figure 2 shows the receiver operating characteristic (ROC) curves for $N = 250$ samples. We notice that Example 4 performs best since all circularity coefficients are equal and, therefore, the two hypotheses are most separated. We also notice that three identical circularity coefficients are easier to detect than two pairs of identical circularity coefficients. Of course, these statements only hold for this particular setup, and the results will change depending on the specific values of the circularity coefficients, in particular, how much separated they are. Figure 3 shows the probability of missed detection p_{m} vs. number of samples for a fixed probability of false alarm $p_{\mathrm{fa}} = 0.01$, where similar conclusions can be drawn.

Finally, let us make some comments on the threshold selection. This is a difficult problem because the null hypothesis is composite, even when invariance techniques are applied. The distribution of the statistic therefore depends on unknown parameters, namely the circularity coefficients. In other works, applying invariance techniques, the null hypothesis becomes simple, which allows them to obtain the threshold through simulations for fixed parameters, and this threshold will still be valid for any other set of parameters [14, 15]. Nevertheless, that idea cannot be applied to this problem because $\mathcal{H}_0$ remains composite, even applying invariance techniques. To illustrate this problem, Figure 4 shows the empirical cumulative distribution function (ECDF) under $\mathcal{H}_0$ in an experiment with $P = 2$ sources, $N = 100$ samples and the following circularity coefficients: $\{0, 0\}$, $\{0.4, 0.4\}$ and $\{0.8, 0.8\}$. As can be seen in the figure, the distribution of $\mathscr{G}$ varies with the circularity coefficients. Nevertheless, since the

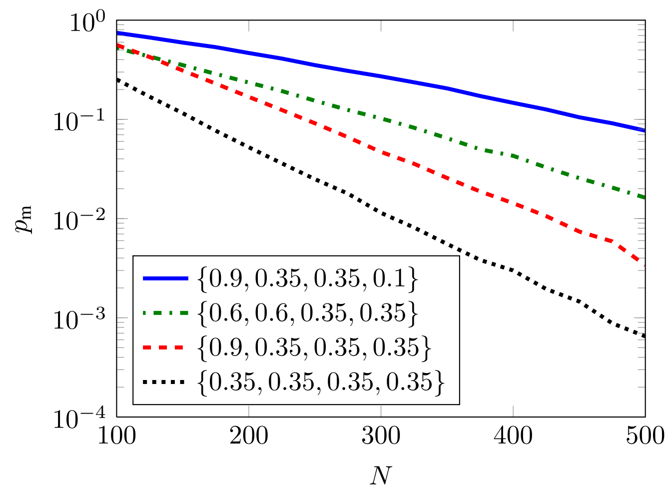

Fig. 3. Probability of missed detection vs. number of samples in four different examples with $P = 4$ sources, for a fixed probability of false alarm $p_{\mathrm{fa}} = 0.01$. The circularity coefficients for each example are shown in the legend of the figure.

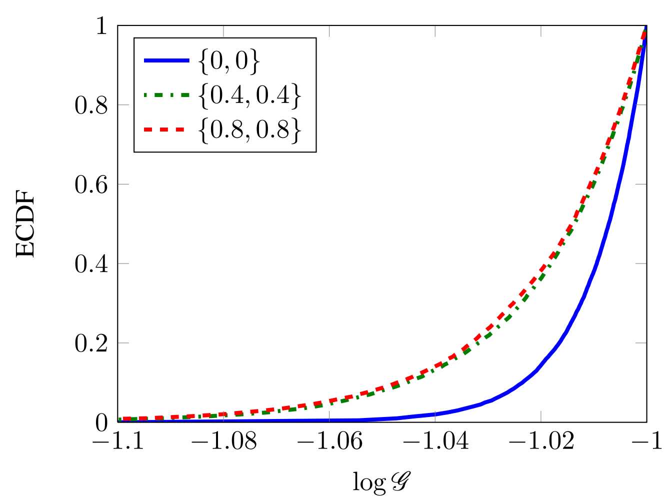

Fig. 4. Empirical cumulative distribution function (ECDF) under $\mathcal{H}_0$ for $P = 2$ sources and $N = 100$ samples. The circularity coefficients for each example are shown in the legend of the figure.

detector is a GLRT, it might be possible to apply Wilks' theorem [16], or some modification of it that works with fewer samples ("Barlett-correction factor" [17]). To conclude, the threshold selection problem deserves further investigation.

6. CONCLUSIONS

The Strong Uncorrelating Transform (SUT) allows blind separation of a complex mixture of independent sources if and only if all sources have distinct circularity coefficients. Since in practice we do not know those circularity coefficients, we have proposed a generalized likelihood ratio test for separability. The full derivation of this test is involved, mainly the derivation of the ML estimates under $\mathcal{H}_0$, due to the complicated structure of the augmented covariance matrix. Simulation results have shown the good performance of the proposed detector. However, threshold selection remains a critical problem to be addressed.

7. ACKNOWLEDGMENTS

The work of P. Schreier was supported by the Alfried Krupp von Bohlen und Halbach Foundation, under its program "Return of German scientists from abroad". The work of I. Santamaría and J. Vía was supported by the Spanish Government, Ministerio de Ciencia e Innovación (MICINN), under project COSIMA (TEC2010-19545-C04-03) and project COMONSENS (CSD2008-00010, CONSOLIDER-INGENIO 2010 Program).

8. REFERENCES

[1] L. DeLathauwer and B. DeMoor, "On the blind separation of non-circular sources," in *European Signal Process. Conf. (EUSIPCO)*, Toulouse, France, 2002.

[2] J. Eriksson and V. Koivunen, "Complex random vectors and ICA models: Identifiability, uniqueness, and separability," *IEEE Trans. Inf. Theory*, vol. 52, no. 3, pp. 1017–1029, Mar. 2006.

[3] P. J. Schreier and L. L. Scharf, *Statistical Signal Processing of Complex-Valued Data*, Cambridge University Press, 2010.

[4] F. D. Neeser and J. L. Massey, "Proper complex random processes with applications to information theory," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1293–1302, Jul. 1993.

[5] B. Picinbono, "Second-order complex random vectors and normal distributions," *IEEE Trans. Signal Process.*, vol. 44, no. 10, pp. 2637–2640, Oct. 1996.

[6] B. Picinbono and P. Bondon, "Second-order statistics of complex signals," *IEEE Trans. Signal Process.*, vol. 45, no. 2, pp. 411–420, Feb. 1997.

[7] A. Yeredor, "Performance analysis of the strong uncorrelating transformation in blind separation of complex-valued sources," *IEEE Trans. Signal Process.*, vol. 60, no. 1, pp. 478–483, Jan. 2012.

[8] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1985.

[9] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*, Addison - Wesley, 1991.

[10] T. S. Ferguson, *Mathematical statistics*, New York: Academic Press, 1967.

[11] E. L. Lehmann, *Testing Statistical Hypotheses*, Springer texts in Statistics, 1986.

[12] K. V. Mardia, J. T. Kent, and J. M. Bibby, *Multivariate Analysis*, New York: Academic, 1979.

[13] L. L. Scharf and C. T. Mullis, "Canonical coordinates and the geometry of inference, rate and capacity," *IEEE Trans. Signal Process.*, vol. 48, no. 3, pp. 824–831, March 2000.

[14] D. Ramírez, G. Vázquez-Vilar, R. López-Valcarce, J. Vía, and I. Santamaría, "Multiantenna detection of rank-P signals in cognitive radio networks," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 3764–3774, Aug. 2011.

[15] A. T. Walden and P. Rubin-Delanchy, "On testing for impropriety of complex-valued Gaussian vectors," *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 825–834, March 2009.

[16] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, vol. II, Prentice Hall, 1998.

[17] M. S. Bartlett, "The effect of standardization on a χ^2 approximation in factor analysis," *Biometrika*, vol. 38, no. 3/4, pp. 337–344, Dec. 1951.