

ON ICA OF IMPROPER AND NONCIRCULAR SOURCES

Peter J. Schreier

Electrical Engr. & Comp. Science
The Univ. of Newcastle, Australia
peter.schreier@newcastle.edu.au

Tülay Adalı

Comp. Science & Electrical Engr.
Univ. of Maryland, Baltimore County
adali@umbc.edu

Louis L. Scharf

Electrical & Computer Engr.
Colorado State Univ.
scharf@engr.colostate.edu

ABSTRACT

We provide a review of independent component analysis (ICA) for complex-valued improper and noncircular random sources. An improper random signal is correlated with its complex conjugate, and a noncircular random signal has a rotationally variant probability distribution. We present methods for ICA using second-order statistics, and higher-order statistics. For ICA based on second-order statistics, we emphasize the key role played by the circularity coefficients, which are the canonical correlations between the source and the complex conjugate. For ICA based on higher-order statistics, we show how to extend algorithms for real-valued ICA to the complex domain using Wirtinger calculus.

Index Terms— Independent component analysis, noncircular, improper, circularity coefficients, Wirtinger calculus

1. INTRODUCTION

Given a random vector \mathbf{x} , independent component analysis (ICA) seeks to determine a transformation \mathbf{W} such that $\mathbf{u} = \mathbf{W}\mathbf{x}$ has components that are as statistically independent as possible. ICA can thus be regarded as a higher-order extension of principal component analysis, where \mathbf{u} has uncorrelated components, and \mathbf{W} is the Karhunen-Loève transform.

In this paper, we provide a review of ICA for complex-valued improper and noncircular random sources. Consider a zero-mean, complex-valued random vector $\mathbf{s} \in \mathbb{C}^N$ with correlation matrix $\mathbf{R}_{ss} = E[\mathbf{s}\mathbf{s}^H]$. In order to fully describe the second-order behavior of \mathbf{s} , we also need the *complementary correlation matrix* $\tilde{\mathbf{R}}_{ss} = E[\mathbf{s}\mathbf{s}^T]$, which characterizes the correlation between \mathbf{s} and its conjugate \mathbf{s}^* . If $\tilde{\mathbf{R}}_{ss} = \mathbf{0}$, \mathbf{s} is called *proper*, otherwise *improper*. It is easy to see that \mathbf{s} is proper if and only if \mathbf{s} and a rotated version $\mathbf{s}e^{j\alpha}$ have the same second-order statistics for all real angles α .

The probability density function (pdf) of a complex random vector \mathbf{s} is defined in terms of the joint pdf of its real part \mathbf{s}_r and imaginary part \mathbf{s}_i as $p_s(\mathbf{s}) \triangleq p_{s_r, s_i}(\mathbf{s}_r, \mathbf{s}_i)$. If \mathbf{s} and $\mathbf{s}e^{j\alpha}$ have the same pdf for all real angles α , then \mathbf{s} is called *circular*, otherwise \mathbf{s} is *noncircular*. Proper random vectors are sometimes also called *second-order circular*, and improper random vectors are obviously noncircular.

Until recently, improper and noncircular signals have received little attention. Yet noncircular signals commonly arise in practice. For instance, many communications signals (such as BPSK, GMSK, OQPSK) as well as biomedical signals such as functional magnetic resonance imaging (fMRI) data are noncircular [1]. Accounting for the noncircular nature of signals also has big potential payoffs. As we will see later on, *noncircular* Gaussian sources may be separated under certain conditions. This is never possible for real Gaussian sources. We will first discuss second-order techniques for complex ICA, and then higher-order techniques.

2. CIRCULARITY COEFFICIENTS ARE CANONICAL CORRELATIONS BETWEEN \mathbf{s} AND \mathbf{s}^*

Propriety is preserved under *linear*, but *not widely linear*, transformation. We may thus be interested in finding a *maximal invariant* (also called complete set of invariants) for the second-order moments $(\mathbf{R}_{ss}, \tilde{\mathbf{R}}_{ss})$ under nonsingular linear transformation. The virtue of such a maximal invariant is that any function of $(\mathbf{R}_{ss}, \tilde{\mathbf{R}}_{ss})$ that is invariant under nonsingular linear transformation *must* be a function of the maximal invariant (and cannot depend on anything else). It has been shown in [14] that the *canonical correlations* [10] between \mathbf{s} and \mathbf{s}^* are such a maximal invariant.

The canonical correlations between \mathbf{s} and \mathbf{s}^* are the singular values of the *coherence matrix*

$$\mathbf{C} = \mathbf{R}_{ss}^{-1/2} \tilde{\mathbf{R}}_{ss} (\mathbf{R}_{ss}^{-*/2})^H = \mathbf{R}_{ss}^{-1/2} \tilde{\mathbf{R}}_{ss} \mathbf{R}_{ss}^{-T/2}, \quad (1)$$

which we will denote by $1 \geq k_1 \geq k_2 \geq \dots \geq k_N \geq 0$. Following Eriksson and Koivunen [7], we call these canonical correlations *circularity coefficients*, even though this term is not entirely accurate. Since the circularity coefficients only characterize improperity (or second-order noncircularity), the name “impropriety coefficients” might be more suitable.

There is a corresponding coordinate system, called the *canonical coordinate system*, where the latent description \mathbf{s}' has identity correlation matrix, $E[\mathbf{s}'\mathbf{s}'^H] = \mathbf{I}$, and diagonal complementary correlation matrix with the circularity coefficients on the diagonal,

$$E[\mathbf{s}'\mathbf{s}'^T] = \mathbf{K} = \mathbf{Diag}(k_1, k_2, \dots, k_N). \quad (2)$$

In [7], vectors that are uncorrelated with unit variance, but possibly improper, are called *strongly uncorrelated*, and the transformation \mathbf{A}_{ss} , which transforms \mathbf{s} into canonical coordinates \mathbf{s}' as $\mathbf{s}' = \mathbf{A}_{ss}\mathbf{s}$, is called the *strong uncorrelating transform* (SUT). The SUT is found as

$$\mathbf{A}_{ss} = \mathbf{F}_{ss}^H \mathbf{R}_{ss}^{-1/2}, \quad (3)$$

where \mathbf{F}_{ss} is obtained from the Takagi factorization of the coherence matrix $\mathbf{C} = \mathbf{F}_{ss} \mathbf{K} \mathbf{F}_{ss}^T$. The Takagi factorization is a special singular value decomposition (SVD) for a symmetric matrix, $\mathbf{C} = \mathbf{C}^T$. If \mathbf{C} has distinct nonzero singular values, then the SVD is unique up to multiplication of the matrix of left singular vectors from the right by a unitary diagonal matrix. For a symmetric matrix, the Takagi factorization determines this unitary diagonal factor in such a way that the matrix of left singular vectors is the conjugate of the matrix of right singular vectors. If all circularity coefficients are distinct and nonzero, the SUT is therefore unique up to the sign of its rows [7]. It is shown in [9, Sec. 4.4] how to compute the Takagi factorization.

The circularity coefficients show up at numerous occasions in a second-order treatment of improper random vectors. For instance, the entropy of an improper Gaussian random vector is [7, 15]

$$H_{\text{improper}} = \underbrace{\log[(\pi e)^N \det \mathbf{R}_{ss}]}_{H_{\text{proper}}} + \underbrace{\frac{1}{2} \log \prod_{i=1}^N (1 - k_i^2)}_{-I(\mathbf{s}; \mathbf{s}^*)}. \quad (4)$$

This explicitly shows the connection between the entropy of an *improper* Gaussian with the entropy of the corresponding *proper* Gaussian random vector with the same \mathbf{R}_{ss} but $\mathbf{R}_{ss} = \mathbf{0}$. The second summand, $I(\mathbf{s}; \mathbf{s}^*)$, is the mutual information between \mathbf{s} and its conjugate \mathbf{s}^* , which depends only on the circularity coefficients (because it is invariant under nonsingular linear transformation of \mathbf{s}). This mutual information is the loss of entropy of \mathbf{s} that is due to impropriety.

3. ICA USING SECOND-ORDER STATISTICS

ICA of improper sources is an interesting application of the invariance property of the circularity coefficients. Suppose that we observe a *linear mixture* \mathbf{x} of *independent components* (sources) \mathbf{s} , as described by

$$\mathbf{x} = \mathbf{M}\mathbf{s}. \quad (5)$$

We will make a few simplifying assumptions. The dimensions of \mathbf{x} and \mathbf{s} are assumed to be equal, and the *mixing matrix* \mathbf{M} is assumed to be nonsingular. The objective is to *blindly* recover the sources \mathbf{s} from the observations \mathbf{x} , without knowledge of \mathbf{M} , using a linear transformation \mathbf{W} . This transformation \mathbf{W} can be regarded as a *blind inverse* of \mathbf{M} ,

which is usually called a *separating matrix*. Note that since the model (5) is linear, it is unnecessary to consider widely linear transformations.

ICA seeks to determine independent components. Arbitrary scaling of \mathbf{s} , *i.e.*, multiplication by a diagonal matrix, and reordering the components of \mathbf{s} , *i.e.*, multiplication by a permutation matrix, preserves the independence of its components. The product of a diagonal and a permutation matrix is a *monomial* matrix, which has exactly one nonzero entry in each column and row. Hence, we can determine \mathbf{W} up to multiplication with a monomial matrix.

In the real-valued case, the blind recovery of \mathbf{s} cannot work if more than one source s_i is Gaussian. If only second-order information is available and there is no sample correlation, the best possible solution is to *decorrelate* the components, rather than to make them independent. This is done by determining the *principal* components $\mathbf{x}' = \mathbf{U}^H \mathbf{x}$ using the eigenvalue decomposition $\mathbf{R}_{xx} = E[\mathbf{x}\mathbf{x}^H] = \mathbf{U}\mathbf{A}\mathbf{U}^H$. However, the restriction to unitary rather than general linear transformations wastes a considerable degree of freedom in designing the blind inverse \mathbf{W} . In the complex case, it was shown by Eriksson and Koivunen [7] that using the SUT it is possible to determine \mathbf{W} using *second-order* information only, provided that \mathbf{s} has distinct circularity coefficients. In this section, we provide an alternative new proof, which exploits the fact that the circularity coefficients of \mathbf{s} are *invariant* under the *linear* mixing transformation \mathbf{M} . (The SUT requires that correlation and complementary correlation matrices exist. For some distributions (e.g., Cauchy distribution) this is not the case. Ollila and Koivunen [12] have presented a generalization of the SUT that works for such distributions.)

The assumption of *independent* components \mathbf{s} implies that the correlation matrix \mathbf{R}_{ss} and the complementary correlation matrix $\tilde{\mathbf{R}}_{ss}$ are both diagonal. It is therefore easy to compute canonical coordinates between \mathbf{s} and \mathbf{s}^* , denoted by $\mathbf{s}' = \mathbf{A}_{ss}\mathbf{s}$. In the strong uncorrelating transform $\mathbf{A}_{ss} = \mathbf{F}_{ss}^H \mathbf{R}_{ss}^{-1/2}$, $\mathbf{R}_{ss}^{-1/2}$ is a *diagonal* scaling matrix, and \mathbf{F}_{ss}^H is a *permutation* matrix that rearranges the canonical coordinates \mathbf{s}' such that s'_1 corresponds to the largest circularity coefficient k_1 , s'_2 to the second largest coefficient k_2 , and so on. This makes the strong uncorrelating transform \mathbf{A}_{ss} monomial. As a consequence, \mathbf{s}' has *independent components*.

The mixture \mathbf{x} has correlation matrix $\mathbf{R}_{xx} = \mathbf{M}\mathbf{R}_{ss}\mathbf{M}^H$ and complementary correlation matrix $\tilde{\mathbf{R}}_{xx} = \mathbf{M}\tilde{\mathbf{R}}_{ss}\mathbf{M}^T$. The canonical coordinates between \mathbf{x} and \mathbf{x}^* are computed as $\mathbf{x}' = \mathbf{A}_{xx}\mathbf{x} = \mathbf{F}_{xx}^H \mathbf{R}_{xx}^{-1/2} \mathbf{x}$, and the strong uncorrelating transform \mathbf{A}_{xx} is determined as explained in the previous section.

Figure 1 shows the connection between the different coordinate systems. The important observation is that \mathbf{s}' and \mathbf{x}' are both in canonical coordinates with the *same* circularity coefficients k_i . It remains to show that \mathbf{s}' and \mathbf{x}' are related by a diagonal unitary matrix $\mathbf{D} = \mathbf{A}_{xx}\mathbf{M}\mathbf{A}_{ss}^{-1}$ as $\mathbf{x}' = \mathbf{D}\mathbf{s}'$, provided that all circularity coefficients are distinct. Since \mathbf{s}' and \mathbf{x}' are both in canonical coordinates with the same diagonal

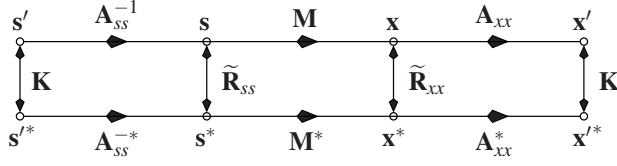


Fig. 1. Two-channel model for complex ICA. The vertical arrows show the complementary correlation matrix between the upper and lower lines.

canonical correlation matrix \mathbf{K} ,

$$E \left\{ \begin{bmatrix} \mathbf{s}' \\ \mathbf{s}'^* \end{bmatrix} \begin{bmatrix} \mathbf{s}' \\ \mathbf{s}'^* \end{bmatrix}^H \right\} = \begin{bmatrix} \mathbf{I} & \mathbf{K} \\ \mathbf{K} & \mathbf{I} \end{bmatrix} \quad (6)$$

$$E \left\{ \begin{bmatrix} \mathbf{x}' \\ \mathbf{x}'^* \end{bmatrix} \begin{bmatrix} \mathbf{x}' \\ \mathbf{x}'^* \end{bmatrix}^H \right\} = \begin{bmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}^* \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{K} \\ \mathbf{K} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{D}^H & \mathbf{0} \\ \mathbf{0} & \mathbf{D}^T \end{bmatrix} \\ = \begin{bmatrix} \mathbf{D}\mathbf{D}^H & \mathbf{D}\mathbf{K}\mathbf{D}^T \\ \mathbf{D}^*\mathbf{K}\mathbf{D}^H & \mathbf{D}^*\mathbf{D}^T \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{K} \\ \mathbf{K} & \mathbf{I} \end{bmatrix}. \quad (7)$$

This shows that \mathbf{D} is unitary and $\mathbf{D}\mathbf{K}\mathbf{D}^T = \mathbf{K}$. The latter can only be true if $D_{ij} = 0$ whenever $k_i \neq k_j$. Therefore, \mathbf{D} is diagonal and unitary if all circularity coefficients are distinct. Moreover, the corresponding diagonal entries of all nonzero circularity coefficients are actually ± 1 . Thus, \mathbf{x}' has independent components because \mathbf{s}' has independent components. Hence, we have shown that the strong uncorrelating transform \mathbf{A}_{xx} is a separating matrix for the complex linear ICA problem if all circularity coefficients are distinct, and the ICA solution is $\mathbf{u} = \mathbf{W}\mathbf{x}$ with $\mathbf{u} = \mathbf{x}'$ and $\mathbf{W} = \mathbf{A}_{xx}$.

We emphasize that this technique assumes that \mathbf{x} is indeed a *linear* mixture. If \mathbf{x} does not satisfy the linear model (5), the objective of ICA is to find components that are as independent as possible. The degree of independence is measured by a *contrast function* such as mutual information or negentropy. It is important to realize that the strong uncorrelating transform \mathbf{A}_{xx} is not guaranteed to optimize any contrast function. Finding maximally independent components in the *nonlinear* case requires the use of higher-order statistics, as discussed in the next section.

4. ICA USING HIGHER-ORDER STATISTICS

Complex ICA can be achieved using only second-order statistical information as long as the sources are improper and have distinct circularity coefficients. Using higher-order statistical information, one can develop ICA algorithms that can achieve separation of sources with any type of distribution, circular or noncircular, as long as the mixing matrix is of full column rank and there are no two complex Gaussian sources with the same circularity coefficient [7]. Algorithms such as joint approximate diagonalization of eigenmatrices (JADE) [6] explicitly calculate the higher-order statistics and

can be directly used for ICA of complex-valued data. These algorithms, which rely on joint diagonalization of cumulant matrices, are robust. However, their performance suffers as the number of sources increases, and the cost of computing and diagonalizing cumulant matrices becomes prohibitive for separating a large number of sources. On the other hand, ICA techniques that use nonlinear functions to implicitly generate the higher-order statistics, such as maximum likelihood (ML) [13], information-maximization (Infomax) [4], and maximization of non-Gaussianity (e.g., the FastICA algorithm) [11], which are all intimately related to each other, present attractive alternatives for performing ICA. These algorithms can be easily extended to the complex domain using Wirtinger calculus as shown in [1].

For the ICA problem given in (5), we determine a separating matrix \mathbf{W} such that the source estimates are recovered as $\mathbf{u} = \mathbf{W}\mathbf{x}$ by optimizing a measure of independence. In the case of ML ICA, we write the log-likelihood function for T independent samples $\mathbf{x}(t) \in \mathbb{C}^N$ as $\mathcal{L}(\mathbf{W}) = \sum_{t=1}^T \ell_t(\mathbf{W})$, where

$$\ell_t(\mathbf{W}) = \log p(\mathbf{x}(t)|\mathbf{W}) = \prod_{n=1}^N p_{s_n}(\mathbf{w}_n^H \mathbf{x}) + \log |\det \bar{\mathbf{W}}|. \quad (8)$$

Here, \mathbf{w}_n is the n th row of \mathbf{W} , $p_{s_n}(u_n) \triangleq p_{s_{n_r}, s_{n_i}}(u_{n_r}, u_{n_i})$ is the joint pdf of source n , $n = 1, \dots, N$, and $\bar{\mathbf{W}} = \begin{bmatrix} \mathbf{W}_r & -\mathbf{W}_i \\ \mathbf{W}_i & \mathbf{W}_r \end{bmatrix}$, where $\mathbf{W} = \mathbf{W}_r + j\mathbf{W}_i$. When writing (8), we used the Jacobian transformation

$$p_{\mathbf{x}}(\mathbf{x}) = |\det \bar{\mathbf{W}}| p_{\mathbf{u}}(\mathbf{W}\mathbf{x}) \quad (9)$$

where $p_{\mathbf{u}}(\mathbf{W}\mathbf{x}) = p_{\mathbf{u}}(\mathbf{u}) = p_{\mathbf{u}}(u_1, u_2, \dots, u_N)$. We also defined $\mathbf{W} = \mathbf{M}^{-1}$ in order to express the likelihood in terms of the inverse mixing matrix, which provides a convenient change of parameter. Note that the time index in $\mathbf{x}(t)$ has been omitted in the expressions for simplicity.

The cost function given in (8) is real-valued, hence not differentiable in the complex plane. The usual practice of calculating separate derivatives of the real and imaginary parts tends to get cumbersome quickly and leads to the need to make simplifying assumptions such as circularity [3]. Instead, we can write the function as $\ell_t(\mathbf{W}) = \ell_t(\mathbf{W}, \mathbf{W}^*)$ and use Wirtinger calculus to compute all derivatives similar to the real-valued case as shown in [1]. We then obtain

$$\Delta \mathbf{W} = (\mathbf{I} - \Psi(\mathbf{u})\mathbf{u}^H)\mathbf{W}, \quad (10)$$

where the vector of score functions $\Psi_n(u)$ for each source n is expressed in vector form as

$$\Psi(\mathbf{u}) = -\frac{1}{2} \left(\frac{\partial \log p_{s_r, s_i}(\mathbf{u}_r, \mathbf{u}_i)}{\partial \mathbf{u}_r} + j \frac{\partial \log p_{s_r, s_i}(\mathbf{u}_r, \mathbf{u}_i)}{\partial \mathbf{u}_i} \right).$$

Another natural cost function for performing ICA is negentropy, which measures the entropic distance of a given distribution from a Gaussian. In this case, independence is

achieved by moving the distribution of the transformed mixture $\mathbf{w}^H \mathbf{x}$ —the independent source estimates—away from Gaussian, *i.e.*, by maximizing non-Gaussianity. With negentropy maximization as the objective, all sources can be estimated by maximizing the cost function

$$J(\mathbf{W}) = \sum_{n=1}^N E \{ \log p_{s_n}(\mathbf{w}_n^H \mathbf{x}) \} \approx \frac{1}{T} \sum_{t=1}^T \sum_{n=1}^N \log p_{s_n}(\mathbf{w}_n^H \mathbf{x}) \quad (11)$$

under the unitary constraint for \mathbf{W} . Note that we have used the mean ergodic theorem to write the approximation in (11). Since $\det(\bar{\mathbf{W}}) = |\det(\mathbf{W})|^2$ [8], when \mathbf{W} is unitary, the second term in (8) vanishes. Thus, when compared to the ML formulation given in (8), it is clear that the two objective functions are equivalent if we constrain the weight matrix \mathbf{W} to be unitary for complex ML.

As in the real case, the two criteria are intimately linked to mutual information. Written as the Kullback-Leibler distance between the joint and factored marginal source densities, the mutual information is given by

$$\begin{aligned} I(\mathbf{W}) &= D \left(p(\mathbf{u}) \parallel \prod_{n=1}^N p_{s_n}(u_n) \right) = \sum_{n=1}^N H(u_n) - H(\mathbf{u}) \\ &= \sum_{n=1}^N H(u_n) - H(\mathbf{x}) - \log |\det \bar{\mathbf{W}}|. \end{aligned} \quad (12)$$

In the last line, we have again used the complex-to-real and the Jacobian transformation for the source density given in (9). Since $H(\mathbf{x})$ is constant, using the mean ergodic theorem for the estimation of entropy, it is easy to see that minimization of mutual information is equivalent to ML, and when the weight matrix is constrained to be unitary, to the negentropy criterion.

All three approaches for achieving ICA, the ML, maximization of non-Gaussianity, and mutual information minimization require that the nonlinearity used in the algorithm is matched as closely as possible to the density for each estimated source.

For the circular case, *i.e.*, when $p_{s_n}(u) = g(|u|)$, the corresponding entry of the score function vector in ML ICA can be easily evaluated as

$$\psi_n(u) = -\frac{\partial \log g(\sqrt{uu^*})}{\partial u^*} = -\frac{u}{2|u|} \left(\frac{g'(|u|)}{g(|u|)} \right).$$

Thus, the score function always has the same phase as its argument. This is the form of the score function proposed in [3] where all sources are assumed to be circular.

When the source is a circular Gaussian source, the score function is linear $\psi(u) = u/2\sigma^2$ as expected, since circular Gaussian sources cannot be separated using ICA. However, when they are noncircular, the score function is nonlinear in u . Hence, noncircular Gaussian sources can be separated using ICA as long as they have unique circularity coefficients.

Given the richer structure of possible distributions in the two-dimensional space compared to the real-valued, *i.e.*, one-dimensional case, the pdf estimation problem becomes more challenging for complex-valued ICA. In this case, in addition to the sub- or super-Gaussian nature of the sources, the circularity properties of the sources have to be taken into account as well. Fully complex nonlinear functions can be shown to provide robust performance for most cases [2]. By using Wirtinger calculus, we can easily extend efficient nonlinear ICA algorithms to the complex domain without the need to make simplifying assumptions, the most common of which is the assumption of circularity [3, 5]. Thus, we can develop an effective class of ICA algorithms for the complex domain for any given source distribution. When the sources are improper with distinct circularity coefficients, we can efficiently separate them—including Gaussian sources—using only second-order statistics.

5. REFERENCES

- [1] T. Adalı, H. Li, M. Novey, and J.-F. Cardoso, "Complex ICA using nonlinear functions," *IEEE Trans. Signal Processing*, vol. 56, no. 9, pp. 4356–4544, Sept. 2008.
- [2] T. Adalı, T. Kim, and V. Calhoun, "Independent component analysis by complex nonlinearities," in *Proc. ICASSP*, vol. V, Montreal, Canada, May 2004, pp. 525–528.
- [3] J. Anemüller, T. J. Sejnowski, and S. Makeig, "Complex independent component analysis of frequency-domain electroencephalographic data," *Neural Networks*, vol. 16, pp. 1311–1323, 2003.
- [4] A. Bell and T. Sejnowski, "An information maximization approach to blind separation and blind deconvolution," *Neural Computation*, vol. 7, pp. 1129–1159, 1995.
- [5] E. Bingham and A. Hyvärinen, "A fast fixed-point algorithm for independent component analysis of complex valued signals," *Int. J. Neural Systems*, vol. 10, pp. 1–8, 2000.
- [6] J.-F. Cardoso and A. Souloumiac, "Blind beamforming for non-Gaussian signals," *IEE Proc. Radar Signal Processing*, vol. 140, pp. 362–370, 1993.
- [7] J. Eriksson and V. Koivunen, "Complex random vectors and ICA models: Identifiability, uniqueness, and separability," *IEEE Trans. Inform. Theory*, vol. 52, no. 3, pp. 1017–1029, Mar. 2006.
- [8] N. R. Goodman, "Statistical analysis based on a certain multivariate complex Gaussian distribution," *Annals Math. Stats.*, vol. 34, pp. 152–176, 1963.
- [9] R. A. Horn, C. R. Johnson, *Matrix Analysis*, Cambridge, UK: Cambridge University Press, 1985.
- [10] H. Hotelling, "Relations between two sets of variates," *Biometrika*, vol. 28, pp. 321–377, 1936.
- [11] A. Hyvärinen, "Fast and robust fixed-point algorithms for independent component analysis," *IEEE Trans. Neural Networks*, vol. 10, no. 3, pp. 626–634, 1999.
- [12] E. Ollila and V. Koivunen, "Complex ICA using generalized uncorrelating transform," *Signal Processing*, vol. 89, pp. 365–377, Apr. 2009.
- [13] D. Pham and P. Garat, "Blind separation of mixtures of independent sources through a quasi maximum likelihood approach," *IEEE Trans. Signal Processing*, vol. 45, no. 7, pp. 1712–1725, 1997.
- [14] P. J. Schreier, L. L. Scharf, and A. Hanssen, "A generalized likelihood ratio test for impropriety of complex signals," *IEEE Signal Processing Lett.*, vol. 13, no. 7, pp. 433–436, July 2006.
- [15] P. J. Schreier, "Bounds on the degree of impropriety of complex random vectors," *IEEE Signal Processing Lett.*, vol. 15, pp. 190–193, 2008.