Cramér-Rao Bound for Circular Complex Independent Component Analysis

Benedikt Loesch and Bin Yang

Institute for Signal Processing and System Theory, University of Stuttgart
{benedikt.loesch, bin.yang}@iss.uni-stuttgart.de

Abstract. Despite an increased interest in complex independent component analysis (ICA) during the last two decades, a closed-form expression for the Cramér-Rao bound (CRB) of the complex ICA problem has not yet been established. In this paper, we fill this gap for the noiseless case and circular sources. The CRB depends on the distributions of the sources only through two characteristic values which can be easily calculated. In addition, we study the CRB for the family of circular complex generalized Gaussian distributions (GGD) in more detail and compare it to simulation results using several ICA estimators.

Key words: Cramér-Rao bound, Fisher Information, independent component analysis, blind source separation, circular complex distribution

1 Introduction

Independent Component Analysis (ICA) is a relatively recent signal processing method to extract unobservable source signals or independent components from their observed linear mixtures. We assume a linear square noiseless mixing model

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

where $\mathbf{x} \in \mathbb{C}^N$ are N linear combinations of the N source signals $\mathbf{s} \in \mathbb{C}^N$. We make the following assumptions:

A1. The mixing matrix $\mathbf{A} \in \mathbb{C}^{N \times N}$ is deterministic and invertible.
A2. $\mathbf{s} = [s_1, \cdots, s_N]^T \in \mathbb{C}^N$ are N independent random variables with zero mean and unit variance (after scaling the rows of $\mathbf{A}$ suitably). The probability density functions (pdfs) $p_i(s_i)$ of s_i can be different. We assume the sources to be circular, i.e. $p_i(s_i) = p_i(s_i e^{j\alpha})\ \forall \alpha \in \mathbb{R}$. Hence $E[s_i^2] = 0$. Furthermore, $p_i(s_i)$ is continuously differentiable with respect to s_i and s_i^* in the sense of Wirtinger derivatives [1] which will be introduced in Sect. 2. The expectations in (15) and (20) exist.

The task of ICA is to demix the signals $\mathbf{x}$ by a demixing matrix $\mathbf{W} \in \mathbb{C}^{N \times N}$

$$\mathbf{y} = \mathbf{W}\mathbf{x} = \mathbf{W}\mathbf{A}\mathbf{s} \tag{2}$$

such that $\mathbf{y}$ is "as close to $\mathbf{s}$" as possible according to some metric. The ideal solution for $\mathbf{W}$ is $\mathbf{A}^{-1}$ neglecting scaling, phase and permutation ambiguity [2].

It is very useful, to have a lower bound for the variance of estimation of $\mathbf{W}$. The Cramér-Rao bound (CRB) provides a lower bound on the covariance matrix of any unbiased estimator of a parameter vector. Although much research in the field of ICA has been undertaken, a closed-form expression for the CRB of the real instantaneous ICA problem has been derived only recently [3, 4]. However, in many practical applications, such as telecommunication or audio processing in frequency domain, the signals are complex. Although many different algorithms for complex ICA have been proposed [5–9], the CRB for this problem has not yet been established. In this paper, we fill this gap by deriving closed-form expressions for the CRB of the vectorized parameter $\boldsymbol{\theta} = \mathrm{vec}(\mathbf{W}^T)$ and for the CRB of $\boldsymbol{\vartheta} = \mathrm{vec}((\mathbf{WA})^T)$. Due to the intrinsic phase ambiguity in circular complex ICA (cf. A2.: $p_i(s_i) = p_i(s_i e^{j\alpha})\ \forall \alpha \in \mathbb{R}$), we can only derive a CRB with the constraint $[\mathbf{WA}]_{ii} \in \mathbb{R}$. The CRB depends on the distributions of the sources only through two scalars defined in (15) which can be easily calculated.

2 Prerequisites

2.1 Complex Functions and Complex Random Vectors

Define the partial derivative of a complex function $\mathbf{g}(\boldsymbol{\theta}) = \mathbf{u}(\boldsymbol{\alpha}, \boldsymbol{\beta}) + j\mathbf{v}(\boldsymbol{\alpha}, \boldsymbol{\beta})$ with respect to $\boldsymbol{\alpha} = \Re[\boldsymbol{\theta}]$ as $\partial\mathbf{g}/\partial\boldsymbol{\alpha} = \partial\mathbf{u}/\partial\boldsymbol{\alpha} + j\partial\mathbf{v}/\partial\boldsymbol{\alpha}$ and with respect to $\boldsymbol{\beta} = \Im[\boldsymbol{\theta}]$ as $\partial\mathbf{g}/\partial\boldsymbol{\beta} = \partial\mathbf{u}/\partial\boldsymbol{\beta} + j\partial\mathbf{v}/\partial\boldsymbol{\beta}$. Then the complex partial differential operators $\partial/\partial\boldsymbol{\theta}$ and $\partial/\partial\boldsymbol{\theta}^*$ are defined as

$$\frac{\partial \mathbf{g}}{\partial \boldsymbol{\theta}} = \frac{1}{2}\left(\frac{\partial \mathbf{g}}{\partial \boldsymbol{\alpha}} - j\frac{\partial \mathbf{g}}{\partial \boldsymbol{\beta}}\right), \quad \frac{\partial \mathbf{g}}{\partial \boldsymbol{\theta}^*} = \frac{1}{2}\left(\frac{\partial \mathbf{g}}{\partial \boldsymbol{\alpha}} + j\frac{\partial \mathbf{g}}{\partial \boldsymbol{\beta}}\right). \tag{3}$$

These differential operators have first been introduced for real valued $\mathbf{g}$ by Wirtinger [1]. As long as the real and imaginary part of a complex function $\mathbf{g}$ are real-differentiable, the two Wirtinger derivatives in (3) also exist [10]. The direction of steepest descent of a real function $\mathbf{g}(\boldsymbol{\theta}) = \mathbf{u}(\boldsymbol{\alpha}, \boldsymbol{\beta})$ is given by $\frac{\partial \mathbf{g}}{\partial \boldsymbol{\theta}^*}$ and not $\frac{\partial \mathbf{g}}{\partial \boldsymbol{\theta}}$ [11]. The complex Jacobian matrix of a complex function $\mathbf{g}\colon \mathbb{C}^M \to \mathbb{C}^N$ is defined as the complex $2N \times 2M$ matrix

$$\mathbf{D}_{\mathbf{g}} = \begin{bmatrix} \frac{\partial \mathbf{g}}{\partial \boldsymbol{\theta}} & \frac{\partial \mathbf{g}}{\partial \boldsymbol{\theta}^*} \\ \left(\frac{\partial \mathbf{g}}{\partial \boldsymbol{\theta}^*}\right)^* & \left(\frac{\partial \mathbf{g}}{\partial \boldsymbol{\theta}}\right)^* \end{bmatrix}, \tag{4}$$

i.e. it is the augmented matrix of $\partial\mathbf{g}/\partial\boldsymbol{\theta}$ and $\partial\mathbf{g}/\partial\boldsymbol{\theta}^*$. The covariance matrix of a complex random vector $\mathbf{x} = \mathbf{x}_R + j\mathbf{x}_I \in \mathbb{C}^N$ is $\mathrm{cov}(\mathbf{x}) = E\big[(\mathbf{x} - E[\mathbf{x}])(\mathbf{x} - E[\mathbf{x}])^H\big]$. The pseudo-covariance matrix of $\mathbf{x}$ is $\mathrm{pcov}(\mathbf{x}) = E\big[(\mathbf{x} - E[\mathbf{x}])(\mathbf{x} - E[\mathbf{x}])^T\big]$.

2.2 Cramér-Rao Bound for a Complex Parameter

We briefly review the CRB for complex parameters (see for example [12]) before we derive the CRB for circular complex ICA. Assume that L observations of

$\mathbf{x}$ are i.i.d. distributed having the pdf $p(\mathbf{x};\boldsymbol{\theta})$ with parameter vector $\boldsymbol{\theta}$. The complex Fisher Information Matrix (FIM) of complex parameter $\boldsymbol{\theta}$ is defined as

$$\mathbf{J}_{\boldsymbol{\theta}} = \begin{bmatrix} \mathcal{I}_{\boldsymbol{\theta}} & \mathcal{P}_{\boldsymbol{\theta}} \\ \mathcal{P}^*_{\boldsymbol{\theta}} & \mathcal{I}^*_{\boldsymbol{\theta}} \end{bmatrix}, \tag{5}$$

where $\mathcal{I}_{\boldsymbol{\theta}} = E\left[\nabla_{\boldsymbol{\theta}^*} \log p(\mathbf{x};\boldsymbol{\theta})\{\nabla_{\boldsymbol{\theta}^*} \log p(\mathbf{x};\boldsymbol{\theta})\}^H\right]$ is called the information matrix and $\mathcal{P}_{\boldsymbol{\theta}} = E\left[\nabla_{\boldsymbol{\theta}^*} \log p(\mathbf{x};\boldsymbol{\theta})\{\nabla_{\boldsymbol{\theta}^*} \log p(\mathbf{x};\boldsymbol{\theta})\}^T\right]$ the pseudo-information matrix. Here $\nabla_{\boldsymbol{\theta}^*} \log p(\mathbf{x};\boldsymbol{\theta}) = \frac{1}{2}\left(\nabla_{\boldsymbol{\alpha}} \log p(\mathbf{x};\boldsymbol{\theta}) + j\nabla_{\boldsymbol{\beta}} \log p(\mathbf{x};\boldsymbol{\theta})\right)$ is the column gradient vector of $\log p(\mathbf{x};\boldsymbol{\theta})$, i.e. $[\partial/\partial\theta_1^*, \cdots, \partial/\partial\theta_N^*]^T \log p(\mathbf{x};\boldsymbol{\theta})$.

The inverse of the FIM of $\boldsymbol{\theta}$ gives, under regularity conditions, the CRB of the augmented covariance matrix of an unbiased estimator $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ and hence

$$\begin{bmatrix} \mathrm{cov}(\hat{\boldsymbol{\theta}}) & \mathrm{pcov}(\hat{\boldsymbol{\theta}}) \\ \mathrm{pcov}(\hat{\boldsymbol{\theta}})^* & \mathrm{cov}(\hat{\boldsymbol{\theta}})^* \end{bmatrix} \geq L^{-1}\mathbf{J}_{\boldsymbol{\theta}}^{-1} = \frac{1}{L}\begin{bmatrix} \mathcal{I}_{\boldsymbol{\theta}} & \mathcal{P}_{\boldsymbol{\theta}} \\ \mathcal{P}^*_{\boldsymbol{\theta}} & \mathcal{I}^*_{\boldsymbol{\theta}} \end{bmatrix}^{-1}. \tag{6}$$

It holds $\mathrm{cov}(\hat{\boldsymbol{\theta}}) \geq L^{-1}(\mathcal{I}_{\boldsymbol{\theta}} - \mathcal{P}_{\boldsymbol{\theta}}\mathcal{I}_{\boldsymbol{\theta}}^{-*}\mathcal{P}^*_{\boldsymbol{\theta}})^{-1} = L^{-1}\mathbf{R}_{\boldsymbol{\theta}}^{-1}$ with $\mathbf{R}_{\boldsymbol{\theta}} = \mathcal{I}_{\boldsymbol{\theta}} - \mathcal{P}_{\boldsymbol{\theta}}\mathcal{I}_{\boldsymbol{\theta}}^{-*}\mathcal{P}^*_{\boldsymbol{\theta}}$. The CRB for a transformed vector $\boldsymbol{\vartheta} = \mathbf{g}(\boldsymbol{\theta})$ is given by the right-hand-side of

$$\begin{bmatrix} \mathrm{cov}(\hat{\boldsymbol{\vartheta}}) & \mathrm{pcov}(\hat{\boldsymbol{\vartheta}}) \\ \mathrm{pcov}(\hat{\boldsymbol{\vartheta}})^* & \mathrm{cov}(\hat{\boldsymbol{\vartheta}})^* \end{bmatrix} \geq L^{-1}\mathbf{D}_{\mathbf{g}}\mathbf{J}_{\boldsymbol{\theta}}^{-1}\mathbf{D}_{\mathbf{g}}^T. \tag{7}$$

3 Derivation of Cramér-Rao Bound

In ICA, the parameter of interest is the demixing matrix $\mathbf{W}$. We form the parameter vector $\boldsymbol{\theta} = \mathrm{vec}(\mathbf{W}^T) = [\mathbf{w}_1^T, \cdots, \mathbf{w}_N^T]^T \in \mathbb{C}^{N^2}$, where $\mathbf{w}_i^T \in \mathbb{C}^N$ are the row vectors of $\mathbf{W}$. The operator $\mathrm{vec}(\cdot)$ stacks the columns of its argument into one long column vector. The pdf of $\mathbf{x} = \mathbf{A}\mathbf{s}$ is defined as $p(\mathbf{x};\boldsymbol{\theta}) = |\det(\mathbf{W})|^2 \prod_{i=1}^N p_i(\mathbf{w}_i\mathbf{x})$, where $p_i(s_i)$ denotes the pdf of s_i and $\mathbf{W} = \mathbf{A}^{-1}$. By using matrix derivatives, we obtain

$$\frac{\partial}{\partial \mathbf{W}^H} \log p(\mathbf{x};\boldsymbol{\theta}) = \mathbf{A}^* - \mathbf{x}^*\boldsymbol{\varphi}^T(\mathbf{W}\mathbf{x}) = \mathbf{A}^*(\mathbf{I} - \mathbf{s}\boldsymbol{\varphi}^H(\mathbf{s}))^* \tag{8}$$

where $\boldsymbol{\varphi}(\mathbf{s}) = [\varphi_1(s_1), \cdots, \varphi_N(s_N)]^T$ and $\varphi_i(s_i) = -\frac{\partial}{\partial s_i^*} \log p_i(s_i)$.

Since $\boldsymbol{\theta} = \mathrm{vec}(\mathbf{W}^T)$, we get $\nabla_{\boldsymbol{\theta}^*} \log p(\mathbf{x};\boldsymbol{\theta}) = \mathrm{vec}\left(\frac{\partial}{\partial \mathbf{W}^H} \log p(\mathbf{x};\boldsymbol{\theta})\right)$ and

$$\mathcal{I}_{\boldsymbol{\theta}} = \left((\mathbf{I}\otimes\mathbf{A})\mathbf{M}_1(\mathbf{I}\otimes\mathbf{A}^H)\right)^*, \quad \mathcal{P}_{\boldsymbol{\theta}} = \left((\mathbf{I}\otimes\mathbf{A})\mathbf{M}_2(\mathbf{I}\otimes\mathbf{A}^T)\right)^*, \tag{9}$$

where $\mathbf{M}_1 = E\left[\mathrm{vec}\{\mathbf{I} - \mathbf{s}\boldsymbol{\varphi}^H(\mathbf{s})\}\mathrm{vec}\{\mathbf{I} - \mathbf{s}\boldsymbol{\varphi}^H(\mathbf{s})\}^H\right]$, $\mathbf{M}_2 = E\left[\mathrm{vec}\{...\}\mathrm{vec}\{...\}^T\right]$ and $\otimes$ denotes the Kronecker product.

3.1 CRB for $\mathbf{G} = \mathbf{W}\mathbf{A}$

For simplicity, we first derive the CRB for the transformed parameter $\boldsymbol{\vartheta} = \mathrm{vec}((\mathbf{W}\mathbf{A})^T) = (\mathbf{I}\otimes\mathbf{A}^T)\boldsymbol{\theta}$. The covariance of $\hat{\boldsymbol{\vartheta}} = \mathrm{vec}((\hat{\mathbf{W}}\mathbf{A})^T)$ is given by $\mathrm{cov}(\hat{\boldsymbol{\vartheta}}) = (\mathbf{I}\otimes\mathbf{A}^T)\mathrm{cov}(\hat{\boldsymbol{\theta}})(\mathbf{I}\otimes\mathbf{A}^*)$ where $\hat{\boldsymbol{\theta}} = \mathrm{vec}(\hat{\mathbf{W}}^T)$. Hence it holds

$$\mathrm{cov}(\hat{\boldsymbol{\vartheta}}) \geq L^{-1}(\mathbf{I}\otimes\mathbf{A}^T)(\mathcal{I}_{\boldsymbol{\theta}} - \mathcal{P}_{\boldsymbol{\theta}}\mathcal{I}_{\boldsymbol{\theta}}^{-*}\mathcal{P}^*_{\boldsymbol{\theta}})^{-1}(\mathbf{I}\otimes\mathbf{A}^*) = L^{-1}\mathbf{R}_{\boldsymbol{\vartheta}}^{-1} \tag{10}$$

with $\mathbf{R}_{\boldsymbol{\vartheta}} = (\mathbf{M}_1 - \mathbf{M}_2\mathbf{M}_1^{-*}\mathbf{M}_2^*)^*$.

As shown in the appendix, $\mathbf{R}_{\boldsymbol{\vartheta}} = \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j\neq i}}^{N} \left(\frac{\kappa_i \kappa_j - 1}{\kappa_j}\right) \mathbf{L}_{ii} \otimes \mathbf{L}_{jj}$, with $\kappa_i = E\left[|\varphi_i(s_i)|^2\right]$. $\mathbf{L}_{ii}$ denotes an $N \times N$ matrix with a 1 at the (i,i) position and 0's elsewhere. $\mathbf{R}_{\boldsymbol{\vartheta}}$ is a diagonal matrix of rank $N^2 - N$. The CRB for G_{ij} yields

$$\mathrm{var}(\hat{G}_{ij}) \geq \frac{1}{L} \frac{\kappa_j}{\kappa_i \kappa_j - 1} \quad i \neq j \tag{11}$$

where $\hat{\mathbf{G}} = \hat{\mathbf{W}}\mathbf{A}$. Eq. (11) looks the same as in the real case [3, 4], but in the complex case κ_i is defined using Wirtinger derivatives instead of real derivatives.

Due to the phase ambiguity in circular complex ICA, the Fisher information for the diagonal elements G_{ii} is 0 and hence their CRB does not exist. However, we can constrain G_{ii} to be real and derive the constrained CRB [13] for $\theta = G_{ii}$: The constraint can be formulated as $f(\theta) = \theta - \theta^* = 0$. We then need to calculate $\mathbf{F}(\theta) = \begin{bmatrix} \partial f/\partial\theta & \partial f/\partial\theta^* \\ \partial f^*/\partial\theta & \partial f^*/\partial\theta^* \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$ and find an orthonormal 2×1 matrix $\mathbf{U}$ in the null-space of $\mathbf{F}(\theta)$, i.e. $\mathbf{FU} = \mathbf{0}$. We choose $\mathbf{U} = 1/\sqrt{2}\left[1\ 1\right]^T$. The CRB for the constrained parameter $\theta = G_{ii}$ then yields

$$\begin{bmatrix} \mathrm{var}(\theta) & \mathrm{pvar}(\theta) \\ \mathrm{pvar}^*(\theta) & \mathrm{var}(\theta) \end{bmatrix} \geq \frac{1}{L}\mathbf{U}\left(\mathbf{U}^H \begin{bmatrix} \mathcal{I}_\theta & \mathcal{P}_\theta \\ \mathcal{P}_\theta^* & \mathcal{I}_\theta \end{bmatrix} \mathbf{U}\right)^{-1} \mathbf{U}^H = \frac{1}{4L(\eta_i - 1)} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \tag{12}$$

where $\mathcal{I}_\theta = \eta_i - 1 = \mathcal{P}_\theta$ and $\eta_i = E\left[|s_i|^2 |\varphi_i(s_i)|^2\right]$. The CRB in (12) is valid for a phase-corrected G_{ii} such that $G_{ii} \in \mathbb{R}$. Eq. (12) matches the real case [3, 4], where $\mathrm{var}(\hat{G}_{ij}) \geq L^{-1}(\bar{\eta}_i - 1)^{-1}$ since η_i is defined using Wirtinger derivatives instead of real derivatives and hence for the real case $4(\eta_i - 1) = \bar{\eta}_i - 1$.

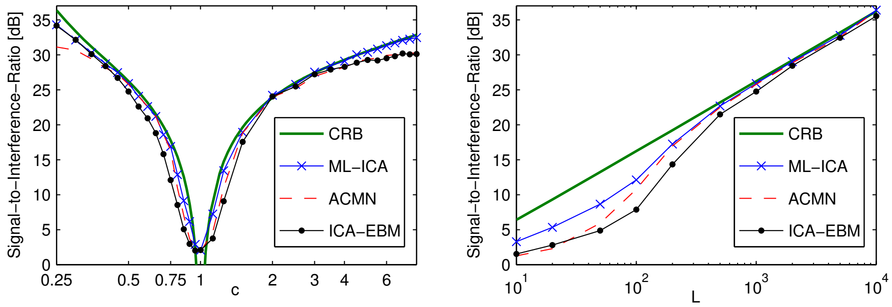
Performance of ICA is often measured using $\hat{\mathbf{G}}$ and hence it can be directly compared to (11), (12). The absolute values of the diagonal elements $|\hat{G}_{ii}|$ should be close to 1. They reflect how well we can estimate the power of each component. The absolute values of the off-diagonal elements $|\hat{G}_{ij}|$ should be close to 0 and reflect how well we can suppress interfering components.

3.2 CRB for W

It holds $\mathrm{vec}(\mathbf{W}^T) = \boldsymbol{\theta} = (\mathbf{I} \otimes \mathbf{A}^T)^{-1}\boldsymbol{\vartheta} = (\mathbf{I} \otimes \mathbf{W}^T)\boldsymbol{\vartheta}$ since $\mathbf{W} = \mathbf{A}^{-1}$. We can estimate the rows of $\mathbf{W}$ only up to an arbitrary phase for each row. We can derive a CRB for the phase-constrained $\mathbf{W}$, for which $[\mathbf{WA}]_{ii} \in \mathbb{R}$: We use the CRB for the constrained G_{ii} (12) together with the CRB for G_{ij} (11) to form the inverse FIM for the constrained $\mathbf{G}$ as $\mathbf{R}_{\boldsymbol{\vartheta}}^{-1} = \sum_{i=1}^{N} \frac{1}{4(\eta_i - 1)} \mathbf{L}_{ii} + \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j\neq i}}^{N} \left(\frac{\kappa_i \kappa_j - 1}{\kappa_j}\right) \mathbf{L}_{ii} \otimes \mathbf{L}_{jj}$. The CRB for constrained $\mathbf{W}$ is then given by $\mathbf{R}_{\boldsymbol{\theta}}^{-1} = (\mathbf{I} \otimes \mathbf{W}^T)\mathbf{R}_{\boldsymbol{\vartheta}}^{-1}(\mathbf{I} \otimes \mathbf{W}^*)$ and $\mathrm{cov}(\hat{\boldsymbol{\theta}}) \geq L^{-1}\mathbf{R}_{\boldsymbol{\theta}}^{-1}$.

4 Results for Generalized Gaussian Distribution (GGD)

A circular complex GGD with zero mean and variance $E[|s|^2] = 1$ is given by the pdf $p(s, s^*) = \frac{c\alpha}{\pi\Gamma(1/c)} \exp\left(-\left[\alpha s s^*\right]^c\right)$ [14], with $\alpha = (\Gamma(2/c))/(\Gamma(1/c))$. $\Gamma(\cdot)$



(a) Varying shape parameter c, $L = 1000$ (b) Varying sample size L, $c = 0.5$

Fig. 1: Comparison of performance of three ICA estimators with CRB

denotes the Gamma function. The shape parameter $c > 0$ varies the form of the pdf from super-Gaussian ($c < 1$) to sub-Gaussian ($c > 1$). For $c = 1$, the pdf is Gaussian. By integration in polar coordinates, we find κ, η and β in (15) and (20) as $\kappa = \frac{c^2 \Gamma(2/c)}{\Gamma^2(1/c)}$, $\eta = \beta = c+1$. For the simulation study, we consider $N = 3$ identically distributed sources with random mixing matrices $\mathbf{A}$ with independent uniform distributions for the real and imaginary parts of each entry (between -1 and 1). We conduct 100 experiments with different $\mathbf{A}$ and different realizations of the source signals and consider the following different ICA estimators: Complex ML-ICA [7], adaptable complex maximization of nongaussianity (ACMN) [9] and complex ICA by entropy bound minimization (ICA-EBM) [8]. We correct for permutation ambiguity and then calculate the signal-to-interference ratio (SIR) averaged over all N sources: $\text{SIR} = \frac{1}{N} \sum_i \left(E\left[|G_{ii}|^2\right] / \sum_{j \neq i} E\left[|G_{ij}|^2\right] \right)$. Fig. 1 (a) compares the SIR given by the CRB with the empirical SIR of the different ICA estimators for varying shape parameter c and a sample size of $L = 1000$. Since all sources are identically distributed, $\text{CRB}(G_{ij}) \to \infty$ and $\text{SIR} \to 0$ for $c \to 1$ (Gaussian). In this case, ICA fails to separate the sources. Clearly, the performance of complex ML-ICA is close to the CRB for a wide range of the shape parameter c. ACMN outperforms ICA-EBM in most cases except for strongly super-Gaussian sources: ACMN uses a GGD model and hence is better suited for separating circular GGD sources. However, ACMN uses prewhitening and then constrains the demixing matrix to be unitary which ICA-EBM does not. Fig. 1 (b) studies the influence of sample size L on ICA performance for $c = 0.5$. Again, complex ML-ICA performs the best as expected. Except for small sample sizes, all algorithms come quite close to the CRB.

5 Conclusion

In this paper, we have derived the CRB for the noiseless ICA problem with circular complex sources. Due to the phase ambiguity in circular complex ICA, the CRB for the diagonal elements of the demixing-mixing-matrix-product $\mathbf{G} = \mathbf{WA}$

does not exist, but a constrained CRB with $G_{ii} \in \mathbb{R}$ can be derived. Simulation results with sources following a circular complex generalized Gaussian distribution have shown that for large enough sample size some ICA estimators can achieve a signal-to-interference ratio close to that given by the CRB.

A Useful Matrix Algebra

Similarly to [4], we make use of some matrix algebra in the derivation of the CRB. We briefly review the required properties here: Let $\mathbf{L}_{ij}$ denote a $N \times N$ matrix with a 1 at the (i,j) position and 0's elsewhere. It is useful to note that

$$\mathbf{A}\mathbf{L}_{ij}\mathbf{A}^T = \mathbf{a}_i\mathbf{a}_j^T, \quad \mathbf{L}_{ij}\mathbf{L}_{kl} = \mathbf{0} \text{ for } j \neq k, \quad \mathbf{L}_{ij}\mathbf{L}_{jl} = \mathbf{L}_{il} \tag{13}$$

where $\otimes$ denotes the Kronecker product. We also note that any $N^2 \times N^2$ block matrix $\mathbf{A}$ can be written using its $N \times N$ diagonal blocks $\mathbf{A}[i,i]$ and $N \times N$ off-diagonal blocks $\mathbf{A}[i,j], i \neq j$ as follows:

$$\mathbf{A} = \sum_{i=1}^{N} \mathbf{L}_{ii} \otimes \mathbf{A}[i,i] + \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} \mathbf{L}_{ij} \otimes \mathbf{A}[i,j]. \tag{14}$$

B Some Steps in the Derivation of the CRB for G

The derivation of the CRB for $\mathbf{G}$, proceeds in three steps: First, we calculate $\mathbf{M}_1$ and $\mathbf{M}_2$. Then, we obtain $\mathbf{R}_{\boldsymbol{\vartheta}} = (\mathbf{M}_1 - \mathbf{M}_2\mathbf{M}_1^{-*}\mathbf{M}_2^*)^*$ and finally invert $\mathbf{R}_{\boldsymbol{\vartheta}}$.

Using $E[\mathbf{s}\boldsymbol{\varphi}^H(\mathbf{s})] = \mathbf{I}$, we can simplify $\mathbf{M}_1$ as

$$\mathbf{M}_1 = E\left[\text{vec}\{\mathbf{I} - \mathbf{s}\boldsymbol{\varphi}^H(\mathbf{s})\}\text{vec}\{\mathbf{I} - \mathbf{s}\boldsymbol{\varphi}^H(\mathbf{s})\}^H\right] = \boldsymbol{\Omega}_1 - \text{vec}\{\mathbf{I}\}\text{vec}\{\mathbf{I}\}^H,$$

where $\boldsymbol{\Omega}_1 = E\left[\text{vec}\{\mathbf{s}\boldsymbol{\varphi}^H(\mathbf{s})\}\text{vec}\{\mathbf{s}\boldsymbol{\varphi}^H(\mathbf{s})\}^H\right]$ is a $N^2 \times N^2$ block matrix. The (i,i) block $\boldsymbol{\Omega}_1[i,i] = E\left[\mathbf{s}\mathbf{s}^H|\varphi_i(s_i)|^2\right]$ is diagonal since the components of $\mathbf{s}$ are independent and zero mean. The diagonal elements $\boldsymbol{\Omega}_1[i,i]_{(j,j)}$ are given by

$$\boldsymbol{\Omega}_1[i,i]_{(j,j)} = \begin{cases} E\left[|s_i|^2|\varphi_i(s_i)|^2\right] =: \eta_i & i = j \\ E\left[|s_j|^2|\varphi_i(s_i)|^2\right] = E\left[|\varphi_i(s_i)|^2\right] =: \kappa_i & i \neq j \end{cases}. \tag{15}$$

κ_i and η_i are real since $E[g(s)]$ with $g(s) \in \mathbb{R}$ is real. The (i,j) block $\boldsymbol{\Omega}_1[i,j]$ $(i \neq j)$ can be calculated as $\boldsymbol{\Omega}_1[i,j] = E\left[\mathbf{s}\mathbf{s}^H\varphi_i^*(s_i)\varphi_j(s_j)\right]$. It has 1 at entry (i,j) and 0 at entry (j,i), since

$$\boldsymbol{\Omega}_1[i,j]_{(i,j)} = E\left[s_i s_j^* \varphi_i^*(s_i)\varphi_j(s_j)\right] = E\left[s_i\varphi_i^*(s_i)\right]E\left[s_j^*\varphi_j(s_j)\right] = 1, \tag{16}$$

$$\boldsymbol{\Omega}_1[i,j]_{(j,i)} = E\left[s_i^* s_j \varphi_i^*(s_i)\varphi_j(s_j)\right] = E\left[s_i^*\varphi_i^*(s_i)\right]E\left[s_j\varphi_j(s_j)\right] = 0. \tag{17}$$

All other entries of $\boldsymbol{\Omega}_1[i,j]$ are zero since the components of $\mathbf{s}$ are independent and zero mean. Using the matrix algebra from appendix A, we can write $\boldsymbol{\Omega}_1$ as

$$\boldsymbol{\Omega}_1 = \sum_{i=1}^{N} \eta_i \mathbf{L}_{ii} \otimes \mathbf{L}_{ii} + \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} \kappa_i \mathbf{L}_{ii} \otimes \mathbf{L}_{jj} + \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} \mathbf{L}_{ij} \otimes \mathbf{L}_{ij}. \tag{18}$$

Using $\text{vec}\{\mathbf{I}\}\text{vec}\{\mathbf{I}\}^H = \sum_{i=1}^N \sum_{\substack{j=1\\j\neq i}}^N \mathbf{L}_{ij} \otimes \mathbf{L}_{ij} + \sum_{i=1}^N \mathbf{L}_{ii} \otimes \mathbf{L}_{ii}$, we get $\mathbf{M}_1$ as

$$\mathbf{M}_1 = \sum_{i=1}^N (\eta_i - 1)\mathbf{L}_{ii} \otimes \mathbf{L}_{ii} + \sum_{i=1}^N \sum_{\substack{j=1\\j\neq i}}^N \kappa_i \mathbf{L}_{ii} \otimes \mathbf{L}_{jj}. \tag{19}$$

We note that $\mathbf{M}_1$ is a real diagonal matrix.

$\mathbf{M}_2$ can be calculated similarly. It holds:

$$\mathbf{M}_2 = E\left[\text{vec}\{\mathbf{I} - \mathbf{s}\boldsymbol{\varphi}^H(\mathbf{s})\}\text{vec}\{\mathbf{I} - \mathbf{s}\boldsymbol{\varphi}^H(\mathbf{s})\}^T\right] = \boldsymbol{\Omega}_2 - \text{vec}\{\mathbf{I}\}\text{vec}\{\mathbf{I}\}^T,$$

where $\boldsymbol{\Omega}_2 = E\left[\text{vec}\{\mathbf{s}\boldsymbol{\varphi}^H(\mathbf{s})\}\text{vec}\{\mathbf{s}\boldsymbol{\varphi}^H(\mathbf{s})\}^T\right]$ is a $N^2 \times N^2$ block matrix. The (i,i) block $\boldsymbol{\Omega}_2[i,i] = E\left[\mathbf{s}\mathbf{s}^T(\varphi_i^*(s_i))^2\right]$ is diagonal since the components of $\mathbf{s}$ are independent and zero mean. The diagonal elements $\boldsymbol{\Omega}_2[i,i]_{(j,j)}$ are given by

$$\boldsymbol{\Omega}_2[i,i]_{(j,j)} = \begin{cases} E\left[s_i^2(\varphi_i^*(s_i))^2\right] =: \beta_i & i = j \\ E\left[s_j^2(\varphi_i^*(s_i)|^2\right] = E\left[s_j^2\right] E\left[(\varphi_i^*(s_i))^2\right] = 0 & i \neq j, \end{cases}. \tag{20}$$

since $E\left[s_j^2\right] = 0$. If s_i is circular, it can be shown that $\beta_i = \eta_i$: For circular $s = s_R + js_I$, $p(-s_R, s_I) = p(s_R, s_I)$, $p(s_R, -s_I) = p(s_R, s_I)$ and $p(s_R, s_I) = g(s_R^2 + s_I^2)$. Let $f(r^2) = f(s_R^2 + s_I^2) = \log p(s_R, s_I)$. It holds

$$\begin{aligned}
\eta &= \frac{1}{4}E\left[(s_R^2 + s_I^2)\left(\left(\frac{\partial f}{\partial s_R}\right)^2 + \left(\frac{\partial f}{\partial s_I}\right)^2\right)\right], \\
\beta &= \frac{1}{4}E\left[(s_R^2 - s_I^2)\left(\left(\frac{\partial f}{\partial s_R}\right)^2 - \left(\frac{\partial f}{\partial s_I}\right)^2\right) + 4s_R s_I \left(\frac{\partial f}{\partial s_R}\right)\left(\frac{\partial f}{\partial s_I}\right)\right], \\
4(\eta - \beta) &= 2E\left[s_R^2\left(\frac{\partial f}{\partial s_I}\right)^2 + s_I^2\left(\frac{\partial f}{\partial s_R}\right)^2 - 2s_R s_I\left(\frac{\partial f}{\partial s_R}\right)\left(\frac{\partial f}{\partial s_I}\right)\right] = 0,
\end{aligned}$$

where we used $E\left[s_R s_I\left(\left(\frac{\partial f}{\partial s_R}\right)^2 - \left(\frac{\partial f}{\partial s_I}\right)^2\right)\right] = 0$ and $E\left[(s_R^2 - s_I^2)\left(\frac{\partial f}{\partial s_R}\right)\left(\frac{\partial f}{\partial s_I}\right)\right] = 0$ in the third line and $\frac{\partial f}{\partial s_R} = 2s_R\frac{\partial f(r^2)}{\partial r^2}$ and $\frac{\partial f}{\partial s_I} = 2s_I\frac{\partial f(r^2)}{\partial r^2}$ in the last line.
The (i,j) block $\boldsymbol{\Omega}_2[i,j]$ $(i \neq j)$ can be calculated as $\boldsymbol{\Omega}_2[i,j] = E\left[\mathbf{s}\mathbf{s}^T\varphi_i^*(s_i)\varphi_j^*(s_j)\right]$. It has 1 at entry (i,j) and (j,i), since

$$\boldsymbol{\Omega}_2[i,j]_{(i,j)} = \boldsymbol{\Omega}_2[i,j]_{(j,i)} = E\left[s_i\varphi_i^*(s_i)\right] E\left[s_j\varphi_j^*(s_j)\right] = 1. \tag{21}$$

All other entries of $\boldsymbol{\Omega}_2[i,j]$ are zero since the components of $\mathbf{s}$ are independent and zero mean. Hence, we can calculate $\mathbf{M}_2 = \boldsymbol{\Omega}_2 - \text{vec}\{\mathbf{I}\}\text{vec}\{\mathbf{I}\}^T$ as

$$\mathbf{M}_2 = \sum_{i=1}^N (\beta_i - 1)\mathbf{L}_{ii} \otimes \mathbf{L}_{ii} + \sum_{i=1}^N \sum_{\substack{j=1\\j\neq i}}^N (\mathbf{L}_{ij} \otimes \mathbf{L}_{ji}). \tag{22}$$

We note that $\mathbf{M}_2$ is a real diagonal matrix.

Since $\mathbf{M}_1$ and $\mathbf{M}_2$ are real matrices, it holds $\mathbf{R}_{\boldsymbol{\vartheta}} = (\mathbf{M}_1 - \mathbf{M}_2\mathbf{M}_1^{-*}\mathbf{M}_2^*)^* = \mathbf{M}_1 - \mathbf{M}_2\mathbf{M}_1^{-1}\mathbf{M}_2$. After some calculations, we get

$$\mathbf{R}_{\boldsymbol{\vartheta}} = \sum_{i=1}^{N} \frac{(\eta_i - 1)^2 - (\beta_i - 1)^2}{\eta_i - 1}\mathbf{L}_{ii} \otimes \mathbf{L}_{ii} + \sum_{i=1}^{N}\sum_{\substack{j=1 \\ j\neq i}}^{N} \left(\frac{\kappa_i\kappa_j - 1}{\kappa_j}\right)\mathbf{L}_{ii} \otimes \mathbf{L}_{jj} \quad (23)$$

which simplifies to $\mathbf{R}_{\boldsymbol{\vartheta}} = \sum_{i=1}^{N}\sum_{\substack{j=1 \\ j\neq i}}^{N} \left(\frac{\kappa_i\kappa_j-1}{\kappa_j}\right)\mathbf{L}_{ii} \otimes \mathbf{L}_{jj}$ due to $\beta_i = \eta_i$.

References

1. W. Wirtinger, "Zur formalen Theorie der Funktionen von mehr komplexen Veränderlichen," *Mathematische Annalen*, vol. 97, no. 1, pp. 357–375, 1927.
2. J. Eriksson and V. Koivunen, "Complex random vectors and ica models: identifiability, uniqueness, and separability," *Information Theory, IEEE Transactions on*, vol. 52, no. 3, pp. 1017 –1029, March 2006.
3. P. Tichavsky, Z. Koldovsky, and E. Oja, "Performance analysis of the FastICA algorithm and Cramér-Rao bounds for linear independent component analysis," *IEEE Trans. on Sig. Proc.*, vol. 54, no. 4, 2006.
4. E. Ollila, H.-J. Kim, and V. Koivunen, "Compact Cramér-Rao bound expression for independent component analysis," *IEEE Trans. on Sig. Proc.*, vol. 56, no. 4, 2008.
5. L. De Lathauwer and B. De Moor, "On the blind separation of non-circular sources," in *EUSIPCO-02*, Toulouse, France, sept. 2002.
6. Scott C. Douglas, "Fixed-point algorithms for the blind separation of arbitrary complex-valued non-gaussian signal mixtures," *EURASIP J. Appl. Signal Process.*, vol. 2007, no. 1, January 2007.
7. H. Li and T. Adali, "Algorithms for complex ML ICA and their stability analysis using Wirtinger calculus," *IEEE Trans. on Sig. Proc.*, vol. 58, no. 12, pp. 6156–6167, Dec. 2010.
8. Xi-Lin Li and T. Adali and, "Complex independent component analysis by entropy bound minimization," *Circuits and Systems I: Regular Papers, IEEE Transactions on*, vol. 57, no. 7, pp. 1417 –1430, july 2010.
9. M. Novey and T. Adali, "Adaptable nonlinearity for complex maximization of nongaussianity and a fixed-point algorithm," *Proc. IEEE Workshop on Machine Learning for Signal Processing*, Sept. 2006.
10. R. Remmert, *Theory of complex functions*, Graduate texts in mathematics. Springer-Verlag, 1991.
11. D. H. Brandwood, "A complex gradient operator and its application in adaptive array theory," *IEE Proc.*, vol. 130, pp. 11–16, 1983.
12. E. Ollila, V. Koivunen, and J. Eriksson, "On the Cramér-Rao bound for the constrained and unconstrained complex parameters," *Proc. IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, 2008.
13. A.K. Jagannatham and B.D. Rao, "Cramér-Rao lower bound for constrained complex parameters," *IEEE Sig. Proc. Letters*, vol. 11, no. 11, nov. 2004.
14. M. Novey, T. Adali, and A. Roy, "A complex generalized gaussian distribution – characterization, generation, and estimation," *IEEE Trans. on Sig. Proc.*, vol. 58, no. 3, pp. 1427 –1433, March 2010.