

Article

Complex-Valued FastICA Estimator with a Weighted Unitary Constraint: A Robust and Equivariant Estimator

Jianwei E  and Mingshu Yang * 

College of Mathematics and Physics, Center for Applied Mathematics of Guangxi, Guangxi Minzu University, Nanning 530006, China; jianwei_math@gxmzu.edu.cn

* Correspondence: msyang22@gxmzu.edu.cn

Abstract: Independent component analysis (ICA), as a statistical and computational approach, has been successfully applied to digital signal processing. Performance analysis for the ICA approach is perceived as a challenging task to work on. This contribution concerns the complex-valued FastICA algorithm in the range of ICA over the complex number domain. The focus is on the robust and equivariant behavior analysis of the complex-valued FastICA estimator. Although the complex-valued FastICA algorithm as well as its derivatives have been widely used methods for approaching the complex blind signal separation problem, rigorous mathematical treatments of the robust measurement and equivariance for the complex-valued FastICA estimator are still missing. This paper strictly analyzes the robustness against outliers and separation performance depending on the global system. We begin with defining the influence function (IF) of complex-valued FastICA functional and followed by deriving its closed-form expression. Then, we prove that the complex-valued FastICA algorithm based on the optimizing cost function is linear-equivariant, depending only on the source signals.

Keywords: complex-valued independent component analysis; FastICA algorithm; robustness; equivariance

MSC: 62F35; 62B15; 62H12



Citation: E, J.; Yang, M. Complex-Valued FastICA Estimator with a Weighted Unitary Constraint: A Robust and Equivariant Estimator. *Mathematics* **2024**, *12*, 1840. <https://doi.org/10.3390/math12121840>

Academic Editor: Manuel Alberto M. Ferreira

Received: 28 May 2024

Revised: 11 June 2024

Accepted: 11 June 2024

Published: 13 June 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Independent component analysis (ICA) is a statistical and computational model, where the observed signals are considered as linear mixtures of underlying source signals. The purpose of ICA is to estimate mutually independent source signals from their linear mixtures without prior knowledge of sources and mixing coefficients [1,2]. Up to the present, ICA has been widely applied in diverse fields, such as signal processing [3,4], financial analysis [5,6], and so on. Hence, there is ever-increasing literature recognizing the essentiality of analyzing the properties of the employed ICA approaches from the aspect of theoretical support. The main purpose of this paper is to analyze the robustness and equivariance of the complex-valued fast fixed-point algorithm for ICA (complex-valued FastICA for short) [7,8], which is one of the most prominent algorithms in the complex number domain due to its faster convergence and easier implementation.

Among many publications, algorithmic property is one of the most fundamental properties for complex-valued ICA. Novey and Adali proposed complex ICA based on negentropy maximization (NM) [9] and provided the local stability condition of the algorithm. Furthermore, Qian and Wei gave a stability analysis from a unique perspective [10]. The authors pointed out that the NM-based ICA algorithm might iterate a poor separation vector even if the source signals meet the stability conditions. Reference [11] presented the Riemann and Lie structures of the complex unitary group, which generalized the topological properties of the complex-valued ICA. Koldovský and Tichavský [12] addressed

the problem regarding the region of convergence of gradient-related ICA algorithms. The results showed that the size of the region of convergence was related to the employed algorithm and relied on the ratio of scales of source signals. E et al. [13] provided a performance analysis of the complex-valued FastICA algorithm with the M-estimator cost function, including stability and local convergence. They proved the existence of local optimal solutions and stability conditions.

Moreover, the statistical property is another fundamental property for the complex-valued ICA. Cramér-Rao bound (CRB) is essential for the ICA algorithms to describe the performance limit, and has been researched, e.g., in [14,15]; the authors derived a closed-form expression for the CRB of the separation parameter for complex-valued ICA. Fu et al. [16] established the theory for complex-valued ICA, providing the Cramér-Rao lower bound and identification conditions, which exploited diversities of non-Gaussianity, non-whiteness, and non-circularity. Furthermore, Koldovský et al. [17] analyzed the accuracy of fast dynamic independent vector analysis, showing asymptotic efficiency under given mixing and statistical models, which coincided with the Cramér-Rao lower bound derived in [15]. Reference [13] also analyzed the statistical properties of the complex-valued FastICA method, including uniformity and robustness (the sequences of the complex-valued FastICA estimator converge in probability to the true demixing vector). However, unfortunately, they addressed the problem concerning the robust property without a rigorous mathematical treatment. In this paper, we focus on deriving a closed-form expression of the robust measurement for the complex-valued FastICA functional.

The circular complex-valued FastICA (c-FastICA) algorithm was introduced by Bingham and Hyvärinen [7], as an extension of FastICA in the real domain [18,19]. It has received wide attention for solving complex-valued source signal separation due to its faster convergence and easier implementation. In order to improve the robustness against outliers, Chao and Douglas [20] selected the Huber M-estimator as nonlinearity in the cost function within the complex-circular FastICA algorithm. The local stability analysis showed that the improved approach was locally stable for Huber's single-parameter M-estimator cost function with circular source signals. The aforementioned complex-valued FastICA algorithm, however, has poor performance when dealing with noncircular source signal separation. To address this obstacle, Novey and Adali extended it to the noncircular source signals separation scenario and analyzed the local convergence of the estimator [8]. Although there have been several attempts to study the statistical properties of the complex-noncircular FastICA algorithm (nc-FastICA for short), a rigorous analysis of the robust and equivariant behavior of nc-FastICA is still missing in the community. To fill this gap, this paper will analyze the robustness against outliers and the separation performance of the complex-valued FastICA estimator with a weighted unitary constraint.

The innovations of this paper are three-fold. First, we define the complex-valued FastICA functional and its influence function (IF) for the deflationary procedure. Then, a closed-form expression of the IF for the complex-valued FastICA functional is derived, which is utilized to measure robustness against outliers from a rigorous mathematical perspective. Third, we prove that the complex-valued FastICA algorithm is equivariant.

This paper is organized as follows: Section 2 provides the preliminaries of the complex-valued ICA model and the deflationary complex FastICA considered argument. Section 3 establishes the theories concerning the complex-valued FastICA functional and its influence function in the range of ICA over the complex number domain. Section 4 analyzes the equivariance of the complex FastICA estimator. Section 5 concludes the whole paper.

2. Preliminaries

2.1. Complex-Valued ICA Model

Assume that a complex-valued random vector $\mathbf{x} \in \mathbb{C}^n$ consists of two real-valued random vectors $\mathbf{x}^R, \mathbf{x}^I \in \mathbb{R}^n$ with imaginary unit j , $\mathbf{x} = \mathbf{x}^R + j\mathbf{x}^I$, and the conjugate of $\mathbf{x}$ can be defined by $\mathbf{x}^* = \mathbf{x}^R - j\mathbf{x}^I$. We denote by

$$\mathbf{m}_{\mathbf{x}} = E\{\mathbf{x}\},$$

$$C_{\mathbf{x}} = E\{(\mathbf{x} - \mathbf{m}_{\mathbf{x}})(\mathbf{x} - \mathbf{m}_{\mathbf{x}})^H\},$$

$$\tilde{C}_{\mathbf{x}} = E\{(\mathbf{x} - \mathbf{m}_{\mathbf{x}})(\mathbf{x} - \mathbf{m}_{\mathbf{x}})^T\},$$

the mean vector, covariance matrix, and complementary covariance matrix of $\mathbf{x}$, respectively, wherein $E\{\mathbf{x}\} = E\{\mathbf{x}^R\} + jE\{\mathbf{x}^I\}$, H denotes the conjugate transpose, and T represents the ordinary transpose. The complex-valued random vector $\mathbf{x}$ is circular if $\tilde{C}_{\mathbf{x}} = \mathbf{0}$, and otherwise noncircular.

Let $f(\mathbf{x}, \mathbf{x}^*)$ be an analytic function of a complex random vector $\mathbf{x}$ and its conjugate $\mathbf{x}^*$. In this paper, we utilize the Wirtinger calculus to obtain the partial derivatives with respect to $\mathbf{x}$ and $\mathbf{x}^*$, considering $\mathbf{x}$ and $\mathbf{x}^*$ as two variables independently [21].

The ICA model established in the complex number field can be considered an extension of a real-valued ICA model [2], which assumes that the observed signals are the linear combinations of the unknown source signals $\mathbf{s} = (s_1, s_2, \dots, s_n)^T$, resulting in the generation of the observed signals vector $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ according to

$$\mathbf{x} = \mathbf{A}\mathbf{s} = \sum_{i=1}^n \mathbf{a}_i s_i, \quad (1)$$

where $\mathbf{A} \in \mathbb{C}^{n \times n}$ is the unknown mixing matrix or coefficients matrix, which regulates the cumulative distribution function (cdf) of the observed signal $\mathbf{x}$ denoted by $F(\mathbf{x})$, and $\mathbf{a}_i$ is the i th column of matrix $\mathbf{A}$. Generally, the following assumptions are needed in the ICA model:

Assumption 1. The source signals are of zero mean and unit variance, i.e., $C_{\mathbf{s}} = \mathbf{I}$ and the sources $s_1, s_2, \dots, s_n$ are statistically mutually independent, and at most, one source is Gaussian.

Assumption 2. The mixing matrix $\mathbf{A}$ is of full rank.

The primary aim of solving the ICA problem is to procure a matrix $\mathbf{B}$ named the demixing matrix, which can be utilized to estimate the source signals $\mathbf{s}$ by $\mathbf{y} = (y_1, y_2, \dots, y_n)^T$ defined as follows:

$$\mathbf{y} = \mathbf{B}\mathbf{x} = \mathbf{B}\mathbf{A}\mathbf{s} = \mathbf{P}\mathbf{\Lambda}\mathbf{s}, \quad (2)$$

where $\mathbf{P}$ and $\mathbf{\Lambda}$ are the permutation and diagonal matrices, respectively, representing the ambiguities of the ICA model (1). This can be obtained by the non-Gaussian-maximization criterion [2] up to the scales, phase, and order of source signals.

The data-centering procedure is needed when solving the ICA problem, as it simplifies the theory and algorithm. Considering the zero-mean source signals $\mathbf{s}$, which can be achieved by subtracting the mean vector $\mathbf{m}_{\mathbf{x}}$ from $\mathbf{x}$, we can similarly define their covariance matrix $C_{\mathbf{s}}$ and complementary covariance matrix $\tilde{C}_{\mathbf{s}}$. Especially, the source signal is circular if $\tilde{C}_{\mathbf{s}} = \mathbf{0}$, otherwise, it is noncircular. In this study, we analyze the robustness and equivariant characteristics of symmetric complex-valued FastICA under the more general case of noncircular sources.

2.2. Deflationary Complex FastICA

To date, various approaches have been developed for solving the complex ICA problem, with the complex-valued FastICA algorithm being one of the most prominent for finding the demixing matrix $\mathbf{B}$.

Let $\mathbf{b}$ be one of the columns of the demixing matrix $\mathbf{B}$, regarding a demixing vector. On the basis of Assumption 1, the estimated source signal $y = \mathbf{b}^H \mathbf{x}$ should be of unit variance by searching for the solution on the weighted unitary constraint group $O(n) = \{\mathbf{b} \in \mathbb{C}^n : \|\mathbf{b}\|_{C_{\mathbf{x}}} = 1\}$ with a weighted vector norm, such that $\|\mathbf{b}\|_{\tilde{C}_{\mathbf{x}}}^2 = \mathbf{b}^H \mathbf{C}_{\mathbf{x}} \mathbf{b}$. Hence, the cost function of the complex-valued FastICA algorithm with the weighted unitary constraint for the one-unit version has the following form:

$$J(\mathbf{b}) = E\{G(|\mathbf{b}^H \mathbf{x}|^2)\}, \quad \mathbf{b} \in O(n), \quad (3)$$

where $G : \mathbb{R}^+ \cup \{0\} \rightarrow \mathbb{R}$ represents a smooth even function defined as the nonlinearity function (which determines the robustness of the complex-valued FastICA algorithm, which will be further involved in the forthcoming Section 3); $\mathbf{x}$ denotes the observed signals vector defined in (1).

We note that (i) the cost function performs on the absolute value $|\mathbf{b}^H \mathbf{x}|^2$ instead of a complex-valued vector, which results from seeking to maximize the expectation of the non-linear cost function; (ii) the optimization of the weighted unitary constraint group $O(n)$ guarantees that the estimated source signal $y = \mathbf{b}^H \mathbf{x}$ has unit variance.

In order to solve the optimization problem (3), Bingham and Hyvärinen [7] presented a fast fixed-point type algorithm (called c-FastICA), which is capable of estimating the circular source signals. However, it has poor performance when dealing with noncircular source separation. To figure out this obstacle, Novey and Adali extended the c-FastICA algorithm to the noncircular source signal separation scenario (called the nc-FastICA algorithm), consisting of the following learning rule [8]:

Step 1. Choose an arbitrary initial value of the unit norm for $\mathbf{b} \in O(n)$;

Step 2. Run iterations:

$$\begin{aligned} \mathbf{b} &\leftarrow \frac{E\{g(|\mathbf{b}^H \mathbf{x}|^2)(\mathbf{b}^H \mathbf{x})^* \mathbf{x}\} + E\{g'(|\mathbf{b}^H \mathbf{x}|^2)|\mathbf{b}^H \mathbf{x}|^2 \\ &\quad + g(|\mathbf{b}^H \mathbf{x}|^2)\}\mathbf{b} + E\{\mathbf{x}\mathbf{x}^T\}E\{g'(|\mathbf{b}^H \mathbf{x}|^2)(\mathbf{b}^H \mathbf{x})^* \mathbf{x}\}^*}{\|\mathbf{b}\|_{\mathbb{C}_x}^2}, \end{aligned} \quad (4)$$

until convergence, where the notations $g(z)$, $g'(z)$ denote $\frac{dG(z)}{dz}$, and $\frac{dg(z)}{dz}$, respectively.

One needs to estimate several source signals; the one-unit complex FastICA algorithm (4) should be run several times with vectors $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n$. To prevent the vectors from converging to the same maxima, they need to be orthogonalized after each iteration using the Gram–Schmidt method [2], resulting in the following deflationary complex FastICA algorithm:

Step 1. Choose an arbitrary initial value of unit norm for $\mathbf{b}_j \in O(n)$; set counter $j = 1$;

Step 2. Run iterations:

$$\begin{aligned} \mathbf{b}_j &\leftarrow \frac{E\{g(|\mathbf{b}_j^H \mathbf{x}|^2)(\mathbf{b}_j^H \mathbf{x})^* \mathbf{x}\} + E\{g'(|\mathbf{b}_j^H \mathbf{x}|^2)|\mathbf{b}_j^H \mathbf{x}|^2 \\ &\quad + g(|\mathbf{b}_j^H \mathbf{x}|^2)\}\mathbf{b}_j + E\{\mathbf{x}\mathbf{x}^T\}E\{g'(|\mathbf{b}_j^H \mathbf{x}|^2)(\mathbf{b}_j^H \mathbf{x})^* \mathbf{x}\}^*}{\|\mathbf{b}_j\|_{\mathbb{C}_x}^2}, \\ \mathbf{b}_j &\leftarrow \mathbf{b}_j - \sum_{i=1}^{j-1} (\mathbf{b}_j^T \mathbf{b}_i) \mathbf{b}_i \\ \mathbf{b}_j &\leftarrow \frac{\mathbf{b}_j}{\|\mathbf{b}_j\|_{\mathbb{C}_x}} \end{aligned} \quad (5)$$

until convergence;

Step 3. Set $j \leftarrow j + 1$ until $j = n$.

3. Robustness of the Complex-Valued FastICA Estimator

3.1. Nonlinearity

From a statistical viewpoint, the nonlinear function G provides information on the higher-order statistics in the expectation format $E\{G(|\mathbf{b}^H \mathbf{x}|^2)\}$, which determines the selection of g (the derivation of the nonlinear function G) in the algorithm (4). Hence, the statistical properties of the complex-valued FastICA estimator $\mathbf{b}_k$ lie essentially in the selection of nonlinear function G . In this paper, we analyze the robustness of the $\mathbf{b}_k$ via the IF.

Generally, robustness against outliers is a desirable property for an estimator, meaning that the estimator is insensitive to individual, highly erroneous observations. In this section, we mainly address the problem of how to measure the robustness of the complex-valued

FastICA estimator $\mathbf{b}_k$? Heuristically, the value of the function $G(x)$ cannot grow fast with the increase in $|x|$ if one needs a robust estimator. Specifically, we list the classical nonlinearities $G(x)$ in Table 1. The curves of the function $G(x)$ and its derivation with respect to x are plotted in Figures 1 and 2.

Table 1. The nonlinearities in the complex-valued FastICA algorithm.

Nonlinear Function $G(x)$	Derivation of $G(x)$
$G_{\log} = \log(1 + x^2)$	$g_{\log} = \frac{2x}{1+x^2}$
$G_{\tanh} = \log \cosh(x)$	$g_{\tanh} = \tanh(x)$
$G_{kurt} = \frac{1}{2}x^4$	$g_{kurt} = 2x^3$
$G_{sqr} = \sqrt{1 + x^2}$	$g_{sqr} = \frac{x}{\sqrt{1+x^2}}$
$G_{Huber} = \begin{cases} \frac{1}{2} x ^2, & x \leq \theta, \\ \theta x - \frac{\theta^2}{2}, & x > \theta. \end{cases}$	$g_{Huber} = \begin{cases} x, & x \leq \theta, \\ \theta \operatorname{sgn}(x), & x > \theta. \end{cases}$
$G_{Tukey} = \begin{cases} \frac{\theta^2}{6} [1 - (1 - \frac{x^2}{\theta^2})^3], & x \leq \theta, \\ \frac{\theta^2}{6}, & x > \theta. \end{cases}$	$g_{Tukey} = \begin{cases} x(1 - \frac{x^2}{\theta^2})^2, & x \leq \theta, \\ 0, & x > \theta. \end{cases}$

Note: The functions with subscripts denote the logarithm, hyperbolic tangent, kurtosis, mean square root, Huber M-estimator, and Tukey M-estimator functions, respectively.

From these Figures, one can yield that the Tukey M-estimator function implements a more robust estimator that is insensitive to outliers, and kurtosis gives a non-robust estimator that may be influenced by individual highly erroneous observations. In fact, the values of $G_{kurt}(x)$ (red line in Figure 1) are increased quickly from the beginning of 1 or -1 without a downward trajectory. With the increase in $|x|$, the outliers do not have much influence on the values of $G_{Tukey}(x)$ in the sense that the complex-valued FastICA estimator based on the Tukey M-estimator cost function is recommended with better robustness. This evidence can also be seen in the curves of the derivation of nonlinearities $G(x)$ in Figure 2. In this paper, we shall, hereafter, analyze the robustness of the complex-valued FastICA estimator via the IF.

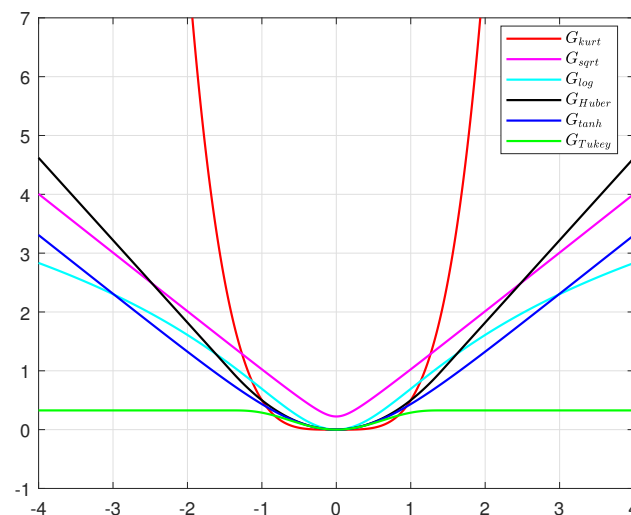


Figure 1. Plot of nonlinearities $G(y)$.

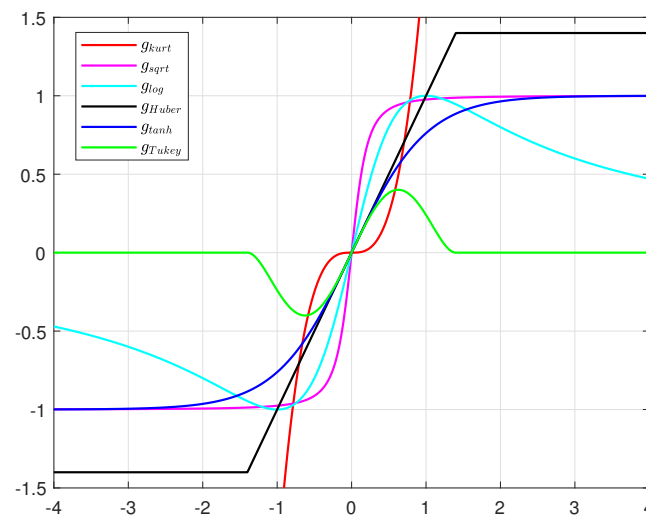


Figure 2. Plot of nonlinearities $g(y)$.

3.2. Influence Function of Complex-Valued FastICA Functional

We analyze the robustness of the complex-valued FastICA estimator by concerning the deflationary version of the complex-valued FastICA algorithm. For convenience, we shall hereafter suppose that $\mathbf{b}_k$ is the k th column of the demixing matrix $\mathbf{B}$, which corresponds to the estimator for finding the k th source signal s_k in the sense that the equation $y_k = \mathbf{b}_k^H \mathbf{x}$ gives an estimation of the source signal s_k .

In the complex-valued ICA model (1), the observed signal vector $\mathbf{x}$ comes from an unknown distribution with the cdf $F(\mathbf{x})$. Hence, in order to give the measurement of the complex-valued FastICA estimator, we first define the complex-valued FastICA functional $\mathbf{b}_k(F)$ for the deflationary procedure and its influence function.

Definition 1. Assume that the observed vector $\mathbf{x}$ follows the complex-valued ICA model (1). The complex-valued FastICA functional $\mathbf{b}_k(F)$ for the deflationary procedure is defined as follows:

$$\mathbf{b}_k(F) = \arg \max_{\|\mathbf{b}_k\|_{C_{\mathbf{x}}} = 1} E\{G(|\mathbf{b}_k^H \mathbf{x}|^2)\}, \quad (6)$$

subject to the following weighted unitary constraint:

$$\mathbf{b}_k(F)^H C_{\mathbf{x}}(F) \mathbf{b}_j(F) = \begin{cases} 0, & j < k, \\ 1, & j = k. \end{cases} \quad (7)$$

where $G : \mathbb{R}^+ \cup \{0\} \rightarrow \mathbb{R}$ represents a smooth even function and $C_{\mathbf{x}}(F)$ denotes the covariance matrix function of the observed vector $\mathbf{x}$ at the distribution $F(\mathbf{x})$.

We note that the constraint condition (7) can be performed by the deflationary orthogonalization process, which separates the source signals one by one.

Definition 2. The influence function (IF) of the complex-valued FastICA functional $\mathbf{b}_k(\cdot)$ at $F(\mathbf{x})$ is given by the following:

$$IF(\mathbf{x}, \mathbf{b}, F(\mathbf{x})) = \lim_{t \rightarrow 0} \frac{\mathbf{b}_k((1-t)F(\mathbf{x}) + t\Delta_{\mathbf{x}}) - \mathbf{b}_k(F(\mathbf{x}))}{t} \quad (8)$$

where $\Delta_{\mathbf{x}}$ denotes the probability measure, which puts mass 1 at point $\mathbf{x} \in \mathbb{C}^n$.

We note that the IF of the complex-valued FastICA functional $\mathbf{b}_k(\cdot)$ is defined based on the fact that it is Gâteaux differentiable [22] at the distribution $F(\mathbf{x})$ in $\mathbb{C}^n$. Thus, it can also be written as follows:

$$IF(\mathbf{x}, \mathbf{b}, F(\mathbf{x})) = \left. \frac{\partial \mathbf{b}_k(F_t(\mathbf{x}))}{\partial t} \right|_{t=0}, \quad (9)$$

where $F_t(\mathbf{x}) = (1 - t)F(\mathbf{x}) + t\Delta_{\mathbf{x}}$.

To simplify the notation, we shall hereafter replace the notations $IF(\mathbf{x}, \mathbf{b}, F(\mathbf{x}))$ with $IF(\mathbf{b})$ to denote the IF of the complex-valued FastICA functional $\mathbf{b}_k(\cdot)$, $\mathbf{C}_{\mathbf{x}}(F)$ (Res. $\mathbf{m}_{\mathbf{x}}(F)$) by the $\mathbf{C}_{\mathbf{x}}$ (Res. $\mathbf{m}_{\mathbf{x}}$) to denote the covariance matrix function (Res. mean vector function) of the observed vector $\mathbf{x}$ at the distribution $F(\mathbf{x})$, F_t by $F_t(\mathbf{x})$, respectively.

The significance of the IF lies in its heuristic explanation: it reports the influence of the contamination at point $\mathbf{x}$ on the estimate. From the robust statistics point of view, the IF quantifies the asymptotic bias resulting from contamination in the observation [22]. Hence, the forthcoming section provides a closed-form expression of the IF of the complex-valued FastICA functional $\mathbf{b}_k(\cdot)$.

3.3. Robustness

In order to analyze the robustness, we use the Lagrangian multiplier method to deduce the specific expression of the influence function of the complex-valued FastICA functional. Plugging the weighted unitary constraint (7) into the cost function (3), the Lagrange function can be obtained as follows:

$$L(\mathbf{b}_k, \lambda) = J(\mathbf{b}_k) - \lambda_k [\mathbf{b}_k^H \mathbf{C}_{\mathbf{x}} \mathbf{b}_k - 1] - \sum_{j=1}^{k-1} \lambda_j \mathbf{b}_j^H \mathbf{C}_{\mathbf{x}} \mathbf{b}_k, \quad (10)$$

where $\lambda_j \in \mathbb{R}$, $j = 1, \dots, k$ is the penalty factor; $J(\mathbf{b}_k)$ is the cost function of the complex-valued FastICA. Note that the observed data, $\mathbf{x}$, are concerned without a preprocessing procedure, in the sense that $J(\mathbf{b}_k) = E\{G(|y_{\mathbf{x},k}|^2)\} = E\{G(y_{\mathbf{x},k}y_{\mathbf{x},k}^*)\}$, wherein $y_{\mathbf{x},k} = \mathbf{b}_k^H(\mathbf{x} - \mathbf{m}_{\mathbf{x}}) = \sum_{i=1}^n b_i^*(x_i - m_i)$, and b_i, x_i, m_i denote the i th elements for counterparts, respectively. After differentiating the $L(\mathbf{b}_k, \lambda)$ with respect to $\mathbf{b}_k = (b_1, \dots, b_n)^T$ and equating to zero, one can yield

$$\nabla_{\mathbf{b}_k} J(\mathbf{b}_k) = \nabla_{\mathbf{b}_k} \lambda_k [\mathbf{b}_k^H \mathbf{C}_{\mathbf{x}} \mathbf{b}_k - 1] + \nabla_{\mathbf{b}_k} \sum_{j=1}^{k-1} \lambda_j \mathbf{b}_j^H \mathbf{C}_{\mathbf{x}} \mathbf{b}_k, \quad (11)$$

where $\nabla_{\mathbf{b}}$ denotes the gradient operator with respect to $\mathbf{b}$, yielding the following:

$$\begin{aligned} \frac{\partial E\{G(y_{\mathbf{x},k}y_{\mathbf{x},k}^*)\}}{\partial b_i} &= E\{g(y_{\mathbf{x},k}y_{\mathbf{x},k}^*)y_{\mathbf{x},k}[x_i - m_i]^*\}, \\ \frac{\partial \lambda_k \mathbf{b}_k^H \mathbf{C}_{\mathbf{x}} \mathbf{b}_k}{\partial b_i} &= \lambda_k \mathbf{b}_k^H \mathbf{c}_i, \\ \frac{\partial \lambda_j \mathbf{b}_j^H \mathbf{C}_{\mathbf{x}} \mathbf{b}_k}{\partial b_i} &= \lambda_j \mathbf{b}_j^H \mathbf{c}_i, \end{aligned}$$

where $\mathbf{c}_i$ represents the i th column of the covariance matrix function $\mathbf{C}_{\mathbf{x}}(F)$. Furthermore,

$$E\{g(|y_{\mathbf{x},k}|^2)y_{\mathbf{x},k}[\mathbf{x} - \mathbf{m}_{\mathbf{x}}]^*\} = \lambda_k \mathbf{C}_{\mathbf{x}}^T \mathbf{b}_k^* + \sum_{j=1}^{k-1} \lambda_j \mathbf{C}_{\mathbf{x}}^T \mathbf{b}_j^*. \quad (12)$$

For the sake of convenience in writing, we shall identify the following:

$$L_k(F) \triangleq E\{g(|y_{\mathbf{x},k}|^2)y_{\mathbf{x},k}[\mathbf{x} - \mathbf{m}_{\mathbf{x}}]^*\} \quad (13)$$

and

$$R_k(F) \triangleq \lambda_k \mathbf{C}_{\mathbf{x}}^T \mathbf{b}_k^* + \sum_{j=1}^{k-1} \lambda_j \mathbf{C}_{\mathbf{x}}^T \mathbf{b}_j^* \quad (14)$$

respectively, yielding the following:

$$L_k(F) = R_k(F). \quad (15)$$

Theorem 1. The influence function of the complex-valued FastICA functional $\mathbf{b}_k(\cdot)$, $k \in \{1, 2, \dots, n\}$, at the centered mixture distribution $F(\mathbf{x})$ is as follows:

$$IF(\mathbf{b}_k) = -\bar{s}_k^* \sum_{i=1}^{k-1} [u_i(\bar{s}_i)^* + \bar{s}_i] \mathbf{b}_i + \frac{1 - |\bar{s}_k|^2}{2} \mathbf{b}_k + u_k(\bar{s}_k) \sum_{i=k+1}^n \bar{s}_i \mathbf{b}_i \quad (16)$$

where $u_k(x) = \frac{(g(|x|^2) - c_k)x^* - \rho_k^*}{c_k - \mu_k}$, $c_k = E\{g(|s_k|^2)|s_k|^2\}$, $\mu_k = E\{g(|s_k|^2)\}$, $\rho_k = E\{g(|s_k|^2)s_k^2\}$, $g(\cdot)$ is the derivative operator of the nonlinearity $G(\cdot)$ in the cost function (3), and $\bar{s}_i = \mathbf{b}_i^H(\mathbf{z} - \mathbf{m}_x)$ denotes the projection of the centered contamination point $\mathbf{z}$ into the direction $\mathbf{b}_i$.

To prove Theorem 1, the following five Lemmas are needed, which will be proved in Appendixes A and B.

Lemma 1. Assume that the observed data, $\mathbf{x}$, obey the distribution $F(\mathbf{x})$ with the mean vector $\mathbf{m}_x$, and the sources $\mathbf{s}$ follow Assumption 1. Let $y_{x,i}(F_t) = \mathbf{b}_i^H(F_t)(\mathbf{x} - \mathbf{m}_x)$, $L_k(F) = E\{g(|y_{x,k}|^2)y_{x,k}[\mathbf{x} - \mathbf{m}_x]^*\}$ and $B_k(F) = E\{g(|y_{x,k}|^2)[\mathbf{x} - \mathbf{m}_x][\mathbf{x} - \mathbf{m}_x]^H\}$, one obtains the following:

- (a) $IF(y_{x,i}) = IF(\mathbf{b}_i)^H(\mathbf{x} - \mathbf{m}_x) - \bar{s}_i$,
- (b) $L_k(F) = c_k \mathbf{a}_k^*$,
- (c) $B_k(F) = \mu_k \mathbf{A} \mathbf{A}^H + (c_k - \mu_k) \mathbf{a}_k \mathbf{a}_k^H$,
- (d) $\mathbf{a}_k^H IF(\mathbf{b}_i) = -IF(\mathbf{b}_k)^H \mathbf{a}_i - \bar{s}_k \bar{s}_i^*$, $i = 1, \dots, k-1$,
- (e) $2\text{Re}(\mathbf{a}_k^H IF(\mathbf{b}_k)) = 1 - |\bar{s}_k|^2$,

where $\bar{s}_i = \mathbf{b}_i^H(\mathbf{z} - \mathbf{m}_x)$, $i = 1, \dots, n$ denotes the projection of the centered contamination point $\mathbf{z}$ into the direction $\mathbf{b}_i$; $c_k = E\{g(|s_k|^2)|s_k|^2\}$, $\mu_k = E\{g(|s_k|^2)\}$. The notation of $\mathbf{a}_k$ represents the k th column of matrix $\mathbf{A}$.

Proof. See Appendix A. $\square$

Lemma 2. Assume that the observed data, $\mathbf{x}$, obey the distribution $F(\mathbf{x})$ with the mean vector $\mathbf{m}_x$, and the sources $\mathbf{s}$ follow Assumption 1. For λ_i , $i = 1, \dots, k$ in (10) at the distribution $F(\mathbf{x})$, we have the following:

- (a) $\lambda_i(F) = 0$, $i = 1, \dots, k-1$,
- (b) $\lambda_k(F) = c_k$,
- (c) $IF(\lambda_i) = (c_k - \mu_k) \mathbf{a}_k^H IF(\mathbf{b}_i) + (g(|\bar{s}_k|^2) - \mu_k) \bar{s}_k \bar{s}_i^* - \rho_k s_i^*$, $i = 1, \dots, k-1$,
- (d) $IF(\lambda_k) = (g(|\bar{s}_k|^2) - c_k) |\bar{s}_k|^2 + E\{g'(|\bar{s}_k|^2) v_k(\bar{s}_k) |\bar{s}_k|^2\} - 2\text{Re}(E\{g(|\bar{s}_k|^2) \bar{s}_k s_k^*\})$,

where $v_k(\bar{s}_k) = |s_k|^2 - |s_k|^2 |\bar{s}_k|^2 - 2\text{Re}(\bar{s}_k s_k^*)$, and c_k , μ_k , $\bar{s}_i$, ρ_k are defined as in Theorem 1.

Proof. See Appendix B. $\square$

Based on Lemmas 1 and 2, we will now prove Theorem 1 from the following three steps:

Proof. Step 1. Differentiating $L_k(F_t)$ and $R_k(F_t)$ with respect to $t = 0$ in (15), yields the following equation:

$$\left. \frac{\partial L_k(F_t)}{\partial t} \right|_{t=0} = \left. \frac{\partial R_k(F_t)}{\partial t} \right|_{t=0} \quad (17)$$

which aims at obtaining a closed-form expression of the influence function of the complex-valued FastICA functional $IF(\mathbf{b}_k)$.

Step 2. Calculating the left-hand side of the Equation (17). Combing $F_t(\mathbf{x}) = (1 - t)F(\mathbf{x}) + t\Delta_x$ with Equation (13), we conclude the following:

$$L_k(F_t) = (1-t)E\{g(|y_{\mathbf{x},k}|^2)y_{\mathbf{x},k}[\mathbf{x} - \mathbf{m}_{\mathbf{x}}]^*\} + tg(|y_{\mathbf{z},k}|^2)y_{\mathbf{z},k}[\mathbf{z} - \mathbf{m}_{\mathbf{x}}]^*, \quad (18)$$

Differentiating $L_k(F_t)$ with respect to $t = 0$ and leveraging Lemma 1, one can obtain the following:

$$\begin{aligned} \left. \frac{\partial L_k(F_t)}{\partial t} \right|_{t=0} &= -L_k(F) + (1-t) \left[E\{g'(|y_{\mathbf{x},k}|^2)(IF(y_{\mathbf{x},k})y_{\mathbf{x},k}^* + y_{\mathbf{x},k}IF(y_{\mathbf{x},k})^*)y_{\mathbf{x},k}(\mathbf{x} - \mathbf{m}_{\mathbf{x}})^*\} \right. \\ &\quad \left. + E\{g(|y_{\mathbf{x},k}|^2)IF(y_{\mathbf{x},k})(\mathbf{x} - \mathbf{m}_{\mathbf{x}})^*\} - E\{g(|y_{\mathbf{x},k}|^2)y_{\mathbf{x},k}IF(\mathbf{m}_{\mathbf{x}})^*\} \right] \Big|_{t=0} \\ &\quad + g(|y_{\mathbf{z},k}|^2)y_{\mathbf{z},k}(\mathbf{z} - \mathbf{m}_{\mathbf{x}})^* \Big|_{t=0} \\ &= -c_k \mathbf{a}_k^* + E\{g'(|s_k|^2)|s_k|^2 v_k(\bar{s}_k)\} \mathbf{a}_k^* + [\mu_k \mathbf{A} \mathbf{A}^H + (c_k - \mu_k) \mathbf{a}_k \mathbf{a}_k^H]^* IF(\mathbf{b}_k)^* \\ &\quad - 2\text{Re}(E\{g(|s_k|^2)\bar{s}_k s_k^*\}) \mathbf{a}_k^* + \rho_k \bar{s}_k^* \mathbf{a}_k^* + [g(|\bar{s}_k|^2)\bar{s}_k - \rho_k](\mathbf{z} - \mathbf{m}_{\mathbf{x}})^*. \end{aligned}$$

Step 3. Calculating the right-hand side of (17). From (14), we have the following:

$$R_k(F_t) = \lambda_k(F_t) \mathbf{C}_{\mathbf{x}}^T \mathbf{b}_k^*(F_t) + \sum_{j=1}^{k-1} \lambda_j(F_t) \mathbf{C}_{\mathbf{x}}^T \mathbf{b}_j^*(F_t), \quad (19)$$

Differentiating $R_k(F_t)$ with respect to $t = 0$ and leveraging Lemma 2 one can obtain the following:

$$\begin{aligned} \left. \frac{\partial R_k(F_t)}{\partial t} \right|_{t=0} &= IF(\lambda_k)(\mathbf{A} \mathbf{A}^H)^T \mathbf{b}_k^* + \lambda_k(F) IF(\mathbf{C}_{\mathbf{x}})^T \mathbf{b}_k^* + \lambda_k(F)(\mathbf{A} \mathbf{A}^H)^T IF(\mathbf{b}_k)^* \\ &\quad + \sum_{j=1}^{k-1} [IF(\lambda_j)(\mathbf{A} \mathbf{A}^H)^T \mathbf{b}_j^* + \lambda_j(F) IF(\mathbf{C}_{\mathbf{x}})^T \mathbf{b}_j^* + \lambda_j(F)(\mathbf{A} \mathbf{A}^H)^T IF(\mathbf{b}_j)^*] \\ &= IF(\lambda_k) \mathbf{a}_k^* + c_k[(\mathbf{z} - \mathbf{m}_{\mathbf{x}})^* \bar{s}_k - \mathbf{a}_k^*] + c_k(\mathbf{A} \mathbf{A}^H)^T IF(\mathbf{b}_k)^* + \sum_{j=1}^{k-1} IF(\lambda_j) \mathbf{a}_j^* \\ &= [g(|\bar{s}_k|^2)|\bar{s}_k|^2 - c_k |\bar{s}_k|^2] \mathbf{a}_k^* - c_k \mathbf{a}_k^* + c_k(\mathbf{z} - \mathbf{m}_{\mathbf{x}})^* \bar{s}_k + c_k(\mathbf{A} \mathbf{A}^H)^T IF(\mathbf{b}_k)^* \\ &\quad + E\{g'(|s_k|^2)|s_k|^2 v_k(\bar{s}_k)\} \mathbf{a}_k^* - 2\text{Re}(E\{g(|s_k|^2)\bar{s}_k s_k^*\}) \mathbf{a}_k^* - \rho_k \sum_{j=1}^{k-1} \bar{s}_j^* \mathbf{a}_j^* \\ &\quad + (c_k - \mu_k) \sum_{j=1}^{k-1} \mathbf{a}_k^H IF(\mathbf{b}_j) \mathbf{a}_j^* + [g(|\bar{s}_k|^2) - \mu_k] \bar{s}_k \sum_{j=1}^{k-1} \bar{s}_j^* \mathbf{a}_j^*. \end{aligned}$$

Step 4. Deriving the specific expression of the IF of the complex-valued FastICA function. Plugging the results derived in step 1 and step 2 into (17), after tedious manipulation, one can yield the following:

$$\begin{aligned} IF(\mathbf{b}_k) &= \frac{g(|\bar{s}_k|^2)\bar{s}_k^* - c_k \bar{s}_k^* - \rho_k^*}{c_k - \mu_k} \mathbf{B}^H \mathbf{B}(\mathbf{z} - \mathbf{m}_{\mathbf{x}}) - \frac{g(|\bar{s}_k|^2) - c_k}{c_k - \mu_k} |\bar{s}_k|^2 \mathbf{B}^H \mathbf{B} \mathbf{a}_k + \frac{\rho_k^* \bar{s}_k}{c_k - \mu_k} \mathbf{B}^H \mathbf{B} \mathbf{a}_k \\ &\quad + \frac{\rho_k^*}{c_k - \mu_k} \sum_{j=1}^{k-1} \bar{s}_j \mathbf{B}^H \mathbf{B} \mathbf{a}_j - \sum_{j=1}^{k-1} \mathbf{B}^H \mathbf{B} \mathbf{a}_k^T IF(\mathbf{b}_j)^* \mathbf{a}_j \\ &\quad - \frac{g(|\bar{s}_k|^2) - \mu_k}{c_k - \mu_k} \bar{s}_k^* \sum_{j=1}^{k-1} \bar{s}_j \mathbf{B}^H \mathbf{B} \mathbf{a}_j + \frac{1 - |\bar{s}_k|^2}{2} \mathbf{B}^H \mathbf{B} \mathbf{a}_k \\ &= \frac{g(|\bar{s}_k|^2)\bar{s}_k^* - c_k \bar{s}_k^* - \rho_k^*}{c_k - \mu_k} \left(\sum_{j \neq k}^n \mathbf{b}_j \bar{s}_j + \bar{s}_k \mathbf{b}_k \right) - \frac{g(|\bar{s}_k|^2) - c_k}{c_k - \mu_k} |\bar{s}_k|^2 \mathbf{b}_k + \frac{\rho_k^* \bar{s}_k}{c_k - \mu_k} \mathbf{b}_k \\ &\quad + \frac{\rho_k^*}{c_k - \mu_k} \sum_{j=1}^{k-1} \bar{s}_j \mathbf{b}_j - \sum_{j=1}^{k-1} \mathbf{a}_k^T IF(\mathbf{b}_j)^* \mathbf{b}_j - \frac{g(|\bar{s}_k|^2) - \mu_k}{c_k - \mu_k} \bar{s}_k^* \sum_{j=1}^{k-1} \bar{s}_j \mathbf{b}_j + \frac{1 - |\bar{s}_k|^2}{2} \mathbf{b}_k \\ &= -s_k^* \sum_{j=1}^{k-1} \bar{s}_j \mathbf{b}_j + u_k(\bar{s}_k) \sum_{j=k+1}^n \bar{s}_j \mathbf{b}_j - \sum_{j=1}^{k-1} \mathbf{a}_k^T IF(\mathbf{b}_j)^* \mathbf{b}_j + \frac{1 - |\bar{s}_k|^2}{2} \mathbf{b}_k. \end{aligned} \quad (20)$$

Combining the preceding equality concludes our proof $\mathbf{a}_k^H IF(\mathbf{b}_j) = u_j(\bar{s}_j)\bar{s}_k$. $\square$

Remark 1. In the proof of Theorem 1 and Lemmas 1–2, we need to consider the following facts:

$$\begin{aligned}\mathbf{A}^H \mathbf{b}_k &= \mathbf{e}_k, \\ \mathbf{A} \mathbf{e}_k &= \mathbf{a}_k, \\ \mathbf{b}_i^H \mathbf{a}_j &= \begin{cases} 0, & j < k \\ 1, & j = k \end{cases}\end{aligned}$$

where $\mathbf{A}$ denotes the mixing matrix in ICA model (1); $\mathbf{a}_k$ is the k th column of matrix $\mathbf{A}$; $\mathbf{e}_k$ represents a vector with 1 in the k th element and 0 elsewhere; and $\mathbf{b}_i$ is the i th column of the demixing matrix $\mathbf{B}$.

Remark 2. Note that pre-multiplying both sides of Equation (20) by $\mathbf{a}_i^H$, we have the following:

$$\mathbf{a}_i^H IF(\mathbf{b}_k) = -\bar{s}_k^* \sum_{j=1}^{k-1} \bar{s}_j \mathbf{a}_i^H \mathbf{b}_j + u_k(\bar{s}_k) \sum_{j=k+1}^n \bar{s}_j \mathbf{a}_i^H \mathbf{b}_j - \sum_{j=1}^{k-1} \mathbf{a}_k^T IF(\mathbf{b}_j)^* \mathbf{a}_i^H \mathbf{b}_j + \frac{1-|\bar{s}_k|^2}{2} \mathbf{a}_i^H \mathbf{b}_k, \quad (21)$$

After analyzing (21), one can observe the following: $\mathbf{a}_i^H IF(\mathbf{b}_k) = u_k(\bar{s}_k)\bar{s}_i$ for $i > k$, or $\mathbf{a}_k^H IF(\mathbf{b}_j) = u_j(\bar{s}_j)\bar{s}_k$ for $k < j$.

Remark 3. The expression of complex-valued FastICA IF can also be written as follows:

$$IF(\mathbf{b}_k) = \begin{cases} \frac{1-|\bar{s}_1|^2}{2} \mathbf{b}_1 + u_1(\bar{s}_1) \sum_{i=2}^n \bar{s}_i \mathbf{b}_i, & k = 1, \\ -\bar{s}_k^* \sum_{i=1}^{k-1} [u_i(\bar{s}_i)^* + \bar{s}_i] \mathbf{b}_i + \frac{1-|\bar{s}_k|^2}{2} \mathbf{b}_k + u_k(\bar{s}_k) \sum_{i=k+1}^n \bar{s}_i \mathbf{b}_i, & 1 < k < n, \\ -\bar{s}_n^* \sum_{i=1}^{n-1} [u_i(\bar{s}_i)^* + \bar{s}_i] \mathbf{b}_i + \frac{1-|\bar{s}_n|^2}{2} \mathbf{b}_n, & k = n. \end{cases} \quad (22)$$

By observing the closed-form expression of the IF for the complex-valued FastICA functional $\mathbf{b}_k$ in Theorem 1, we obtain that the IF of $\mathbf{b}_k$ is the weighted sum of the separation vector $\mathbf{b}_1, \dots, \mathbf{b}_n$ with the unbounded weight coefficient function, with respect to the projection of the contaminated point $\bar{s}_1, \dots, \bar{s}_n$. This finding confirms that the values of $IF(\mathbf{b}_k)$ are large when the outliers are present in the source signals. As can also be seen in Theorem 1, the greater the values of the contaminated point $\bar{s}_1, \dots, \bar{s}_n$ presenting in the source signals, the higher the values of $IF(\mathbf{b}_k)$.

4. Equivariance of the Complex-Valued FastICA Estimator

4.1. Equivariance of Complex-Valued ICA

Equivariance is a fundamental property in statistics [23], which makes the transformation of the parameter in the model equal to that of the sample data, which results in the following definition.

Definition 3. Assume that $\hat{\mathbf{A}} = \mathcal{A}(\mathbf{X})$ is an estimator of the mixing matrix $\mathbf{A}$ in the complex-valued ICA model (1), which relies on the sample matrix $\mathbf{X} = [\mathbf{x}(1), \dots, \mathbf{x}(T)]$ with T sample points. Then, we call the estimator $\hat{\mathbf{A}}$ equivariant if

$$\mathcal{A}(\mathbf{T}\mathbf{X}) = \mathbf{T}\mathcal{A}(\mathbf{X}) \quad (23)$$

for any invertible transformation $\mathbf{T}$.

In order to analyze the equivariance of the complex-valued FastICA estimator, the following Theorem is needed.

Theorem 2. Let $\mathbf{C} = \mathbf{B}\mathbf{A}$ be the global mixing–demixing system of the complex-valued ICA model (1) and $\hat{\mathbf{A}}$ be equivariant. Then the global mixing–demixing system $\mathbf{C}$ of ICA algorithms based

on the optimizing cost function depends only on the source signal matrix $\mathbf{S} = [\mathbf{s}(1), \dots, \mathbf{s}(T)]$, instead of the mixing system $\mathbf{A}$ and demixing system $\mathbf{B}$.

Proof. Denote by $\mathbf{C}$ the mixing–demixing matrix $\mathbf{BA}$; the estimation of the source can be rewritten as follows:

$$\mathbf{y} = \mathbf{C}\mathbf{s}. \quad (24)$$

In fact, the main property of the equivariant estimator for the BSS problem is that it provides uniform performance in the sense that

$$\mathbf{y} = \mathbf{B}\mathbf{x} = \mathcal{A}(\mathbf{X})^{-1}\mathbf{A}\mathbf{s} = \mathcal{A}(\mathbf{AS})^{-1}\mathbf{A}\mathbf{s} = \mathcal{A}(\mathbf{S})^{-1}\mathbf{s}, \quad (25)$$

where $\mathbf{S} = [\mathbf{s}(1), \dots, \mathbf{s}(T)]$ is the source signal matrix with T sample points.

By comparing Equations (24) and (25), the result follows from the fact that $\mathbf{C} = \mathcal{A}(\mathbf{S})^{-1}$. $\square$

Remark 4. From Theorem 2, the current estimation of source $\mathbf{y}_t$ estimated by the equivariant estimator relies only on the source signal matrix $\mathbf{S}$ and the source $\mathbf{s}_t$, rather than the mixing matrix $\mathbf{A}$.

Remark 5. Due to the fact that the estimator based on the optimizing cost function is equivariant [24], we only need to consider the equivariance of the complex-valued FastICA based on the global mixing–demixing system $\mathbf{C} = \mathbf{BA}$, i.e., the update learning rule of the estimated parameter for the ICA algorithm depends on the composition system.

4.2. Equivariance of Symmetric Complex-Valued FastICA

For convenience, we analyze the equivariance of the complex-valued FastICA algorithm by focusing on the symmetric version, in the sense that the source signals are estimated in parallel. In particular, the demixing vectors $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n$ are iterated simultaneously, as follows:

$$\mathbf{b}_1 \leftarrow E\{g(|\mathbf{b}_1^H \mathbf{x}|^2)(\mathbf{b}_1^H \mathbf{x})^* \mathbf{x}\} + E\{g'(|\mathbf{b}_1^H \mathbf{x}|^2)|\mathbf{b}_1^H \mathbf{x}|^2 + g(|\mathbf{b}_1^H \mathbf{x}|^2)\}\mathbf{b}_1 + E\{\mathbf{x}\mathbf{x}^T\}E\{g'(|\mathbf{b}_1^H \mathbf{x}|^2)(\mathbf{b}_1^H \mathbf{x})^{*2}\}\mathbf{b}_1^*,$$

$$\vdots$$

$$\mathbf{b}_n \leftarrow E\{g(|\mathbf{b}_n^H \mathbf{x}|^2)(\mathbf{b}_n^H \mathbf{x})^* \mathbf{x}\} + E\{g'(|\mathbf{b}_n^H \mathbf{x}|^2)|\mathbf{b}_n^H \mathbf{x}|^2 + g(|\mathbf{b}_n^H \mathbf{x}|^2)\}\mathbf{b}_n + E\{\mathbf{x}\mathbf{x}^T\}E\{g'(|\mathbf{b}_n^H \mathbf{x}|^2)(\mathbf{b}_n^H \mathbf{x})^{*2}\}\mathbf{b}_n^*,$$

where $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n$ is an orthonormal set. In order to prevent different parameters from converging to the same maxima, the estimate sources $y_1 = \mathbf{b}_1^H \mathbf{x}$, $y_2 = \mathbf{b}_2^H \mathbf{x}$, $\dots$, $y_n = \mathbf{b}_n^H \mathbf{x}$ need to be decorrelated, which can be accomplished by the following classic symmetric orthogonalization approach involving matrix square roots

$$\mathbf{B} \leftarrow (\mathbf{B}\mathbf{B}^H)^{-\frac{1}{2}}\mathbf{B}, \quad (26)$$

where $\mathbf{B} = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ is the matrix of the vectors. Note that the inverse square root $(\mathbf{B}\mathbf{B}^H)^{-\frac{1}{2}}$ can be acquired by the EVD of $\mathbf{B}\mathbf{B}^H = \mathbf{E}\text{diag}(d_1, d_2, \dots, d_n)\mathbf{E}^H$. For any vector $\mathbf{h} = (h_1, \dots, h_n)^T$, we denote the following:

$$\text{diag}(\mathbf{h}) = \begin{pmatrix} h_1 & & 0 \\ & \ddots & \\ 0 & & h_n \end{pmatrix}.$$

Hence, we have the following:

$$(\mathbf{B}\mathbf{B}^H)^{-\frac{1}{2}} = \mathbf{E}\text{diag}(d_1^{-\frac{1}{2}}, d_2^{-\frac{1}{2}}, \dots, d_n^{-\frac{1}{2}})\mathbf{E}^H. \quad (27)$$

Using the matrix representation, the symmetric version of the complex-valued FastICA algorithm can be rewritten by the following learning rule:

Step 1 Choose arbitrary initial values for the $\mathbf{b}_1, \dots, \mathbf{b}_n \in O(n)$; orthogonalize matrix $\mathbf{B} = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ as in Step 2 below.

Step 2 Let $\mathbf{y} = \mathbf{B}\mathbf{x}$, run iteration

$$\begin{aligned} \mathbf{B} &\leftarrow -E\{\text{diag}(g(|\mathbf{y}|^2))\mathbf{y}\mathbf{x}^H\} + E\{\text{diag}(g'(|\mathbf{y}|^2))|\mathbf{y}|^2 \\ &\quad + g(|\mathbf{y}|^2)\}\mathbf{B} + E\{\text{diag}(g'(|\mathbf{y}|^2))\mathbf{y}^2\}E\{\mathbf{y}^*\mathbf{x}^H\}, \\ \mathbf{B} &\leftarrow (\mathbf{B}\mathbf{B}^H)^{-\frac{1}{2}}\mathbf{B} \end{aligned} \quad (28)$$

until convergence.

Definition 4. Assume that $\mathbf{B} = \mathcal{B}(\mathbf{X})$ is an estimator of the demixing matrix $\mathbf{B}$ in the complex-valued ICA model (1), which relies on the sample matrix $\mathbf{X} = [\mathbf{x}(1), \dots, \mathbf{x}(T)]$ with T sample points. Then, we call the estimator $\mathcal{B}(\mathbf{X})$ linear-equivariant, if

$$\mathbf{C} = \mathcal{B}(\mathbf{S}), \quad (29)$$

where $\mathbf{x}(i) = \mathbf{A}\mathbf{s}(i)$, $i = 1, \dots, T$, $\mathbf{S} = [\mathbf{s}(1), \dots, \mathbf{s}(T)]$ and $\mathbf{C} = \mathbf{B}\mathbf{A}$.

From Definition 4, we can obtain the equivariance of the symmetric complex-valued FastICA algorithm:

Theorem 3. The algorithm (28) for the complex-valued BSS is linear-equivariant and $\mathcal{B}(\mathbf{S}) = \mathcal{A}(\mathbf{S})^{-1}$.

Proof. We denote by $\mathbf{C}$ the mixing–demixing matrix $\mathbf{B}\mathbf{A}$, and the estimation of sources can be rewritten as follows:

$$\mathbf{y} = \mathbf{C}\mathbf{s}. \quad (30)$$

Plugging the symmetric orthogonalization into the algorithm (28), by post-multiplying $\mathbf{A}$ on both sides of the complex-valued FastICA algorithm, one can yield

$$\begin{aligned} \mathbf{C} &\leftarrow -E\{\text{diag}(g(|\mathbf{C}\mathbf{s}|^2))\mathbf{C}\mathbf{s}\mathbf{s}^H\mathbf{C}^H\mathbf{C}\} + E\{\text{diag}(g'(|\mathbf{C}\mathbf{s}|^2))|\mathbf{C}\mathbf{s}|^2 \\ &\quad + g(|\mathbf{C}\mathbf{s}|^2)\}\mathbf{C} + E\{\text{diag}(g'(|\mathbf{C}\mathbf{s}|^2))(\mathbf{C}\mathbf{s})^2\}E\{(\mathbf{C}\mathbf{s})^*\mathbf{s}^H\mathbf{C}^H\mathbf{C}\}. \end{aligned} \quad (31)$$

From the above equation, we have $\mathbf{C} = \mathcal{B}(\mathbf{S})$. Hence, the complex-valued BSS is linear-equivariant. The residual conclusion $\mathcal{B}(\mathbf{S}) = \mathcal{A}(\mathbf{S})^{-1}$ can be obtained by combining Theorem 2, Definition 4 and Theorem 3. $\square$

The results obtained demonstrate the following: (1) Theorems 2 and 3 suggest that the complex-valued FastICA estimator is of equivariance, namely, the performance of the complex-valued FastICA algorithm depends only on the source signal $\mathbf{s}$, rather than the mixing matrix $\mathbf{A}$ and demixing matrix $\mathbf{B}$; (2) There is an invertible relationship between the mixing matrix $\mathbf{A}$ and demixing matrix $\mathbf{B}$ under Assumption 2.

5. Conclusions

In order to provide a rigorous mathematical treatment of the robust measurement and equivariance for the complex-valued FastICA estimator, this paper analyzed the statistical properties of the complex-valued FastICA algorithm in the context of ICA over the complex number domain. Firstly, a closed-form expression of the complex-valued FastICA functional was derived and used to measure robustness against outliers. We found that the complex-valued FastICA algorithm based on Tukey's single-parameter M-estimator cost function had the best separation performance in both circular and noncircular scenarios. Then, we proved that the complex-valued FastICA algorithm is equivariant in the sense that the global mixing-demixing system of the algorithm depended only on the source signals.

Author Contributions: Conceptualization, J.E. and M.Y.; methodology, J.E. and M.Y.; software, M.Y. writing—original draft preparation, J.E. and M.Y.; validation, J.E. and M.Y.; writing—review and editing, J.E. and M.Y.; visualization, J.E. and M.Y.; supervision, M.Y.; funding acquisition, M.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Guangxi Natural Science Foundation under grant 2023GXNSFBA026180, Middle-aged and Young Teachers' Basic Ability Promotion Project of Guangxi Province under grant 2024KY0181, and the Natural Science Basic Research Program of Shaanxi under grant 2024JC-YBMS-043.

Data Availability Statement: The data used to support the findings of this study are included within the article.

Acknowledgments: Sincere thanks are given to anonymous referees for their insightful comments which are valuable for improving this study.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

Proof of Lemma 1. (a) By determining the derivative $y_{x,i}(F_t)$ with respect to t at point $t = 0$, we have the following:

$$\begin{aligned} IF(y_{x,i}) &= \left. \frac{\partial y_{x,i}(F_t)}{\partial t} \right|_{t=0} \\ &= \left. \frac{\partial \mathbf{b}_i^H(F_t)}{\partial t} \right|_{t=0} \cdot (\mathbf{x} - \mathbf{m}_x) \Big|_{t=0} + \mathbf{b}_i^H(F_t) \Big|_{t=0} \cdot \left. \frac{\partial (\mathbf{x} - \mathbf{m}_x)}{\partial t} \right|_{t=0} \\ &= IF(\mathbf{b}_i)^H (\mathbf{x} - \mathbf{m}_x) - \mathbf{b}_i^H IF(\mathbf{m}_x) \\ &= IF(\mathbf{b}_i)^H (\mathbf{x} - \mathbf{m}_x) - \mathbf{b}_i^H (\mathbf{z} - \mathbf{m}_x) \\ &= IF(\mathbf{b}_i)^H (\mathbf{x} - \mathbf{m}_x) - \bar{s}_i. \end{aligned}$$

(b) $y_{x,k}(F) = \mathbf{b}_k^H(F) [\mathbf{x} - \mathbf{m}_x] = \mathbf{b}_k^H \mathbf{A} \mathbf{s} = \mathbf{e}_k^H \mathbf{s} = s_k$, the notation of $\mathbf{e}_k$ represents a real-valued vector with element 1 in the k th component, and element 0 elsewhere. By $E\{g(|s_k|^2)s_k s_l\} = E\{g(|s_k|^2)s_k\}E\{s_l\} = 0$, $l \neq k$, we have the following:

$$\begin{aligned} L_k(F) &= E\{g(|s_k|^2)s_k(\mathbf{A}\mathbf{s})^*\} = \mathbf{A}^* E\{g(|s_k|^2)|s_k|^2 \mathbf{e}_k\} \\ &= E\{g(|s_k|^2)|s_k|^2\} \mathbf{a}_k^* = c_k \mathbf{a}_k^*, \end{aligned}$$

where $\mathbf{A}\mathbf{e}_k = \mathbf{a}_k$.

(c) By the assumption $E\{\mathbf{s}\} = \mathbf{0}$, $E\{\mathbf{s}\mathbf{s}^H\} = \mathbf{I}$; that is, $E\{s_i\} = 0$, $E\{s_i s_j^*\} = 0$, $E\{|s_i|^2\} = 1$, $i \neq j = 1, \dots, n$, one can obtain the following:

$$\begin{aligned} B_k(F_A) &= E\{g(|s_k|^2)\mathbf{A}\mathbf{s}(\mathbf{A}\mathbf{s})^H\} = \mathbf{A} E\{g(|s_k|^2)(\mathbf{s}\mathbf{s})^H\} \mathbf{A} \\ &= \mathbf{A} \begin{pmatrix} E\{g(|s_k|^2)\} & & & \\ & \ddots & & \\ & & E\{g(|s_k|^2)|s_k|^2\} & \\ & & & \ddots & \\ & & & & E\{g(|s_k|^2)\} \end{pmatrix} \mathbf{A}^H \\ &= \mathbf{A}(c_k \mathbf{e}_k \mathbf{e}_k^T + \mu_k \sum_{l \neq k} \mathbf{e}_l \mathbf{e}_l^T) \mathbf{A}^H \\ &= c_k \mathbf{A} \mathbf{e}_k \mathbf{e}_k^H \mathbf{A}^H + \mu_k \mathbf{A} \mathbf{A}^H - \mu_k \mathbf{a}_k \mathbf{a}_k^H \\ &= \mu_k \mathbf{A} \mathbf{A}^H + (c_k - \mu_k) \mathbf{a}_k \mathbf{a}_k^H. \end{aligned}$$

(d) According to the constraint condition (7), if $i < k$: $\mathbf{b}_k(F_t)^H \mathbf{C}_x(F_t) \mathbf{b}_i(F_t) = 0$, after calculating the derivation with respect to $t = 0$ on both sides, we have the following:

$$\begin{aligned}
0 &= IF(\mathbf{b}_k)^H \mathbf{A} \mathbf{A}^H \mathbf{b}_i + \mathbf{b}_k^H IF(\mathbf{C}_x) \mathbf{b}_i + \mathbf{b}_k^H \mathbf{A} \mathbf{A}^H IF(\mathbf{b}_i) \\
&= IF(\mathbf{b}_k)^H \mathbf{a}_i + \mathbf{b}_k^H [(\mathbf{z} - \mathbf{m}_x)(\mathbf{z} - \mathbf{m}_x)^H - \mathbf{A} \mathbf{A}^H] \mathbf{b}_i + \mathbf{a}_k^H IF(\mathbf{b}_i) \\
&= IF(\mathbf{b}_k)^H \mathbf{a}_i + \bar{s}_k \bar{s}_i^* - \mathbf{b}_k^H \mathbf{A} \mathbf{A}^H \mathbf{b}_i + \mathbf{a}_k^H IF(\mathbf{b}_i),
\end{aligned}$$

which is equivalent to $\mathbf{a}_k^H IF(\mathbf{b}_i) = -IF(\mathbf{b}_k)^H \mathbf{a}_i - \bar{s}_k \bar{s}_i^*$, where $\bar{s}_i = \mathbf{b}_i^H (\mathbf{z} - \mathbf{m}_x)$, $i = 1, \dots, k-1$ denotes the projection of the centered contamination point $\mathbf{z}$ into the direction $\mathbf{b}_i$.

(e) According to the constraint condition (7), if $j = k$: $\mathbf{b}_k(F_t)^H \mathbf{C}_x(F_t) \mathbf{b}_k(F_t) = 1$, after calculating the derivation with respect to $t = 0$ on both sides, we have the following:

$$\begin{aligned}
0 &= IF(\mathbf{b}_k)^H \mathbf{A} \mathbf{A}^H \mathbf{b}_k + \mathbf{b}_k^H IF(\mathbf{C}_x) \mathbf{b}_k + \mathbf{b}_k^H \mathbf{A} \mathbf{A}^H IF(\mathbf{b}_k) \\
&= IF(\mathbf{b}_k)^H \mathbf{a}_k + \mathbf{b}_k^H [(\mathbf{z} - \mathbf{m}_x)(\mathbf{z} - \mathbf{m}_x)^H - \mathbf{A} \mathbf{A}^H] \mathbf{b}_k + \mathbf{a}_k^H IF(\mathbf{b}_k) \\
&= IF(\mathbf{b}_k)^H \mathbf{a}_k + \bar{s}_k \bar{s}_k^* - \mathbf{b}_k^H \mathbf{A} \mathbf{A}^H \mathbf{b}_k + \mathbf{a}_k^H IF(\mathbf{b}_k) \\
&= IF(\mathbf{b}_k)^H \mathbf{a}_k + \mathbf{a}_k^H IF(\mathbf{b}_k) + |\bar{s}_k|^2 - 1,
\end{aligned}$$

which is equivalent to $2\text{Re}(\mathbf{a}_k^H IF(\mathbf{b}_k)) = 1 - |\bar{s}_k|^2$, where $\bar{s}_k = \mathbf{b}_k^H (\mathbf{z} - \mathbf{m}_x)$ denotes the projection of the centered contamination point $\mathbf{z}$ into the direction $\mathbf{b}_k$. Note that $\mathbf{A}^H \mathbf{b}_k = \mathbf{e}_k$; $\mathbf{A} \mathbf{e}_k = \mathbf{a}_k$; $\mathbf{a}_k^H \mathbf{a}_k = 0$, $\mathbf{a}_k^H \mathbf{a}_i = 0$ ($i < k$). $\square$

Appendix B

Proof of Lemma 2. (a) and (b) Recalling

$$R_k(F) = \lambda_k \mathbf{C}_x^T \mathbf{b}_k^* + \sum_{j=1}^{k-1} \lambda_j \mathbf{C}_x^T \mathbf{b}_j^*,$$

by multiplying $\mathbf{b}_i^T$ from the right-hand side of the above equation and plugging the constraint condition (7), one can obtain the following: $\lambda_i(F) = \mathbf{b}_i^T R_k(F)$, $i = 1, \dots, k$. We note that $L_k(F) = R_k(F)$ and recalling Lemma 1 (b), we have the following:

$$\lambda_i(F_A) = c_k (\mathbf{b}_i^H \mathbf{a}_k)^*.$$

Hence, $\lambda_k(F) = 0$, if $i = 1, \dots, k-1$; $\lambda_i(F) = c_k$ if $i = k$.

(c) and (d) According to the proof of Lemma 1, we have the following:

$$\begin{aligned}
\lambda_i(F) &= \mathbf{b}_i^T L_k(F) \\
&= \mathbf{b}_i^T E\{g(|y_{x,k}|^2) y_{x,k} [\mathbf{x} - \mathbf{m}_x]^*\} \\
&= E\{g(|y_{x,k}|^2) y_{x,k} y_{x,i}^*\}, \quad i = 1, \dots, k
\end{aligned}$$

due to

$$\lambda_i(F_t) = (1-t)E\{g(|y_{x,k}|^2) y_{x,k} y_{x,i}^*\} + tE\{g(|y_{z,k}|^2) y_{z,k} y_{z,i}^*\},$$

after calculating the derivation with respect to $t = 0$, one can achieve the following:

$$\begin{aligned}
IF(\lambda_i) &= -\lambda_i(F) + 2E\{g'(|y_{x,k}|^2) \text{Re}(IF(y_{x,k}) y_{x,k}^*) y_{x,k} y_{x,i}^*\} \\
&\quad + E\{g(|y_{x,k}|^2) IF(y_{x,k}) y_{x,i}^*\} + E\{g(|y_{x,k}|^2) y_{x,k} IF(y_{x,i})^*\} + g(|\bar{s}_k|^2) \bar{s}_k \bar{s}_i^*.
\end{aligned}$$

When $i = 1, \dots, k-1$, after tedious calculations, we have the following:

$$\begin{aligned}
IF(\lambda_i) &= -\mu_k [IF(\mathbf{b}_i)^T \mathbf{a}_k^* - \bar{s}_k \bar{s}_i^*] + c_k \mathbf{a}_k^H IF(\mathbf{b}_i) - \rho_k \bar{s}_i^* + g(|\bar{s}_k|^2) \bar{s}_k \bar{s}_i^* \\
&= (c_k - \mu_k) \mathbf{a}_k^H IF(\mathbf{b}_i) + (g(|\bar{s}_k|^2) - \mu_k) \bar{s}_k \bar{s}_i^* - \rho_k \bar{s}_i^*.
\end{aligned}$$

which completes the proof of Lemma 2 (c).

When $i = k$, after tedious calculations, we have the following:

$$\begin{aligned}
IF(\lambda_i) &= -c_k + E\{g'(|s_k|^2)(|s_k|^2 - |s_k|^2|\bar{s}_k|^2 - 2\text{Re}(\bar{s}_k s_k^*))|s_k|^2\} + \mu_k \mathbf{a}_k^T IF(\mathbf{b}_k)^* \\
&\quad + (c_k - \mu_k) \mathbf{a}_k^T IF(\mathbf{b}_k)^* - E\{g(|s_k|^2)\bar{s}_k s_k^*\} + c_k \mathbf{a}_k^H IF(\mathbf{b}_k) \\
&\quad - E\{g(|s_k|^2)s_k \bar{s}_k^*\} + g(|\bar{s}_k|^2)|\bar{s}_k|^2 \\
&= -c_k + E\{g'(|s_k|^2)v_k(\bar{s}_k)|s_k|^2\} + 2c_k \text{Re}(\mathbf{a}_k^H IF(\mathbf{b}_k)) \\
&\quad - 2\text{Re}(E\{g(|s_k|^2)\bar{s}_k s_k^*\}) + g(|\bar{s}_k|^2)|\bar{s}_k|^2 \\
&= (g(|\bar{s}_k|^2) - c_k)|\bar{s}_k|^2 + E\{g'(|\bar{s}_k|^2)v_k(\bar{s}_k)|\bar{s}_k|^2\} - 2\text{Re}(E\{g(|\bar{s}_k|^2)\bar{s}_k s_k^*\})
\end{aligned}$$

which completes the proof of Lemma 2 (d). $\square$

References

- Comon, P. Independent component analysis, A new concept? *Signal Process.* **1994**, *36*, 287–314. [\[CrossRef\]](#)
- Hyvärinen, A.; Karhunen, J.; Oja, E. *Independent Component Analysis*; Wiley and Sons: New York, NY, USA, 2001.
- Chen, Y.H.; Wang, S.P. Low-cost implementation of independent component analysis for biomedical signal separation using very-large-scale integration. *IEEE Trans. Circuit Syst. II* **2020**, *67*, 3437–3441. [\[CrossRef\]](#)
- Schell, A.; Oberhauser, H. Nonlinear independent component analysis for discrete-time and continuous-time signals. *Ann. Stat.* **2023**, *51*, 487–518. [\[CrossRef\]](#)
- Liu, J.; Ye, J.; E, J. A multi-scale forecasting model for CPI based on independent component analysis and non-linear autoregressive neural network. *Phys. A* **2023**, *609*, 128369. [\[CrossRef\]](#)
- E, J.; He, K.; Liu, H.; Ji, Q. A novel separation-ensemble analyzing and forecasting method for the gold price forecasting based on RLS-type independent component analysis. *Expert Syst. Appl.* **2023**, *232*, 120852. [\[CrossRef\]](#)
- Bingham, E.; Hyvärinen, A. A fast fixed-point algorithm for independent component analysis of complex valued signals. *Int. J. Neural Syst.* **2000**, *10*, 1–8. [\[CrossRef\]](#) [\[PubMed\]](#)
- Novey, M.; Adali, T. On extending the complex FastICA algorithm to noncircular sources. *IEEE Trans. Signal Process.* **2008**, *56*, 2148–2154. [\[CrossRef\]](#)
- Novey, M.; Adali, L. Complex ICA by negentropy maximization. *IEEE Trans. Neural Netw.* **2008**, *19*, 596–609. [\[CrossRef\]](#) [\[PubMed\]](#)
- Qian, G.; Wei, P. Stability analysis of complex ica by negentropy maximization: A unique perspective. *Neurocomputing* **2016**, *214*, 80–85. [\[CrossRef\]](#)
- Mika, D. Fast gradient algorithm with toral decomposition for complex ICA. *Mech. Syst. Signal Process.* **2022**, *178*, 109266. [\[CrossRef\]](#)
- Koldovský, Z.; Tichavský, P. Gradient algorithms for complex non-Gaussian independent component/vector extraction, question of convergence. *IEEE Trans. Signal Process.* **2019**, *67*, 1050–1064. [\[CrossRef\]](#)
- E, J.; Ye, J.; He, L.; Jin, H. Performance analysis for complex-valued FastICA and its improvement based on the Tukey M-estimator. *Digit. Signal Process.* **2021**, *115*, 103077. [\[CrossRef\]](#)
- Loesch, B.; Yang, B. Cramér-Rao bound for circular and noncircular complex independent component analysis. *IEEE Trans. Signal Process.* **2012**, *61*, 365–379. [\[CrossRef\]](#)
- Kautský, V.; Koldovský, Z.; Tichavský, P.; Zarzoso, V. Cramér-Rao bounds for complex-valued independent component extraction: Determined and piecewise determined mixing models. *IEEE Trans. Signal Process.* **2020**, *68*, 5230–5243. [\[CrossRef\]](#)
- Fu, G.S.; Phlypo, R.; Anderson, M.; Adali, T. Complex independent component analysis using three types of diversity: Non-Gaussianity, nonwhiteness and noncircularity. *IEEE Trans. Signal Process.* **2015**, *63*, 794–805. [\[CrossRef\]](#)
- Koldovský, Z.; Kautský, V.; Tichavský, P.; Čmejla, J.; Málek, J. Dynamic independent component/vector analysis: Time-variant linear mixtures separable by time-invariant beamformers. *IEEE Trans. Signal Process.* **2021**, *69*, 2158–2173. [\[CrossRef\]](#)
- Hyvärinen, A.; Oja, E. A fast fixed-point algorithms for independent component analysis. *Neural Comput.* **1997**, *9*, 1483–1492. [\[CrossRef\]](#)
- Hyvärinen, A. Fast and robust fixed-point algorithms for independent component analysis. *IEEE Trans. Neural Netw.* **1999**, *10*, 626–634. [\[CrossRef\]](#) [\[PubMed\]](#)
- Chao, J.C.; Douglas, S.C. *A Robust Complex FastICA Algorithm Using the Huber M-Estimator Cost Function*; Springer: Berlin/Heidelberg, Germany, 2007.
- Adali, T.; Schreier, P.J.; Scharf, L.L. Complex-valued signal processing: The proper way to deal with impropriety. *IEEE Trans. Signal Process.* **2011**, *59*, 5101–5125. [\[CrossRef\]](#)
- Hampel, F.R.; Ronchetti, E.M.; Stahel, W.A. *Robust Statistics: The Approach Based on Influence Functions*; Wiley and Sons: New York, NY, USA, 1986.
- Lehmann, E.L. *Testing Statistical Hypothesis*; Wiley: New York, NY, USA, 1959.
- Cardoso, J.F. Blind signal separation: Statistical principles. *Proc. IEEE* **1998**, *86*, 2009–2025. [\[CrossRef\]](#)

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.