

Research Article

Blind Source Separation for Complex-Valued Signals Using Generalized Autocorrelation

Xiaogang Tang,^{1,2} Sun'an Wang,¹ and Jiong Li²

¹School of Mechanical Engineering, Xian Jiaotong University, No.28, Xianning West Road, Xian, Shaanxi 710049, China

²School of Space Information, Space Engineering University of PLA, No.1, Bayi Road, Huairou District, Beijing 101416, China

Correspondence should be addressed to Jiong Li; lij_2015@126.com

Received 11 February 2018; Revised 24 May 2018; Accepted 20 June 2018; Published 12 July 2018

Academic Editor: Bogdan Dumitrescu

Copyright © 2018 Xiaogang Tang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We introduce a new complex-valued blind source separation approach, based on generalized autocorrelations of sources, to improve the spectrum efficiency for the next-generation wireless communications system. The proposed algorithm considers the temporal structures of communication signals and the natural gradient-based method is used to optimize the demixing matrix. In addition, the local stability condition is proved. Simulation results are presented showing the superior performance of the proposed algorithm in the intersymbol interference of the estimated signals.

1. Introduction

For the next-generation mobile communications system (5G), spectrum efficiency is a critical performance metric, which aims at achieving more than 10 times spectrum efficiency compared with the current communications system (4G) [1]. Since radios must either transmit or receive on the same channel, but not simultaneously, in previous wireless communications systems, the frequency resource is not utilized sufficiently. Fortunately, an emerging technique, co-frequency and co-time full-duplex (CCFD) [2–4], is able to address this issue. CCFD enables radios to transmit and receive signals on the same channel simultaneously and, thus, theoretically, can double the spectrum efficiency. CCFD technique claims to be the most potential duplex scheme for the 5G network. But, there is a challenge which lies in the application of CCFD technique, i.e., mitigating the local self-interference (SI) [3, 4]. Since the transmit and receive antenna work on the same frequency band, traditional interference cancellation techniques are invalid. As our previous work shown [5], blind source separation (BSS) [6] has an excellent capability to address such issue.

Since BSS is able to estimate original signals from their observed sensors' signals without knowing both the mixing process and the sources, it has been one of the most attractive

technologies in many fields, such as communications [7–9], speech signal processing [10, 11], and medical signal processing [12–14]. Depending on the application, the sources may be real or complex. This paper focuses on the separation of baseband communication signals which are usually complex-valued.

Independent component analysis (ICA) is a common method to realize BSS, which can estimate demixing matrix and separate signals up to an arbitrary scaling and permutation ambiguities if the source signals are mutually independent and no more than one Gaussian source exists. And most of complex ICA algorithms exploit the non-Gaussianity or sample dependency of sources. Non-Gaussianity is often measured by high order statistics. Fourth-order moments is used in fourth-order blind identification (FOBI) [15], which is very efficient but requires that the sources have different kurtosis. However, sources have identical distribution in many applications, and then FOBI is invalid. Joint approximate diagonalization of eigenmatrices (JADE) [16] is based on fourth-order cumulants and performance suffers when the number of sources increases. The strong uncorrelating transform (SUT) [17] and the generalized uncorrelating transform (GUT) [18] estimate mixing matrix by jointly diagonalizing the covariance matrix and pseudo-covariance matrix of mixed signals. Also, it is important to note that,

for noncircular sources, SUT can achieve separation of them-selves as long as they have distinct noncircularity indices.

The above methods make use of the non-Gaussianity property and neglect sample dependency. In this paper, we take the temporal characteristics of communication signals into account. We develop a new complex BSS algorithm motivated by generalized autocorrelations of the source signals. A natural gradient-based algorithm is derived for the generalized autocorrelations contrast function. Stability conditions are then derived in the complex domain for the generalized autocorrelations contrast function.

This paper is organized as follows. First, the BSS problem is formulated in Section 2. We then introduce the new contrast function based on the generalized autocorrelations of source signals and derive a natural gradient-based algorithm in Section 3. The performance of the algorithm is demonstrated with simulations in Section 4. Conclusions are drawn in Section 5.

2. Problem Formulation

Considering there are N sensors and N independent sources, the instantaneous linear mixtures of these sources are observed at the sensors:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) \quad (1)$$

where $\mathbf{s}(t) = [s_1(t), s_2(t), \dots, s_N(t)]^T$ (superscript T denotes transpose) is a vector of unknown zero mean and unit-variance source signals, $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_N(t)]^T$ is the observation vector, and \mathbf{A} is the unknown $N \times N$ mixing matrix. We impose another assumption on the source signals that the source signals have specific temporal structures, linear or nonlinear autocorrelations.

In general, preprocessing operations of observed signals are needed before performing BSS algorithm. Two common preprocessing operations are removing mean and whitening. The whitening matrix \mathbf{Q} can be obtained using the eigenvalue decomposition of covariance matrix $\mathbf{R}_x = E[\mathbf{x}(t)\mathbf{x}(t)^H]$ (superscript H denotes conjugate transpose; $E[\bullet]$ denotes expectation).

$$\mathbf{Q} = \mathbf{D}^{-1/2}\mathbf{V}^H \quad (2)$$

where \mathbf{D} and \mathbf{V} are the eigenvalue matrix and the eigenvector matrix of the covariance matrix \mathbf{R}_x . The observations are whitened by \mathbf{Q} :

$$\mathbf{z}(t) = \mathbf{Q}\mathbf{x}(t) \quad (3)$$

Then, $\mathbf{R}_z = E[\mathbf{z}(t)\mathbf{z}(t)^H] = \mathbf{Q}\mathbf{A}\mathbf{R}_s(\mathbf{Q}\mathbf{A})^H$, due to $\mathbf{R}_s = \mathbf{I}$, $\mathbf{R}_z = \mathbf{I}$, and $\mathbf{Q}\mathbf{A}$ is an orthogonal matrix which imposes orthogonal constraint on the demixing matrix \mathbf{W} . The sources can be estimated by

$$\mathbf{y}(t) = \mathbf{W}\mathbf{z}(t) \quad (4)$$

where $\mathbf{y}(t) = [y_1(t), y_2(t), \dots, y_N(t)]^T$ is the estimation of $\mathbf{s}(t)$.

3. Proposed Algorithm

Now we generalize the framework in [19] for complex-valued signals. We start from an arbitrary nonlinear contrast function and prove that its extrema coincide with the sources. The generalized autocorrelation of estimated signal is calculated with the selected nonlinear function and the nonlinear function can be chosen freely. Our contrast function based on complex generalized autocorrelation functions is

$$J(\mathbf{w}_n) = E \left\{ G \left(|\mathbf{w}_n^H \mathbf{z}(t)|^2 \right) G \left(|\mathbf{w}_n^H \mathbf{z}(t-\tau)|^2 \right) \right\} \quad (5)$$

where G is a differentiable function which measures the generalized autocorrelation degree of the source signal, \mathbf{w}_n is the n th volume vector of demixing matrix \mathbf{W} , and $|\mathbf{w}_n|^2 = 1$, τ is a delay in time. Finding the extrema of a contrast is a well-defined problem only if the function is real. So we make our contrast functions operate on absolute values rather than complex values. Examples of choices of the differentiable function G are $G_1(u) = u$, $G_2(u) = u^2$, $G_3(u) = \log[\cosh(u)]$. For simplicity, the time index t is omitted in the following, i.e., $\mathbf{z}(t) = \mathbf{z}$, $\mathbf{z}(t-\tau) = \mathbf{z}_\tau$.

Theorem 1 gives the local stability of the estimators, the proof of which is given in Appendix.

Theorem 1. Assume that the input data follows model (1). The prewhitening operation is performed on the observation signals as (3). Furthermore, assume that $\{s_n, s_{n\tau}\}$ and $\{s_l, s_{l\tau}\}$ are mutually independent and $E\{|s_n|^2\} = 1$, $E\{|s_n|^2\} = 0$. Then, the local maximum (respectively, minimum) of $J(\mathbf{w}_n)$ under the constraint $\|\mathbf{w}_n\| = 1$ includes those rows of the inverse of the global mixing matrix $\mathbf{Q}\mathbf{A}$, so the corresponding sources satisfy

$$\begin{aligned} & E \left\{ \alpha_3 - \alpha_1 \right. \\ & \left. + 2 \operatorname{Re} \{ s_n s_{n\tau} s_1^* s_{1\tau}^* + s_n s_{n\tau}^* s_1^* s_{1\tau} \} g(|s_1|^2) g(|s_{1\tau}|^2) \right\} \quad (6) \\ & < 0 \quad (\text{resp. } > 0) \end{aligned}$$

where we define

$$\begin{aligned} \alpha_1 &= E \left\{ |s_1|^2 g(|s_1|^2) G(|s_{1\tau}|^2) \right. \\ & \left. + |s_{1\tau}|^2 G(|s_1|^2) g(|s_{1\tau}|^2) \right\} \quad (7) \end{aligned}$$

$$\begin{aligned} \alpha_2 &= 2E \left\{ |s_1|^4 g'(|s_1|^2) G(|s_{1\tau}|^2) \right. \\ & \left. + 2 |s_1|^2 |s_{1\tau}|^2 g(|s_1|^2) g(|s_{1\tau}|^2) \right\} \quad (8) \end{aligned}$$

$$\begin{aligned} \alpha_3 &= E \left\{ g(|s_1|^2) G(|s_{1\tau}|^2) + G(|s_1|^2) g(|s_{1\tau}|^2) \right. \\ & \left. + |s_1|^2 g'(|s_1|^2) G(|s_{1\tau}|^2) \right\} \quad (9) \end{aligned}$$

$$\begin{aligned} & + |s_{1\tau}|^2 G(|s_1|^2) g'(|s_{1\tau}|^2) \end{aligned}$$

where function g' is the derivative of g .

We now give the conjugate gradient algorithm for complex signals under model (1). The problem given in (5) can be written as the following Lagrangian function:

$$J_0(\mathbf{w}_n) = J(\mathbf{w}_n) + \lambda (\|\mathbf{w}_n\|^2 - 1) \quad (10)$$

where the Lagrangian multiplier λ is a real number. The conjugate gradient of the contrast function J with respect to \mathbf{w}_n can be obtained as

$$\frac{\partial J_0(\mathbf{w}_n)}{\partial \mathbf{w}_n^*} = \frac{\partial J(\mathbf{w}_n)}{\partial \mathbf{w}_n^*} + \lambda \mathbf{w}_n \quad (11)$$

$$\begin{aligned} \frac{\partial J(\mathbf{w}_n)}{\partial \mathbf{w}_n^*} &= E \left\{ g \left(\left| \mathbf{w}_n^H \mathbf{z} \right|^2 \right) G \left(\left| \mathbf{w}_n^H \mathbf{z}_\tau \right|^2 \right) \mathbf{w}_n^T \mathbf{z}^* \mathbf{z} \right. \\ &\quad \left. + G \left(\left| \mathbf{w}_n^H \mathbf{z} \right|^2 \right) g \left(\left| \mathbf{w}_n^H \mathbf{z}_\tau \right|^2 \right) \mathbf{w}_n^T \mathbf{z}_\tau^* \mathbf{z}_\tau \right\} \end{aligned} \quad (12)$$

Function g is the derivative of G ; \mathbf{z}^* designates the complex conjugate of \mathbf{z} . Thus, the conjugate gradient-based update rule of \mathbf{w}_n can be written as

$$\begin{aligned} \mathbf{w}_n &= \mathbf{w}_n - \mu \left[\frac{\partial J(\mathbf{w}_n)}{\partial \mathbf{w}_n^*} - \operatorname{Re} \left\{ \mathbf{w}_n^H \frac{\partial J(\mathbf{w}_n)}{\partial \mathbf{w}_n^*} \right\} \mathbf{w}_n \right] \\ \mathbf{w}_n &= \frac{\mathbf{w}_n}{\|\mathbf{w}_n\|} \end{aligned} \quad (13)$$

where $\mu > 0$ is the real-valued step-size.

It is known that, in general, relative gradient-based method converges slower than natural gradient-based method. To speed up the convergence, the natural gradient-based update rule of (13) can be written as

$$\begin{aligned} \mathbf{W} &= \mathbf{W} - \mu \frac{\partial J_0(\mathbf{W})}{\partial \mathbf{W}} \mathbf{W}^H \mathbf{W} \\ \mathbf{W} &= \mathbf{W} \left(\mathbf{W}^H \mathbf{W} \right)^{-1/2} \end{aligned} \quad (14)$$

4. Simulation and Results

The proposed algorithm is compared with three competitive complex ICA algorithms: complex FastICA [20], EBM [21], and EASI [22]. FastICA is a Newton-based ICA algorithm which converges fast. EBM is a conjugate gradient-based algorithm. EASI is a relative gradient-based algorithm. In general, Newton method converges faster than conjugate gradient and relative gradient methods, and relative gradient method converges slowest. The performance index is defined as (15) which means the average intersymbol interference of the estimation sources.

$$\begin{aligned} PI &= \frac{1}{2N} \left[\sum_{k=1}^N \left(\sum_{l=1}^N \frac{|U_{kl}|^2}{\max_n |U_{kn}|^2} - 1 \right) \right. \\ &\quad \left. + \sum_{l=1}^N \left(\sum_{k=1}^N \frac{|U_{kl}|^2}{\max_n |U_{nl}|^2} - 1 \right) \right] \end{aligned} \quad (15)$$

where $\mathbf{U} = \mathbf{WQA}$ is the combined separation-whitening-mixing matrix, and U_{kl} is the (k, l) th entry of \mathbf{U} .

This article considers the situation where two traditional and classic modulation signals are mixed, i.e., two 8PSK signals. Figure 1 shows the constellation figures of two original 8PSK signals, their mixtures, and the estimations using the proposed algorithm employing the nonlinear function $G_1(u) = u$. In this simulation, the two 8PSK are with the following parameters: symbol rate $R_s = 198 \text{ kBps}$, raised cosine filter which has 0.5 roll-off factor, and sampling frequency $16R_s$. The mixing matrix is randomly generated as

$$\mathbf{A} = \begin{bmatrix} 0.5559 - 0.4087i & -0.8623 + 0.0475i \\ -0.8571 - 1.1729i & 1.0418 + 0.9968i \end{bmatrix} \quad (16)$$

From Figure 1, we can see that the source signals are properly estimated. As most complex BSS algorithms, there are some ambiguities, including amplitude, order, and phase. The combined separation-whitening-mixing matrix is estimated as

$$\mathbf{U} = \begin{bmatrix} -0.0012 - 0.0050i & \mathbf{0.7640 - 0.6459i} \\ \mathbf{0.5893 + 0.8082i} & -0.0092 - 0.0167i \end{bmatrix} \quad (17)$$

In order to exploit the influence of time delay τ to the proposed algorithm, the autocorrelation of the source signal is calculated $\xi(\tau) = E\{|s_n(t)s_n^*(t-\tau)|\}$ with respect to different τ and the separation performance of the proposed algorithm with different time delay τ is simulated. Since the source signals are two 8PSK signals with the same modulation parameters, they have the same autocorrelation function. The autocorrelation function is shown in Figure 2 and the separation performance of the proposed algorithm is given in Figure 3. From Figures 2 and 3, we can find that the separation results remain at a certain level when time delay $\tau < 10$ and then $\xi(\tau) > 0.5$. For proposed algorithm with the three nonlinear functions, the proposed algorithm with $G_1(u)$ is sensitive to the time delay τ if $\tau > 10$; the influence of the time delay to the proposed algorithm with $G_3(u)$ is less than that with $G_2(u)$. In the rest simulations, we set the time delay $\tau = 1$.

Figure 4 shows the convergence speed of different algorithms. The sources are two 8PSK signals with the same parameters as shown previously and the mixing matrix is randomly generated. It is clear that the convergence speed of EASI is slower than the other algorithm since EASI is a relative gradient-based algorithm with linear convergence speed. FastICA and EBM have a similar convergence speed and the convergence value of FastICA is lower than EBM, which means the residual intersymbol interference of EBM is higher than FastICA. For the proposed algorithm, we can see that its convergence speed is slightly slower than FastICA and EBM. It is because that the proposed algorithm is a natural gradient-based algorithm while FastICA is based on Newton method. In addition, it also can be seen that the convergence speed and convergence values of the proposed algorithm using the three different nonlinear functions are similar, and the convergence values of the proposed algorithm are lower than the two comparison algorithms.

Figure 5 shows the averaged PI values for the separation of two 8PSK sources with different sample size. The mixing matrix is randomly generated in every individual experiment. From Figure 5, we observe that, with the increase of sample

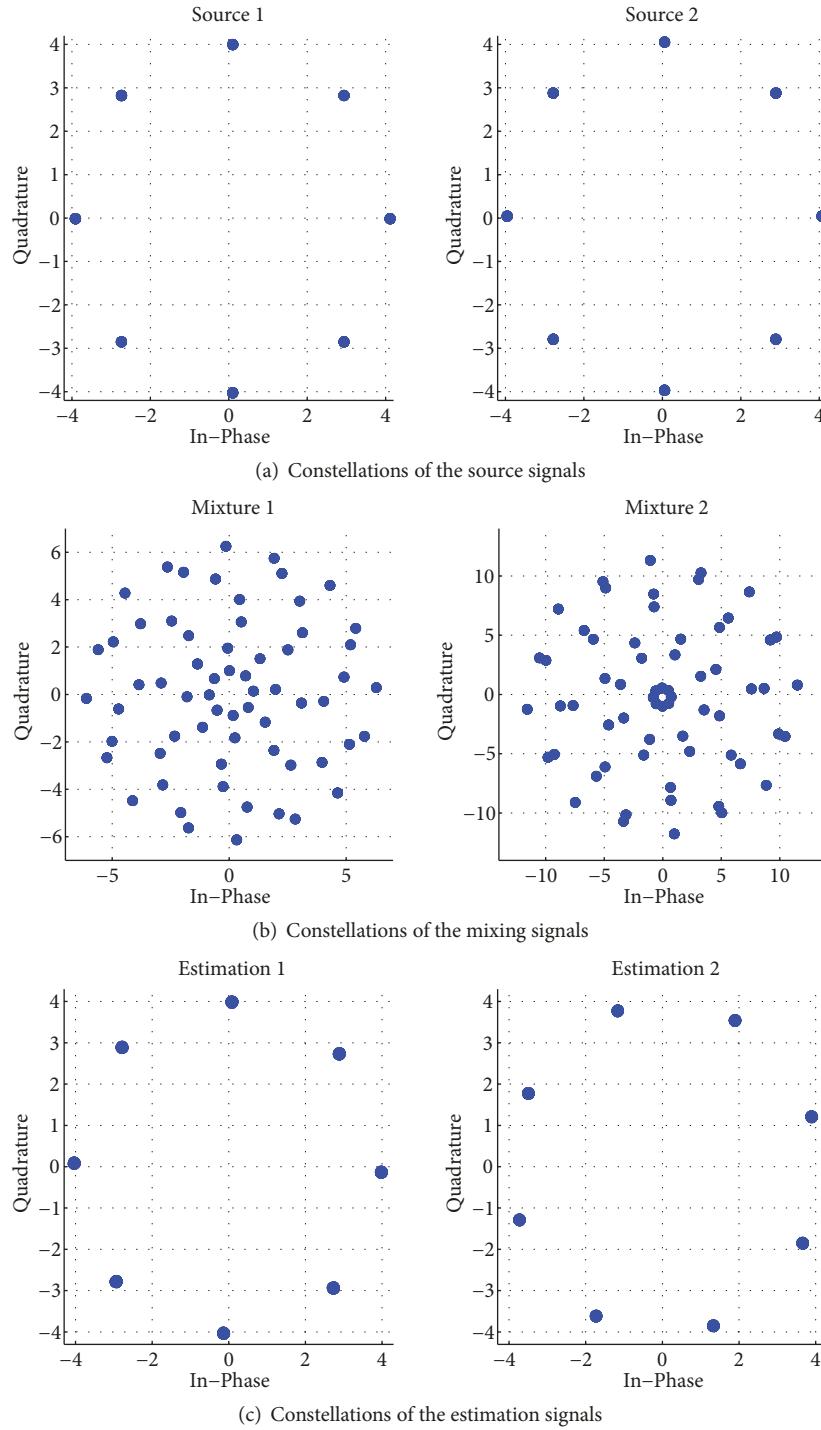


FIGURE 1: Simulation result of the proposed algorithm for mixtures of two 8PSK signals, time delay, $\tau = 1$. From (a) to (c), the source signals 1-2, the mixtures 1-2, and the estimation signals 1-2.

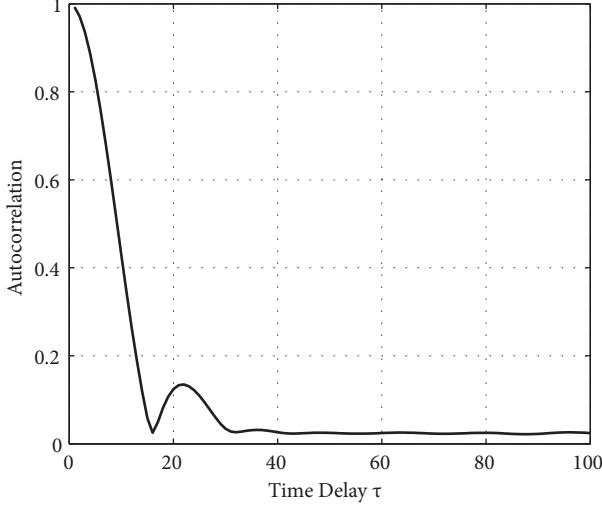


FIGURE 2: Autocorrelation function of the source signal.

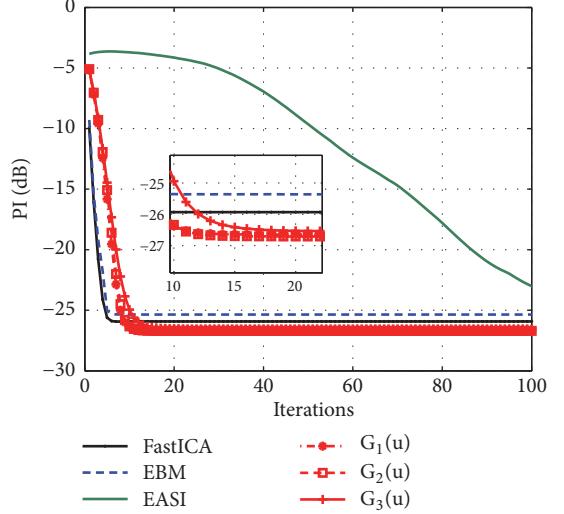


FIGURE 4: Comparison of average PIs over 100 independent runs for the four algorithms.

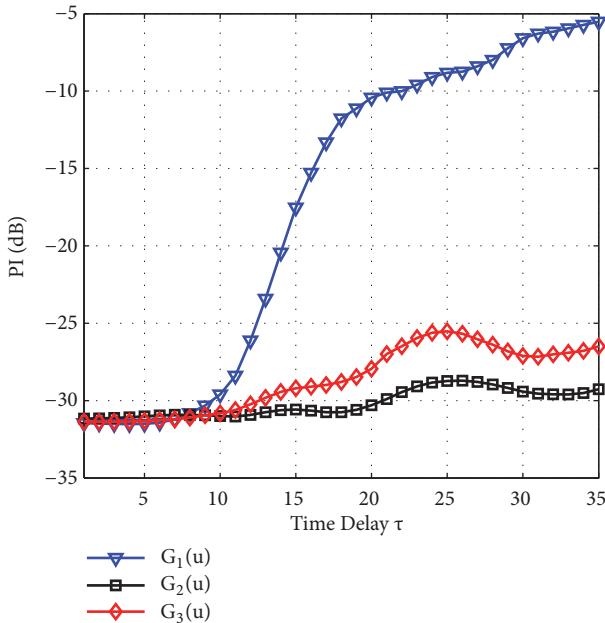
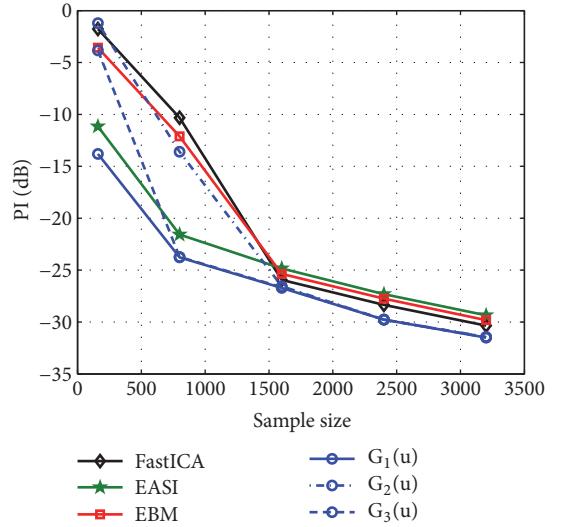
FIGURE 3: Separation performance of the proposed algorithm at different time delay τ .

FIGURE 5: Comparison of average PIs of the four algorithms over 100 independent runs at different sample size for 8PSK signals.

size, the PI values decrease. The proposed algorithm using the nonlinear function $G_1(u)$ is the best algorithm and shows almost the same performance as the proposed algorithm using nonlinear functions $G_2(u)$ and $G_3(u)$ when the sample size is bigger than 1600. FastICA and EBM show a similar performance. The performance of EASI is similar to EBM when sample size is bigger than 1600 even better than FastICA and EBM when sample size is smaller than 1600. However, as shown in Figure 2, the convergence speed of EASI is so slow that it would not be the first choice to perform sources separation.

5. Conclusion

In this paper, we address the complex blind source separation problem using generalized autocorrelations of the sources with temporal structures. A natural gradient-based algorithm is proposed and the stability condition is proved. In the simulations, we perform the proposed algorithm on the mixtures of two 8PSK signals. The simulation results show that although the convergence speed of the proposed algorithm is slightly slower than complex FastICA and EBM algorithms, its convergence value is smaller than the comparison algorithms, which means the averaged signal-to-interference ratio of the estimated signals is higher than the comparison algorithms.

Appendix

Proof of Theorem 1

We make the orthogonal change of coordinates $\mathbf{q} = (\mathbf{AQ})^H \mathbf{w}$ result in the contrast function

$$\nabla J(\mathbf{q}) = \begin{pmatrix} \frac{\partial}{\partial q_{1r}} \\ \frac{\partial}{\partial q_{1i}} \\ \vdots \\ \frac{\partial}{\partial q_{Nr}} \\ \frac{\partial}{\partial q_{Ni}} \end{pmatrix} J(\mathbf{q})$$

$$= 2 \begin{pmatrix} E \left\{ \operatorname{Re} \left\{ s_1 (\mathbf{q}^H \mathbf{s})^* g(|\mathbf{q}^H \mathbf{s}|^2) G(|\mathbf{q}^H \mathbf{s}_\tau|^2) + s_{1\tau} (\mathbf{q}^H \mathbf{s}_\tau)^* G(|\mathbf{q}^H \mathbf{s}|^2) g(|\mathbf{q}^H \mathbf{s}_\tau|^2) \right\} \right\} \\ E \left\{ \operatorname{Im} \left\{ s_1 (\mathbf{q}^H \mathbf{s})^* g(|\mathbf{q}^H \mathbf{s}|^2) G(|\mathbf{q}^H \mathbf{s}_\tau|^2) + s_{1\tau} (\mathbf{q}^H \mathbf{s}_\tau)^* G(|\mathbf{q}^H \mathbf{s}|^2) g(|\mathbf{q}^H \mathbf{s}_\tau|^2) \right\} \right\} \\ \vdots \\ E \left\{ \operatorname{Re} \left\{ s_N (\mathbf{q}^H \mathbf{s})^* g(|\mathbf{q}^H \mathbf{s}|^2) G(|\mathbf{q}^H \mathbf{s}_\tau|^2) + s_{N\tau} (\mathbf{q}^H \mathbf{s}_\tau)^* G(|\mathbf{q}^H \mathbf{s}|^2) g(|\mathbf{q}^H \mathbf{s}_\tau|^2) \right\} \right\} \\ E \left\{ \operatorname{Im} \left\{ s_N (\mathbf{q}^H \mathbf{s})^* g(|\mathbf{q}^H \mathbf{s}|^2) G(|\mathbf{q}^H \mathbf{s}_\tau|^2) + s_{N\tau} (\mathbf{q}^H \mathbf{s}_\tau)^* G(|\mathbf{q}^H \mathbf{s}|^2) g(|\mathbf{q}^H \mathbf{s}_\tau|^2) \right\} \right\} \end{pmatrix} \quad (\text{A.1})$$

where $q_j = q_{jr} + iq_{ji}$. The Hessian matrix of $J(\mathbf{q})$ is now a $2N \times 2N$ real matrix. Denote

$$J_{Rn} = E \left\{ \operatorname{Re} \left\{ s_1 (\mathbf{q}^H \mathbf{s})^* g(|\mathbf{q}^H \mathbf{s}|^2) G(|\mathbf{q}^H \mathbf{s}_\tau|^2) + s_{1\tau} (\mathbf{q}^H \mathbf{s}_\tau)^* G(|\mathbf{q}^H \mathbf{s}|^2) g(|\mathbf{q}^H \mathbf{s}_\tau|^2) \right\} \right\} \quad (\text{A.2})$$

$$+ s_{1\tau} (\mathbf{q}^H \mathbf{s}_\tau)^* G(|\mathbf{q}^H \mathbf{s}|^2) g(|\mathbf{q}^H \mathbf{s}_\tau|^2) \}$$

$$J_{In} = E \left\{ \operatorname{Im} \left\{ s_1 (\mathbf{q}^H \mathbf{s})^* g(|\mathbf{q}^H \mathbf{s}|^2) G(|\mathbf{q}^H \mathbf{s}_\tau|^2) + s_{1\tau} (\mathbf{q}^H \mathbf{s}_\tau)^* G(|\mathbf{q}^H \mathbf{s}|^2) g(|\mathbf{q}^H \mathbf{s}_\tau|^2) \right\} \right\} \quad (\text{A.3})$$

$$+ s_{1\tau} (\mathbf{q}^H \mathbf{s}_\tau)^* G(|\mathbf{q}^H \mathbf{s}|^2) g(|\mathbf{q}^H \mathbf{s}_\tau|^2) \}$$

Hence the Hessian matrix of $J(\mathbf{q})$ is

$J(\mathbf{q}) = E\{G(|\mathbf{q}^H \mathbf{s}|^2)G(|\mathbf{q}^H \mathbf{s}_\tau|^2)\}$. We seek a Taylor series expansion of $J(\mathbf{q})$ in the extrema. Since $J(\mathbf{q})$ is in general not analytic and thus it cannot be expanded as a Taylor series in the complex form, we do not use complex differentiation operators. The gradient of $J(\mathbf{q})$ with respect to \mathbf{q} is

$$\nabla^2 J(\mathbf{q}) = \begin{pmatrix} \frac{\partial J_{R1}}{\partial q_{1r}} & \frac{\partial J_{R1}}{\partial q_{1i}} & \dots & \frac{\partial J_{R1}}{\partial q_{Nr}} & \frac{\partial J_{R1}}{\partial q_{Ni}} \\ \frac{\partial J_{I1}}{\partial q_{1r}} & \frac{\partial J_{I1}}{\partial q_{1i}} & \dots & \frac{\partial J_{I1}}{\partial q_{Nr}} & \frac{\partial J_{I1}}{\partial q_{Ni}} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{\partial J_{RN}}{\partial q_{1r}} & \frac{\partial J_{RN}}{\partial q_{1i}} & \dots & \frac{\partial J_{RN}}{\partial q_{Nr}} & \frac{\partial J_{RN}}{\partial q_{Ni}} \\ \frac{\partial J_{IN}}{\partial q_{1r}} & \frac{\partial J_{IN}}{\partial q_{1i}} & \dots & \frac{\partial J_{IN}}{\partial q_{Nr}} & \frac{\partial J_{IN}}{\partial q_{Ni}} \end{pmatrix} \quad (\text{A.4})$$

Without loss of generality, we assume that the optimal solution for s_1 is at $\mathbf{q}_1 = q\mathbf{e}_1 = [q, 0, \dots, 0]^T$. Now

$q = q_r + iq_i$ and $|\mathbf{q}^H \mathbf{s}|^2 = |s_1|^2$. We now search for a Taylor series expansion of the contrast function $J(\mathbf{q})$ around the

optimal solution \mathbf{q}_1 . Evaluating gradient equation (A.1) at point $\mathbf{q}_1 = q\mathbf{e}_1$, we get

$$\nabla J(\mathbf{q}_1) = 2 \begin{pmatrix} q_r E \{ |s_1|^2 g(|s_1|^2) G(|s_{1\tau}|^2) + |s_{1\tau}|^2 G(|s_1|^2) g(|s_{1\tau}|^2) \} \\ q_i E \{ |s_1|^2 g(|s_1|^2) G(|s_{1\tau}|^2) + |s_{1\tau}|^2 G(|s_1|^2) g(|s_{1\tau}|^2) \} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad (\text{A.5})$$

The hessian matrix at point $\mathbf{q}_1 = q\mathbf{e}_1$ is

$$\nabla^2 J(\mathbf{q}_1) = \begin{pmatrix} \alpha_1 + q_r^2 \alpha_2 & q_r q_i \alpha_2 & 0 & 0 & 0 & 0 & 0 \\ q_r q_i \alpha_2 & \alpha_1 + q_i^2 \alpha_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \alpha_3 + \beta_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \alpha_3 + \beta_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \alpha_3 + \beta_1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & \alpha_3 + \beta_2 \end{pmatrix} \quad (\text{A.6})$$

where

$$\alpha_1 = E \{ |s_1|^2 g(|s_1|^2) G(|s_{1\tau}|^2) + |s_{1\tau}|^2 G(|s_1|^2) g(|s_{1\tau}|^2) \} \quad (\text{A.7})$$

$$\alpha_2 = 2E \{ |s_1|^4 g'(|s_1|^2) G(|s_{1\tau}|^2) + 2|s_1|^2 |s_{1\tau}|^2 \\ \cdot g(|s_1|^2) g(|s_{1\tau}|^2) + |s_{1\tau}|^4 G(|s_1|^2) g'(|s_{1\tau}|^2) \} \quad (\text{A.8})$$

$$\alpha_3 = E \{ g(|s_1|^2) G(|s_{1\tau}|^2) + G(|s_1|^2) g(|s_{1\tau}|^2) \\ + |s_1|^2 g'(|s_1|^2) G(|s_{1\tau}|^2) + |s_{1\tau}|^2 G(|s_1|^2) \\ \cdot g'(|s_{1\tau}|^2) \} \quad (\text{A.9})$$

$$\beta_1 = 2E \{ \operatorname{Re} \{ s_n s_{n\tau} s_1^* s_{1\tau}^* + s_n s_{n\tau}^* s_1^* s_{1\tau} \} g(|s_1|^2) \\ \cdot g(|s_{1\tau}|^2) \} \quad (\text{A.10})$$

$$\beta_2 = 2E \{ \operatorname{Re} \{ s_n s_{n\tau}^* s_1^* s_{1\tau} - s_n s_{n\tau} s_1^* s_{1\tau}^* \} g(|s_1|^2) \\ \cdot g(|s_{1\tau}|^2) \} \quad (\text{A.11})$$

We now make a small perturbation $\varepsilon = [\varepsilon_{1r}, \varepsilon_{1i}, \dots, \varepsilon_{Nr}, \varepsilon_{Ni}]^T$, where ε_{nr} and ε_{ni} are the real and imaginary parts of $\varepsilon_n \in \mathbb{C}$ and the Taylor expansion of $J(\mathbf{q}_1 + \varepsilon)$ at point \mathbf{q}_1 is

$$\begin{aligned} J(\mathbf{q}_1 + \varepsilon) &= J(\mathbf{q}_1) + \varepsilon^T \nabla J(\mathbf{q}_1) + \varepsilon^T \nabla^2 J(\mathbf{q}_1) \varepsilon \\ &\quad + o(\|\varepsilon\|^2) = J(\mathbf{q}_1) + 2(\varepsilon_{1r} q_r + \varepsilon_{1i} q_i) E\{\alpha_1\} \\ &\quad + (\varepsilon_{1r}^2 + \varepsilon_{1i}^2) E\{\alpha_1\} + (\varepsilon_{1r} q_r + \varepsilon_{1i} q_i)^2 E\{\alpha_2\} \\ &\quad + \sum_{n=2}^N 2(\varepsilon_{nr}^2 + \varepsilon_{ni}^2) \\ &\quad \cdot E\{\operatorname{Re}\{s_n s_{n\tau} s_1^* s_{1\tau}^*\} g(|s_1|^2) g(|s_{1\tau}|^2)\} \\ &\quad + \sum_{n=2}^N 2(\varepsilon_{nr}^2 - \varepsilon_{ni}^2) \\ &\quad \cdot E\{\operatorname{Re}\{s_n s_{n\tau} s_1^* s_{1\tau}^*\} g(|s_1|^2) g(|s_{1\tau}|^2)\} \\ &\quad + o(\|\varepsilon\|^2) \end{aligned} \quad (\text{A.12})$$

Furthermore, due to the constraint $\|\mathbf{w}_i\| = 1$ and thus $\|\mathbf{q}_1 + \varepsilon\| = 1$, we get

$$2(\varepsilon_{1r} q_r + \varepsilon_{1i} q_i) = - \sum_{n=1}^N (\varepsilon_{nr}^2 + \varepsilon_{ni}^2) \quad (\text{A.13})$$

Substituting (A.13) into (A.12), we get

$$\begin{aligned}
J(\mathbf{q}_1 + \boldsymbol{\varepsilon}) &= J(\mathbf{q}_1) + \sum_{n=2}^N (\varepsilon_{nr}^2 + \varepsilon_{ni}^2) E \{ \alpha_3 - \alpha_1 \\
&\quad + 2 \operatorname{Re} \{ s_n s_{n\tau}^* s_1^* s_{1\tau} \} g(|s_1|^2) g(|s_{1\tau}|^2) \} \\
&\quad + \sum_{n=2}^N 2 (\varepsilon_{nr}^2 - \varepsilon_{ni}^2) \\
&\quad \cdot E \{ \operatorname{Re} \{ s_n s_{n\tau} s_1^* s_{1\tau}^* \} g(|s_1|^2) g(|s_{1\tau}|^2) \} \\
&\quad + o(\|\boldsymbol{\varepsilon}\|^2)
\end{aligned} \tag{A.14}$$

Due to the fact that $|\varepsilon_{nr}^2 + \varepsilon_{ni}^2| > |\varepsilon_{nr}^2 - \varepsilon_{ni}^2|$, (A.14) is closer to (A.15) than (A.16), that is to say, if $J(\mathbf{q}_1)$ is an extrema, (A.14) is closer to the extrema than (A.16).

$$\begin{aligned}
J_1(\mathbf{q}_1 + \boldsymbol{\varepsilon}) &= J(\mathbf{q}_1) + \sum_{n=2}^N (\varepsilon_{nr}^2 + \varepsilon_{ni}^2) E \{ \alpha_3 - \alpha_1 \\
&\quad + 2 \operatorname{Re} \{ s_n s_{n\tau}^* s_1^* s_{1\tau} \} g(|s_1|^2) g(|s_{1\tau}|^2) \} \\
&\quad + o(\|\boldsymbol{\varepsilon}\|^2)
\end{aligned} \tag{A.15}$$

$$\begin{aligned}
J_2(\mathbf{q}_1 + \boldsymbol{\varepsilon}) &= J(\mathbf{q}_1) + \sum_{n=2}^N (\varepsilon_{nr}^2 + \varepsilon_{ni}^2) E \{ \alpha_3 - \alpha_1 + \beta_1 \} \\
&\quad + o(\|\boldsymbol{\varepsilon}\|^2)
\end{aligned} \tag{A.16}$$

Obviously, \mathbf{q}_1 is an extrema of (A.16), and it is the maximum (minimum) if

$$\begin{aligned}
E \{ \alpha_3 - \alpha_1 + 2 \operatorname{Re} \{ s_n s_{n\tau} s_1^* s_{1\tau}^* + s_n s_{n\tau}^* s_1^* s_{1\tau} \} g(|s_1|^2) \\
\cdot g(|s_{1\tau}|^2) \} < 0 \quad (\text{resp. } > 0)
\end{aligned} \tag{A.17}$$

\mathbf{q}_1 is also extrema of (A.14) and $J(\mathbf{q}_1)$ is the maximum (minimum) if (A.17) can be satisfied.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] Q. C. Li, H. Niu, A. T. Papathanassiou, and G. Wu, “5G network capacity: key elements and technologies,” *IEEE Vehicular Technology Magazine*, vol. 9, no. 1, pp. 71–78, 2014.
- [2] L. Wang, F. Tian, T. Svensson, D. Feng, M. Song, and S. Li, “Exploiting full duplex for device-to-device communications in heterogeneous networks,” *IEEE Communications Magazine*, vol. 53, no. 5, pp. 146–152, 2015.
- [3] D. Korpi, J. Tammisen, M. Turunen et al., “Full-duplex mobile device: Pushing the limits,” *IEEE Communications Magazine*, vol. 54, no. 9, pp. 80–87, 2016.
- [4] Z. Zhang, X. Chai, K. Long, A. V. Vasilakos, and L. Hanzo, “Full duplex techniques for 5G networks: self-interference cancellation, protocol design, and relay selection,” *IEEE Communications Magazine*, vol. 53, no. 5, pp. 128–137, 2015.
- [5] J. Li, H. Zhang, and M. Fan, “Digital self-interference cancellation based on independent component analysis for co-time co-frequency full-duplex communication systems,” *IEEE Access*, vol. 5, pp. 10222–10231, 2017.
- [6] C. Chenot and J. Bobin, “Blind separation of sparse sources in the presence of outliers,” *Signal Processing*, vol. 138, pp. 233–243, 2017.
- [7] Y. Duan and H. Zhang, “Noisy blind signal-jamming separation algorithm based on VBICA,” *Wireless Personal Communications*, vol. 74, no. 2, pp. 307–324, 2014.
- [8] Z. Cui, H. Zhang, F. Gu, and J. Zhang, “Blind multiuser detection and external interference cancellation of asynchronous DS-CDMA system with single antenna,” *Information*, vol. 15, no. 10, pp. 4261–4276, 2012.
- [9] H. Yang, H. Zhang, J. Li, L. Yang, and W. Ding, “Baseband communication signal blind separation algorithm based on complex nonparametric probability density estimation,” *IEEE Access*, vol. 6, pp. 22434–22440, 2018.
- [10] S. Gannot, E. Vincent, S. Markovich-Golan, and A. Ozerov, “A Consolidated Perspective on Multimicrophone Speech Enhancement and Source Separation,” *IEEE/ACM Transactions on Audio Speech and Language Processing*, vol. 25, no. 4, pp. 692–730, 2017.
- [11] L. Wang, “Multi-band multi-centroid clustering based permutation alignment for frequency-domain blind speech separation,” *Digital Signal Processing*, vol. 31, no. 5, pp. 79–92, 2014.
- [12] T. Adali, Y. Levin-Schwartz, and V. D. Calhoun, “Multimodal Data Fusion Using Source Separation: Application to Medical Imaging,” *Proceedings of the IEEE*, vol. 103, no. 9, pp. 1494–1506, 2015.
- [13] R. Chai, G. R. Naik, T. N. Nguyen et al., “Driver fatigue classification with independent component by entropy rate bound minimization analysis in an eeg-based system,” *IEEE Journal of Biomedical Health Informatics*, vol. 21, no. 3, pp. 715–724, 2017.
- [14] G. R. Naik, K. G. Baker, and H. T. Nguyen, “Dependence independence measure for posterior and anterior EMG sensors used in simple and complex finger flexion movements: Evaluation using SDICA,” *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 5, pp. 1689–1696, 2014.
- [15] J.-F. Cardoso, “Source separation using higher order moments,” in *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, vol. 4, pp. 2109–2112, May 1989.
- [16] J. F. Cardoso and A. Souloumiac, “Blind beamforming for non-Gaussian signals,” *IEE Proceedings F Radar and Signal Processing*, vol. 140, no. 6, pp. 362–370, 1993.
- [17] J. Eriksson and V. Koivunen, “Complex-valued ICA using second order statistics,” in *Proceedings of the IEEE Workshop on Machine Learning for Signal Processing*, pp. 183–192, São Luís, Brazil, October 2004.
- [18] E. Ollila and V. Koivunen, “Complex ICA using generalized uncorrelating transform,” *Signal Processing*, vol. 89, no. 4, pp. 365–377, 2009.
- [19] Z. Shi and C. Zhang, “Blind source extraction using generalized autocorrelations,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 18, no. 5, pp. 1516–1524, 2007.

- [20] E. Bingham and A. Hyvärinen, “A fast fixed-point algorithm for independent component analysis of complex valued signals,” *International Journal of Neural Systems*, vol. 10, no. 1, pp. 1–8, 2000.
- [21] X.-L. Li and T. Adali, “Complex independent component analysis by entropy bound minimization,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 57, no. 7, pp. 1417–1430, 2010.
- [22] J. M. Ye, H. H. Jin, S. T. Lou, and K. J. You, “An optimized EASI algorithm,” *Signal Processing*, vol. 89, no. 3, pp. 333–338, 2009.

