

A Fast Algorithm for Blind Separation of Complex Valued Signals with Nonlinear Autocorrelation

Pengcheng Xu and Yuehong Shen

College of Communications Engineering, PLA University of Science and Technology, 210007, Nanjing, P.R. China
Email: pchxu2005@gmail.com, gmajjie03@163.com

Abstract—Blind source separation of complex valued signals has been a hot issue especially in the field of multi-input/multi-output (MIMO) digital communications. Many contrast functions based on the nonlinear structure of the signals have been proposed to extract the unknown sources. However, these researches usually focused on the real-valued case, but ignoring the complex problem. This paper proposes a novel algorithm based on Newton iterations to solve the complex-valued case. The method has a potential capability of extracting complex sources with nonlinear autocorrelation. We also analyze the convergence conditions of the algorithm in theory. Numerical simulations for artificial complex signals corroborate the efficiency of the proposed method. Moreover, our algorithm performs more robust with lower computational cost than classical cumulant-based approach using the nonstationarity of variance (CANSV). Finally, experiments for the separation of single sideband signals illustrate that our method might have good prospects in real-world applications.

Index Terms—Blind source separation, complex valued signals, nonlinear autocorrelation, convergence conditions.

I. INTRODUCTION

Blind source separation (BSS), an active area of research, aims to recover original sources from their mixtures with minimal priors and has found wide applicability. In particular, blind source separation of complex valued signals has found utility in many applications such as digital communications [1]-[3], artificial intelligence [4], analysis of functional magnetic resonance imaging [5], and radar data [6], [7]. In these applications, the signals are usually complex-valued with non-Gaussian or Gaussian distribution.

Independent component analysis (ICA), as a kind of successful algorithms in separation [8], [9], generally assume that all sources are non-Gaussian or at most one is Gaussian, dealing with the separation problem by maximizing the non-Gaussianity. Therefore, most conventional ICA algorithms like Fast ICA and

Infomax are incapable to provide the desired source separation when the assumed sources are all Gaussian distributed. Alternative methods that explore other statistical properties instead of non-Gaussianity have been proposed, such as linear predictability or smoothness [10], [11], linear autocorrelation [12]-[15], temporal predictability [16], etc. These algorithms mainly based on linear statistical properties are known to provide good results for the separation of Gaussian distributed sources. Recently, deeper latent statistical properties of sources have been developed, like nonstationarity [17]-[19], nonlinear autocorrelation [20], nonlinear predictability [21] and so on. The nonlinear statistical properties are more common in biomedicine [22], communications [18], and speech areas [10]. Nevertheless, it is a pity that these studies, only considering the real-valued case, can not be directly applied to solving the complex-valued problem. In this context, we investigated the separation of complex sources with nonlinear autocorrelation in our previous work [23]. The gradient algorithm has been proposed, and the stability conditions about the contrasts in complex case have been derived.

In this paper, we extend our previous work into a novel algorithm based on Newton iterations to solve complex-valued BSS problem. The learning algorithm is implemented in second-order iterative, therefore the convergence condition of the proposed algorithm is analyzed in theory, which is valuable to choose proper nonlinear functions. When sources have square temporal autocorrelation, we demonstrate the efficiency of the method. Furthermore, the method performs more robust at lower computational time than classical CANSV method [17]. Finally, experiments on the separation of single sideband signals illustrate that our method may have good prospects in real-world applications. What's more, this work can be seemed as an important complement of the previous work.

The remainder of the paper is organized as follows. Section II introduces the contrast function based on nonlinear autocorrelation and the optimization algorithm. We analyze the convergence condition of the proposed method in Sect. III. Section IV presents simulation results on different datasets and discussion. Finally, Sect. V concludes the paper. In this article, the superscripts $()^*$, $()^T$, $()^H$ denote conjugate, transpose and conjugate transpose, respectively. Bold upper

Manuscript received November 24, 2014; revised March 25, 2015.

This work was supported by the National Natural Science Foundation of China under Grant No. 61172061 and the Natural Science Foundation of JiangSu Province in China under Grant No. BK2011117.

Corresponding author email: pchxu2005@gmail.com.
doi:10.12720/jcm.10.3.170-177

(respectively, lower) case letters are used for matrices (respectively, vectors).

II. PROPOSED ALGORITHM

A. Contrast Function

In the noise-free instantaneous case, we assume that n unknown statistically independent source signals $\mathbf{s}(t) = [s_1(t), \dots, s_n(t)]^T \in \mathbb{C}^{n \times 1}$ of zero mean pass through an unknown full-column rank mixing matrix $\mathbf{A} \in \mathbb{C}^{m \times n}$ ($m \geq n$), therefore m mixed signals $\mathbf{x}(t) = [x_1(t), \dots, x_m(t)]^T$ can be modeled as

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

where t is the sample index. Besides, the sources are assumed to have certain unknown temporal structure with nonlinear autocorrelation, i.e., $|\mathbb{E}\{I(s_i(t))I(s_i(t-\tau)^*)\}|^2 \neq 0$ for $\forall i$, where τ and $I(\cdot)$ denotes respectively certain time lags and an unknown nonlinear function that characterize the nonlinear autocorrelation.

To simplify the problem, we further assume that the number of sources matches the number of mixtures i.e. $m=n$, an exactly determined problem. The prewhitening of the observations by the matrix \mathbf{U} yields the new observations vector, $\mathbf{z}(t) = \mathbf{U}\mathbf{x}(t)$, whose covariance $\mathbb{E}[\mathbf{z}\mathbf{z}^H]$ is the identity matrix of dimension $n \times n$. Then the output signal $y(t)$ that estimates one of the sources is obtained as $y(t) = \mathbf{w}^H \mathbf{z}(t)$, where \mathbf{w} is the unit norm extraction column vector (enforcing the output to have unit variance). To estimate all source signals $\mathbf{y}(t) = [y_1(t), \dots, y_n(t)]^T$, we should find an $n \times n$ separating matrix $\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_n]$.

Based on nonlinear autocorrelation of the desired complex-valued source, we present the following contrast function in complex case [23]

$$J(\mathbf{w}) = \left| \mathbb{E}\{G(y(t))G(y(t-\tau)^*)\} \right|^2 \quad (2)$$

where $G(\cdot)$ is the nonlinear function. The contrast function measures the nonlinear autocorrelation of the desired source. Most of the time, the contrast function though using only one time lagged nonlinear autocorrelation (the delay τ usually equals to 1) can obtain satisfying enough source separation performance. The nonlinear function choices are important for our method to obtain a stable separation point, further analysis of which is provided in Section 3. Examples of proper nonlinear functions are $G(u) = u^2$ and $G(u) = \log(\cosh(u))$. Hence, the solution of BSS for complex sources relies on the maximization of the contrast function.

B. Learning Algorithm

Basically, in order to maximize the contrast function (2), we can introduce a steepest ascent algorithm which moves from one point to another following the gradient direction of the criterion. However, improper choosing of step-size in gradient algorithm usually leads to slow convergence speed and poor stability performance [24]. Therefore, we derive a fast fixed-point algorithm through the Newton iteration, which is similar to the Complex FastICA [25] in ICA. Then we have

$$\mathbf{w} \leftarrow \mathbf{w} - \left[\frac{\partial}{\partial \mathbf{w}^T} \left(\frac{\partial J(\mathbf{w})}{\partial \mathbf{w}^*} \right) \right]^{-1} \frac{\partial J(\mathbf{w})}{\partial \mathbf{w}^*} \quad (3)$$

$$\mathbf{w} \leftarrow \mathbf{w} / \|\mathbf{w}\| \quad (4)$$

To derive the conjugate gradient of $J(\mathbf{w})$, we write $J(\mathbf{w})$ as

$$J(\mathbf{w}) = |J(\mathbf{w})|^2 \quad (5)$$

where $J(\mathbf{w}) = \mathbb{E}\{G(y(t))G(y(t-\tau)^*)\}$.

Then the partial derivative of $J(\mathbf{w})$ with respect to the conjugate of \mathbf{w} can be obtained as (see [26] for more details)

$$\begin{aligned} \frac{\partial J(\mathbf{w})}{\partial \mathbf{w}^*} &= \left(\frac{\partial J(\mathbf{w})}{\partial \mathbf{w}} \right)^* = \left(2J(\mathbf{w}) \frac{\partial J(\mathbf{w})}{\partial \mathbf{w}} \right)^* \\ &= \mathbb{E}\{G(y(t))^* G(y(t-\tau)^*)\} \times \\ &\quad \mathbb{E}\{G(y(t))^* g(y(t-\tau)^*)^* \mathbf{z}(t-\tau)\} \end{aligned} \quad (6)$$

where $g(\cdot)$ is the derivative of $G(\cdot)$. The Hessian matrix of $J(\mathbf{w})$ with respect to \mathbf{w} can be obtained as

$$\begin{aligned} \frac{\partial}{\partial \mathbf{w}^T} \left(\frac{\partial J(\mathbf{w})}{\partial \mathbf{w}^*} \right) &= \mathbb{E}\{g(y(t))^* G(y(t-\tau)^*)^* (\mathbf{z}(t))^H\} \times \\ &\quad \mathbb{E}\{G(y(t))^* g(y(t-\tau)^*)^* \mathbf{z}(t-\tau)\} + \\ &\quad \mathbb{E}\{G(y(t))^* G(y(t-\tau)^*)^*\} \times \\ &\quad \mathbb{E}\{g(y(t))^* g(y(t-\tau)^*)^* \mathbf{z}(t-\tau)(\mathbf{z}(t))^H\} \end{aligned} \quad (7)$$

To simplify the inversion of this matrix, we decide to approximate the first term in (7). An approximation that turns out to be very suitable here is

$$\begin{aligned} &\mathbb{E}\{g(y(t))^* G(y(t-\tau)^*)^* (\mathbf{z}(t))^H\} \times \\ &\quad \mathbb{E}\{G(y(t))^* g(y(t-\tau)^*)^* \mathbf{z}(t-\tau)\} \\ &\approx \mathbb{E}\{g(y(t))^* G(y(t-\tau)^*)^*\} \mathbb{E}\{(\mathbf{z}(t))^H\} \times \end{aligned} \quad (8)$$

$$\begin{aligned} &\mathbb{E}\{G(y(t))^* g(y(t-\tau)^*)^*\} \mathbb{E}\{\mathbf{z}(t-\tau)\} \\ &= 0 \end{aligned}$$

Thus, the formula (3) becomes

$$\mathbf{w} \leftarrow \mathbf{w} - \left(\mathbb{E} \left\{ \mathbf{g}(y(t))^* \mathbf{g}(y(t-\tau))^* \mathbf{z}(t-\tau) (\mathbf{z}(t))^H \right\} \right)^{-1} \times \mathbb{E} \left\{ \mathbf{G}(y(t))^* \mathbf{g}(y(t-\tau))^* \mathbf{z}(t-\tau) \right\} \quad (9)$$

As a result, formula (9) can be further simplified by multiplying both sides of (9) by

$$\mathbb{E} \left\{ \mathbf{g}(y(t))^* \mathbf{g}(y(t-\tau))^* \mathbf{z}(t-\tau) (\mathbf{z}(t))^H \right\} \\ \approx \mathbb{E} \left\{ \mathbf{g}(y(t))^* \mathbf{g}(y(t-\tau))^* \right\} \mathbb{E} \left\{ \mathbf{z}(t-\tau) (\mathbf{z}(t))^H \right\}$$

This leads to the following fixed-point algorithm:

$$\mathbf{w} \leftarrow \mathbb{E} \left\{ \mathbf{G}(y(t))^* \mathbf{g}(y(t-\tau))^* \mathbf{z}(t-\tau) \right\} \\ - \mathbb{E} \left\{ \mathbf{g}(y(t))^* \mathbf{g}(y(t-\tau))^* \right\} \mathbb{E} \left\{ \mathbf{z}(t-\tau) (\mathbf{z}(t))^H \right\} \mathbf{w} \quad (10) \\ \mathbf{w} \leftarrow \mathbf{w} / \|\mathbf{w}\|$$

Hence, our algorithm (C-FastNA: Complex Fast Algorithm based on Nonlinear Autocorrelation) can be summarized as follows:

C-FastNA Algorithm (estimating all sources using symmetric orthogonalization [25])

1) Initialization

Center and whiten the observations. Set random initial matrix $\mathbf{W}^{(0)} = [\mathbf{w}_1^{(0)}, \dots, \mathbf{w}_n^{(0)}]$, and choose proper nonlinear functions and delays τ (equals to 1).

2) Iteration

At the i th iteration for obtaining \mathbf{W} ,

For $p = 1, \dots, n$

$$\mathbf{w}_p^{(i)} \leftarrow \mathbb{E} \left\{ \mathbf{G} \left((\mathbf{w}_p^{(i-1)})^H \mathbf{z}(t) \right)^* \mathbf{g} \left((\mathbf{z}(t-\tau_k))^H \mathbf{w}_p^{(i-1)} \right)^* \mathbf{z}(t-\tau) \right\} \\ - \mathbb{E} \left\{ \mathbf{g} \left((\mathbf{w}_p^{(i-1)})^H \mathbf{z}(t) \right)^* \mathbf{g} \left((\mathbf{z}(t-\tau_k))^H \mathbf{w}_p^{(i-1)} \right)^* \right\} \\ \mathbb{E} \left\{ \mathbf{z}(t-\tau) (\mathbf{z}(t))^H \right\} \mathbf{w}_p^{(i-1)} \quad (11)$$

End for

$$\mathbf{W}^{(i)} \leftarrow \mathbf{W}^{(i)} \left((\mathbf{W}^{(i)})^H \mathbf{W}^{(i)} \right)^{-1/2}$$

3) Termination

The iteration is terminated when the relative change $\|\mathbf{W}^{(i)} - \mathbf{W}^{(i-1)}\|$ is less than a specified tolerance.

Furthermore, one can also use a deflation scheme (one-by-one estimation) to estimate all the source signals [27].

III. CONVERGENCE ANALYSIS

In this section, we derive the convergence properties of the proposed algorithm (10). For the simplicity of expression, $s_{i,\tau}$ denotes $s_i(t-\tau)$. We make the orthogonal change of coordinates $\mathbf{p} = (p_1, \dots, p_n)^T =$

$\mathbf{A}^H \mathbf{U}^H \mathbf{w}$ so that $y = \mathbf{w}^H \mathbf{x} = \mathbf{p}^H \mathbf{s}$. Without loss of generality, we analyze the convergence of the point $\mathbf{p}_1 = (e^{j\theta}, \dots, 0)^T$ where θ is an arbitrary phase shift. The following theorem tells us in which conditions the proposed algorithm can converge.

Theorem. If the following conditions are satisfied: (a) $\{s_i, s_{i,\tau}\}$ and $\{s_j, s_{j,\tau}\}$ are mutually independent ($\forall j \neq i$); (b) sources have no temporally linear autocorrelation, i.e., $\mathbb{E}\{s_i s_{i,\tau}^*\} = 0, \forall i$; (c) all the sources are second-order circular, i.e., $\mathbb{E}\{s_i^2\} = 0, \forall i$; (d) $\mathbb{E}\left\{ \mathbf{G}(e^{-j\theta} s_i)^* \mathbf{g}(e^{j\theta} s_{i,\tau}^*)^* s_{i,\tau} \right\} \neq 0$, the fast algorithm (10) converges. That is to say, the vector \mathbf{w} converges, up to the phase ambiguity, to one row of the inverse of the mixing matrix $\mathbf{U}\mathbf{A}$, and the convergence speed is at least quadratic.

Proof. To begin with, we can rewrite the fast algorithm (10) as

$$\mathbf{p} \leftarrow \mathbb{E} \left\{ \mathbf{G}(\mathbf{p}^H \mathbf{s})^* \mathbf{g}(\mathbf{s}_\tau^H \mathbf{p})^* \mathbf{s}_\tau \right\} - \\ \mathbb{E} \left\{ \mathbf{g}(\mathbf{p}^H \mathbf{s})^* \mathbf{g}(\mathbf{s}_\tau^H \mathbf{p})^* \right\} \mathbb{E} \left\{ \mathbf{s}_\tau \mathbf{s}_\tau^H \right\} \mathbf{p} \quad (12)$$

$$\mathbf{p} \leftarrow \mathbf{p} / \|\mathbf{p}\|$$

Using the complex Taylor series expansion for \mathbf{G} and \mathbf{g} , we obtain

$$\mathbf{G}(\mathbf{p}^H \mathbf{s}) = \mathbf{G}(p_1^* s_1) + \mathbf{g}(p_1^* s_1) \mathbf{p}_{-1}^H \mathbf{s}_{-1} + \\ \frac{1}{2} \mathbf{g}'(p_1^* s_1) (\mathbf{p}_{-1}^H \mathbf{s}_{-1})^2 + o(\|\mathbf{p}_{-1}\|^2) \quad (13)$$

$$\mathbf{g}(\mathbf{p}^H \mathbf{s}) = \mathbf{g}(p_1^* s_1) + \mathbf{g}'(p_1^* s_1) \mathbf{p}_{-1}^H \mathbf{s}_{-1} + \\ \frac{1}{2} \mathbf{g}''(p_1^* s_1) (\mathbf{p}_{-1}^H \mathbf{s}_{-1})^2 + o(\|\mathbf{p}_{-1}\|^2) \quad (14)$$

where \mathbf{p}_{-1} and \mathbf{s}_{-1} are the vectors \mathbf{p} and \mathbf{s} without their first components. Using assumptions (a) in Theorem, also doing some algebraic manipulations, we have

$$p_1 = \mathbb{E} \left\{ \mathbf{G}(e^{-j\theta} s_1)^* \mathbf{g}(e^{j\theta} s_{1,\tau}^*)^* s_{1,\tau} \right\} + o(\|\mathbf{p}_{-1}\|^2) \quad (15)$$

$$p_i = \mathbb{E} \left\{ \mathbf{G}(e^{-j\theta} s_1)^* \mathbf{g}'(e^{-j\theta} s_1) s_{i,\tau}^* \right\} p_i^* + \\ \mathbb{E} \left\{ \mathbf{g}(e^{-j\theta} s_1)^* \mathbf{g}(e^{-j\theta} s_1) s_{i,\tau}^* \right\} p_i + \\ \frac{1}{2} \mathbb{E} \left\{ \mathbf{G}(e^{-j\theta} s_1)^* \mathbf{g}''(e^{-j\theta} s_{1,\tau}) s_{i,\tau}^3 \right\} (p_i^*)^2 - \\ \frac{1}{2} \mathbb{E} \left\{ \mathbf{g}'(e^{-j\theta} s_1)^* \mathbf{g}(e^{-j\theta} s_{1,\tau}) s_{i,\tau} (s_i^*)^2 \right\} p_i^2 + \\ o(\|\mathbf{p}_{-1}\|^2) \quad (16)$$

Noting that by assumptions (b) and (c), we find that

$$\begin{aligned} & \mathbb{E} \left\{ \mathbf{G} \left(e^{-j\theta} s_1 \right)^* \mathbf{g}' \left(e^{-j\theta} s_1 \right) s_{i,\tau}^2 \right\} \\ & \approx \mathbb{E} \left\{ \mathbf{G} \left(e^{-j\theta} s_1 \right)^* \mathbf{g}' \left(e^{-j\theta} s_1 \right) \right\} \mathbb{E} \left\{ s_{i,\tau}^2 \right\} = 0 \end{aligned} \quad (17)$$

$$\begin{aligned} & \mathbb{E} \left\{ \mathbf{g} \left(e^{-j\theta} s_1 \right)^* \mathbf{g} \left(e^{-j\theta} s_1 \right) s_{i,\tau}^* s_{i,\tau} \right\} \\ & \approx \mathbb{E} \left\{ \mathbf{g} \left(e^{-j\theta} s_1 \right)^* \mathbf{g} \left(e^{-j\theta} s_1 \right) \right\} \mathbb{E} \left\{ s_{i,\tau}^* s_{i,\tau} \right\} = 0 \end{aligned} \quad (18)$$

Substituting (17) and (18) into (16), we get

$$\begin{aligned} p_i &= \frac{1}{2} \mathbb{E} \left\{ \mathbf{G} \left(e^{-j\theta} s_1 \right)^* \mathbf{g}^n \left(e^{-j\theta} s_{1,\tau} \right) s_{i,\tau}^3 \right\} \left(p_i^* \right)^2 \\ & - \frac{1}{2} \mathbb{E} \left\{ \mathbf{g}' \left(e^{-j\theta} s_1 \right)^* \mathbf{g} \left(e^{-j\theta} s_{1,\tau} \right) s_{i,\tau} \left(s_i^* \right)^2 \right\} p_i^2 + o \left(\left\| \mathbf{p}_{-1} \right\|^2 \right) \end{aligned} \quad (19)$$

Equation (15) and (19) indicate that under the assumption (d) $\mathbb{E} \left\{ \mathbf{G} \left(e^{-j\theta} s_i \right)^* \mathbf{g} \left(e^{j\theta} s_{i,\tau}^* \right) s_{i,\tau} \right\} \neq 0$, the proposed algorithm converges to the vector \mathbf{p}_1 , and the convergence is at least quadratic.

Moreover, the **Theorem** is meaningful for real applications to choose proper nonlinear functions. For instance, assume that the square function $\mathbf{G}(y) = y^2$ is chosen for the extraction of independent circular sources with no temporally linear autocorrelation, we have the convergence condition

$$\begin{aligned} & \mathbb{E} \left\{ \mathbf{G} \left(e^{-j\theta} s_i \right)^* \mathbf{g} \left(e^{j\theta} s_{i,\tau}^* \right) s_{i,\tau} \right\} = e^{j\theta} \mathbb{E} \left\{ \left(s_i^* \right)^2 s_{i,\tau} \right\} \neq 0 \\ & \Rightarrow \mathbb{E} \left\{ s_i^2 \left(s_{i,\tau}^* \right)^2 \right\} = \text{cov} \left(s_i^2(t), s_i^2(t-\tau) \right) \neq 0 \end{aligned} \quad (20)$$

It implies that if the energies of sources are time-autocorrelated, the choosing of the square function will properly lead to the convergence of the algorithm. Besides, although the true nonlinear function that describes nonlinear autocorrelation of sources is unavailable, we can achieve the separation of sources by choosing a proper nonlinear function (e.g. $\mathbf{G}(u) = \log(\cosh(u))$). This can be demonstrated in the next Section.

IV. SIMULATION RESULTS AND DISCUSSION

Here, several sets of simulation results are provided to demonstrate the performance of the proposed algorithm. In order to measure the separation performance, we use the performance index [28].

$$PI = \frac{1}{n^2} \left\{ \sum_{i=1}^n \left(\sum_{j=1}^n \frac{|c_{ij}|}{\max_k |c_{ik}|} - 1 \right) + \sum_{j=1}^n \left(\sum_{i=1}^n \frac{|c_{ij}|}{\max_k |c_{kj}|} - 1 \right) \right\} \quad (21)$$

where c_{ij} is the ij th element of $n \times n$ combined mixing-separating matrix $\mathbf{C} = \mathbf{W}^H \mathbf{U} \mathbf{A}$, and $PI \in [0, 1]$.

As \mathbf{C} converges to a generalized permutation matrix, PI will converge to zero, and sources will be retrieved exactly.

Besides, we also utilized the mean square error (MSE) in decibel to measure the performance in some experiments.

$$MSE_i = -10 \log \left(\frac{1}{T} \sum_{t=1}^T \| y_i(t) - s_i(t) \|^2 \right) \quad (22)$$

where s_i is the original signal and y_i is the recovered corresponding signal (both are normalized to have zero mean and unit variance), T denotes the data length of samples. The higher MSE_i is (usually over 20dB), the better the performance is [21].

A. Experiments on Artificial Complex Sources with Nonlinear Autocorrelation

In the simulation, we test the performance of C-FastNA on square temporal autocorrelated complex Gaussian sources. Firstly, the signals are modeled using a standard autoregressive model

$$s_i(t) = \rho_i s_i(t-1) + \zeta_i(t) \quad (23)$$

where ρ_i is the correlation coefficient of the i th source, and $\zeta_i(t)$ is a complex Gaussian random number. The correlation coefficient ρ_i was set 0.8 for all sources. Then the signs of the signals are completely randomized by multiplying each signal by a binary i.i.d. signal that takes the values ± 1 with equal probabilities. Thus the source signals are created with square temporal autocorrelations (or variance nonstationarity), which could not be separated by ordinary source separation methods based on non-Gaussianity such as Complex FastICA [8], [9] as well as linear autocorrelation methods such as AMUSE [14], JADE [29] and SOBI [12].

We mixed ten complex valued sources described above, with data length varied from 500 to 5000 samples. The mixing matrix \mathbf{A} in data model of (1) and the initial separating matrix $\mathbf{W}^{(0)}$ were created randomly. The ten sources with 5000 samples are shown in Fig. 1. We chose the time delay $\tau=1$, which is suitable for all experiments.

The C-FastNA algorithm with the nonlinear functions ($\mathbf{G}(u) = u^2$ and $\mathbf{G}(u) = \log(\cosh(u))$) were used to estimate the separating matrix. For comparison, we also carry out the cumulant-based approach using the nonstationarity of variance (CANSV) [17], C-FastICA [8], and SOBI [12]. The total number of iterations was set to proper value 100 for the convergence of the methods. The results involving 100 realizations are shown in Table. I. From the comparison of PI among different separation methods, we can observe that traditional methods like SOBI and C-FastICA fail to achieve the separation, while successful separation is obtained by CANSV and the proposed method.

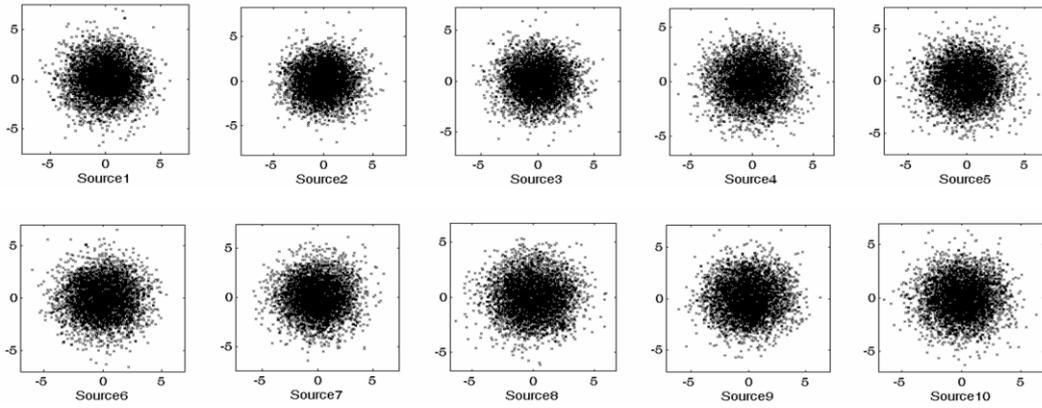


Fig. 1. Ten sources with square temporal autocorrelation

TABLE I: PI AND EXECUTION TIME FOR SEPARATION METHODS. (TEN COMPLEX GAUSSIAN SOURCES, RANDOM MIXING MATRIX. 100 MONTE CARLO REALIZATIONS)

Separation Methods	Average Performance Index				Average Execution Time (s)			
	Number of Samples				Number of Samples			
	500	1000	2000	5000	500	1000	2000	5000
C-FastICA	0.87	0.87	0.88	0.86	0.87	0.36	0.53	1.33
SOBI	0.84	0.81	0.81	0.80	0.79	0.21	0.24	0.79
CANSV	0.12	0.07	0.05	0.03	2.41	4.09	7.21	17.61
C-FastNA $G(u) = \log(\cosh(u))$	0.36	0.17	0.12	0.06	1.05	2.80	5.73	10.01
C-FastNA $G(u) = u^2$	0.07	0.04	0.03	0.02	0.66	1.62	3.23	6.01

From the comparison of PI among three separation methods, we can observe that similar separation performance is achieved, which confirm the validity of our method. The proposed algorithm using $G(u) = u^2$ behaves a slightly better separation performance than the other two algorithms. Furthermore, the comparison of execution time demonstrates that a suitable choice of the nonlinear function for our method might lead to significantly lower computational time. Although our algorithm in (10) has the same quantity level of computational complexity $o(n^2)$ as CANSV algorithm, the latter needs to compute more terms from the conjugate gradient of the cross-cumulant every iterative than our method.

The robustness of the methods was also investigated by choosing different single time lag τ . Fig. 2 illustrates the average performance indexes against iteration numbers by the proposed algorithm ($G(u) = u^2$) with different time lags τ ($\tau=1, 2, 3, 4$, and 5 , respectively). Ten sources above with 5000 samples were employed. We can see that the performance of the proposed algorithm with $\tau=1$ is best, however, the performance with other lag τ (whose value is close to $\tau=1$) is also similar to $\tau=1$. Thus, the experiments show that the performance of the C-FastNA is still good if the choice of the time delay τ is not far from the time delay $\tau=1$. Nevertheless, if the time delays τ are far from the true one, the performance degrades.

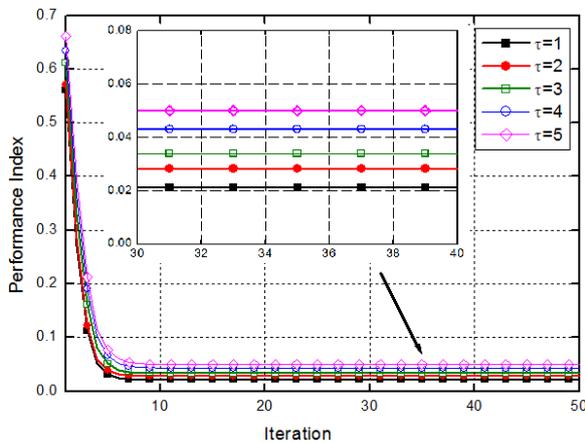


Fig. 2. The average performance indexes over 100 Monte Carlo Realizations against iteration numbers by the proposed C-FastNA algorithm with different time delay τ

Moreover, we randomly added complex outliers whose values are $10+10j$ in each source signal in order to further examine the robustness of algorithms. Fig. 3 and Fig. 4 shows the average performance indexes against iteration numbers when the numbers of outliers are 10 and 20 respectively ($\tau=1$). Only the method C-FastNA which employs the nonlinear function $G(u) = \log(\cosh(u))$ is not affected dramatically by outliers and is seen to perform well. The method CANSV, which use a cross-cumulant as a measure of nonstationarity, behaves similar to the method C-FastNA ($G(u) = u^2$) especially when the number of outliers is greater, i.e., both of them fail completely in separating the sources. Indeed, both the methods CANSV and C-FastNA ($G(u) = u^2$) utilize the

forth-order statistics of signals, which typically result in an estimate sensitive to outliers. This is the key reason why the two algorithms perform non-robust in terms of separation accuracy.

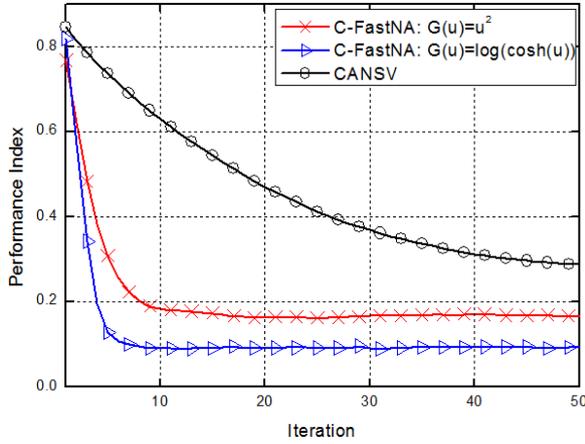


Fig. 3. The average performance indexes over 100 Monte Carlo Realizations for 10 sources with square temporal autocorrelation when 10 outliers are introduced

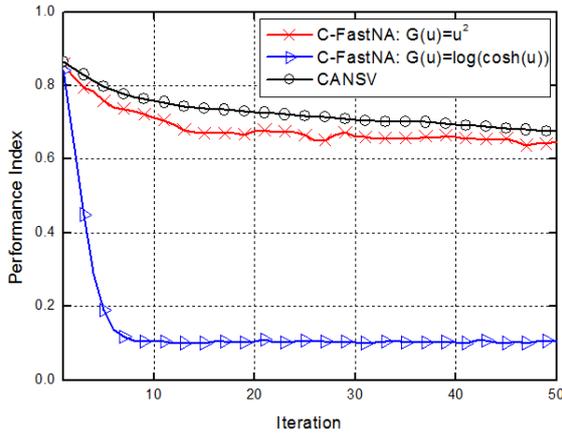


Fig. 4. The average performance indexes over 100 Monte Carlo Realizations for 10 sources with square temporal autocorrelation when 20 outliers are introduced

B. Experiments on the Real-World Single Sideband (SSB) Signals

Finally we show the separation validity and efficiency of our proposed algorithms for practical single sideband (SSB) signals carrying the music information (typically non-Gaussian distribution sources).

To begin with, three baseband sources with three types were utilized: (1) bass; (2) tenor; and (3) jazz music. The sources are all with sampling frequency 8 kHz and time span 2.5 s. Then all the baseband sources were represented as equivalent low-pass signals with single sideband

$$\mathbf{s}_l(t) = \mathbf{s}(t) + \hat{j}\hat{\mathbf{s}}(t) \quad (24)$$

where $\hat{\mathbf{s}}(t)$ denotes the Hilbert transform of $\mathbf{s}(t)$.

As we know, in wireless communications, the equivalent low-pass representation is meaningful in that

the signals which are single-side band modulated can save considerable spectrum and power resources comparing to the double-side band signals [30]. Furthermore, we assume that the communication channel is linear without complicated factors like IQ unbalance, multi-path effect and non-linear distortion. Thus in the simulation, the MIMO channel can be characterized by the time-invariant mixing matrix \mathbf{A} , which was created randomly in data model of (1). We chose the nonlinearity $G(u) = u^2$.

Fig. 5 gives the separation results provided by C-FastNA in the waveforms of the real parts of signals, whose imaginary parts need not be redundantly presented here since the imaginary parts are the Hilbert transform of the real parts. It illustrates that our method seems to have good prospects in real-world applications.

We also investigated the effect of noise on the separation performance when received signals are contaminated with complex white Gaussian noise, i.e., $\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{v}(t)$, where $\mathbf{v}(t) = [v_1(t), v_2(t), \dots, v_n(t)]^T \in \mathbb{C}^{n \times 1}$ is the receiver noise vector. A detailed performance of MSE for different signal to noise ratios (10dB, 20dB, 30dB, 40dB, 50dB) was obtained among CANSV, C-FastNA with $G(u) = \log(\cosh(u))$ and $G(u) = u^2$. The results involving 100 realizations are shown in Table. II, which demonstrate that our method behaves equally well as CANSV algorithm. This illustrates the validity of the applications in noise case.

TABLE II: MSE FOR SEPARATION METHODS IN THE COMPLEX WHITE GAUSSIAN NOISE CASE. (TWO SSB SOURCES, RANDOM MIXING MATRIX. 100 MONTE CARLO REALIZATIONS)

Separation Methods	Average MSE (dB)				
	Signal to Noise Ratio (dB)				
	10	20	30	40	50
CANSV	17.1	20.3	35.6	48.2	52.2
C-FastNA $G(u) = \log(\cosh(u))$	14.4	22.1	33.7	46.9	50.1
C-FastNA $G(u) = u^2$	16.4	27.4	40.8	49.5	55.2

V. CONCLUSION

In this paper, we have proposed a new method for blind source separation of complex sources through the optimization of nonlinear autocorrelation contrast. Depending on temporal structure with nonlinear autocorrelation of the signals, our method holds a potential capability of extracting complex sources with arbitrary distribution (Gaussian and non-Gaussian), which is unattainable for traditional BSS methods. We have also provided a theoretical analysis about the convergence of the algorithm based on Newton

iterations. Therefore, this paper is an important complement for our previous work. It has been shown through computer simulations that the algorithm offers more robust performance and lower computational cost than classical CANSV method. The experiments for SSB signals in Section 4.2 indicate that the method might be applied to wireless broadcast system.

However, the channel model in this paper is simply represented by instantaneous mixtures, such that further research is still needed to determine if our method can be extended to implement in more practical and complicated scenarios like convolutive mixing case or fast time-variant channel.

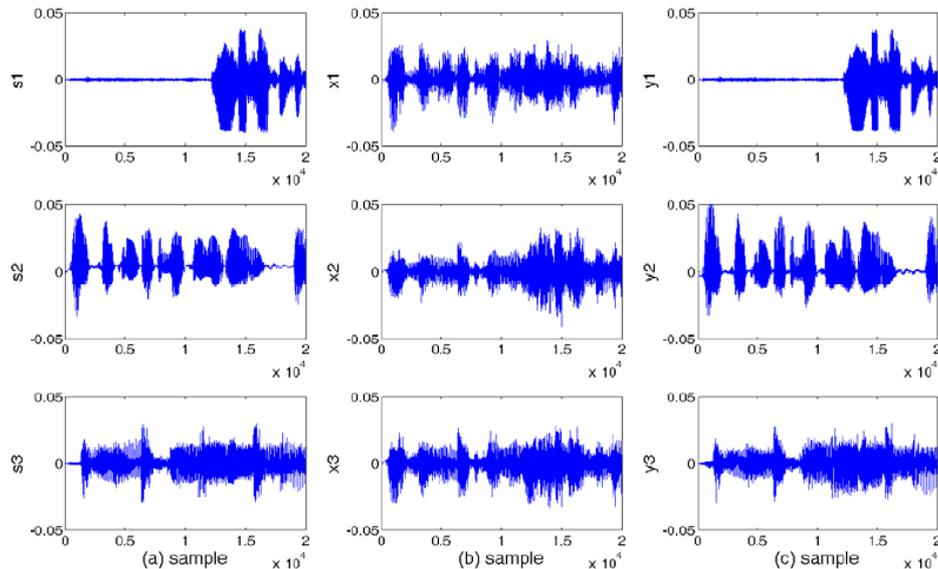


Fig. 5. Separation result of three SSB signals with bass, tenor and jazz music information respectively. (a) Original sources, (b) Observations, (c) Extracted sources.

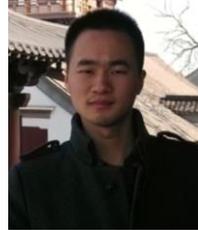
ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China under Grant No. 61172061.

REFERENCES

- [1] L. R. Arnaut and C. S. Obiekiezie, "Source separation for wideband energy emissions using complex independent component analysis," *IEEE Transactions on Electromagnetic Compatibility*, vol. 56, no. 3, pp. 559-570, 2014.
- [2] H. D. Han and Z. Ding, "Steepest descent algorithm implementation for multichannel blind signal recovery," *IET Communications*, vol. 6, no. 18, pp. 3196-3203, 2012.
- [3] G. Chabriel and J. Barrere, "Non-symmetrical joint zero-diagonalization and MIMO zero-division multiple access," *IEEE Transactions on Signal Processing*, vol. 59, no. 5, pp. 2296-2307, 2011.
- [4] J. C. Wang, C. H. Lin, E. Siahhaan, B. W. Chen, and H. L. Chuang, "Mixed sound event verification on wireless sensor network for home automation," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 1, pp. 803-812, 2014.
- [5] Y. Q. Li, Z. L. Yu, N. Bi, Y. Xu, Z. H. Gu, and S. I. Amari, "Sparse representation for brain signal processing: A tutorial on methods and applications," *IEEE Signal Processing Magazine*, vol. 31, no. 3, pp. 96-106, 2014.
- [6] D. Z. Feng, H. Zhang, and W. X. Zheng, "Bi-iterative algorithm for extracting independent components from array signals," *IEEE Transactions on Signal Processing*, vol. 59, no. 8, pp. 3636-3646, 2011.
- [7] G. Fabrizio and A. Farina, "An adaptive filtering algorithm for blind waveform estimation in diffuse multipath channels," *IET Radar, Sonar & Navigation*, vol. 5, no. 3, pp. 322-330, 2011.
- [8] H. L. Li, N. M. Correa, P. A. Rodriguez, V. D. Calhoun, and T. Adali, "Application of independent component analysis with adaptive density model to complex-valued fMRI data," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 10, pp. 2794-2803, 2011.
- [9] P. Comon and C. Jutten, *Handbook of Blind Source Separation, Independent Component Analysis and Applications*, New York: Eds. Academic, 2010, ch. 1, pp. 11-15.
- [10] A. Cichocki and S. I. Amari, *Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications*, New York: Wiley, 2002, ch. 5.
- [11] S. Hosseini and Y. Deville, "Blind separation of parametric nonlinear mixtures of possibly autocorrelated and non-stationary sources," *IEEE Transactions on Signal Processing*, vol. 62, no. 24, pp. 6521-6533, 2014.
- [12] A. Belouchrani, K. A. Meraim, J. F. Cardoso, and E. Moulines, "A blind source separation technique based on second order statistics," *IEEE Transactions on Signal Processing*, vol. 45, no. 2, pp. 434-444, 1997.
- [13] L. Molgedey and H. G. Schuster, "Separation of a mixture of independent signals using time delayed correlations," *Physical Review Letters*, vol. 72, no. 23, pp. 3634-3637, 1994.
- [14] L. Tong, R. W. Liu, V. Soon, and Y. F. Huang, "Indeterminacy and identifiability of blind identification," *IEEE Transactions on Circuits and Systems*, vol. 38, no. 5, pp. 499-509, 1991.
- [15] K. T. Herring, A. V. Mueller, and D. H. Staelin, "Blind separation of noisy multivariate data using second-order statistics: Remote-sensing applications," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 10, pp. 3406-3415, 2009.
- [16] J. V. Stone, "Blind source separation using temporal predictability," *Neural Computation*, vol. 13, pp. 1559-1574, 2001.

- [17] A. Hyvarinen, "Blind source separation by nonstationarity of variance: A cumulant-based approach," *IEEE Transactions on Neural Networks*, vol. 12, no. 6, pp. 1471-1474, 2001.
- [18] J. T. Chien and H. L. Hsieh, "Nonstationary source separation using sequential and variational bayesian learning," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 24, no. 5, pp. 681-694, 2013.
- [19] J. Gunther, "Learning echo paths during continuous double-talk using semi-blind source separation," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, no. 2, pp. 646-660, 2012.
- [20] Z. Shi and C. Zhang, "Fast nonlinear autocorrelation algorithm for source separation," *Pattern Recognition*, vol. 42, no. 9, pp. 1732-1741, 2009.
- [21] Z. Shi, H. J. Zhang, and Z. G. Jiang, "Hybrid linear and nonlinear complexity pursuit for blind source separation," *Journal of Computational and Applied Mathematics*, vol. 236, pp. 3434-3444, 2012.
- [22] A. Hyvarinen, "A unifying model for blind separation of independent sources," *Signal Processing*, vol. 85, no. 7, pp. 1419-1427, 2005.
- [23] P. C. Xu, Y. H. Shen, W. Jian, W. Zhao, and C. Peng, "Blind source separation of complex valued signals based on nonlinear autocorrelation," *Circuits, Systems and Signal Processing*, Published, February 14, 2015.
- [24] S. Amari, "Natural gradient works efficiently in learning," *Neural Computation*, vol. 10, pp. 251-276, 1998.
- [25] E. Bingham and A. Hyvarinen, "A fast fixed-point algorithm for independent component analysis of complex valued signals," *International Journal of Neural Systems*, vol. 10, pp. 1-8, 2000.
- [26] K. B. Petersen and M. S. Pedersen. *The Matrix Cookbook*. [Online]. ch 2. pp. 10-12. Available: <http://www2.imm.dtu.dk/pubdb/p.php?3274>.
- [27] N. Delfosse and P. Loubaton, "Adaptive blind separation of independent sources: a deflation approach," *Signal Processing*, vol. 45, pp. 59-83, 1995.
- [28] S. I. Amari, A. Cichocki, and H. Yang, "A new learning algorithm for blind source separation," *Advances in Neural Information Processing Systems*, vol. 8, pp. 757-763, 1996.
- [29] J. F. Cardoso and A. Souloumiac, "Blind beamforming for non-Gaussian signals," *IEE Proceedings F (Communications, Radar and Signal Processing)*, vol. 140, pp. 362-370, 1993.
- [30] J. G. Proakis, *Digital Communications*, Asia: McGraw-Hill Education Co., 2007, ch. 2, pp. 18-27.



Pengcheng Xu was born in Zhejiang Province, China, in 1987. He received his M.S. degree from College of Communications Engineering, PLA University of Science and Technology in 2012. He is currently pursuing the Ph.D. degree in College of Communications Engineering, PLA University of Science and Technology. His research interests include blind signal processing and wireless communications.



Yuehong Shen was born in 1959. He received his Ph.D. degree in Communication Engineering from Nanjing University of Science and Technology in 1999. And now, he is a Professor and doctor advisor of College of Communications Engineering, PLA University of Science and Technology. His current research interests include wireless communication systems, digital communications and signal processing.