Journal of Information Hiding and Multimedia Signal Processing ©2017 ISSN 2073-4212
Ubiquitous International Volume **8**, Number **2**, March **2017**

A Fast Convergence Adaptive Complex-Valued Blind Source Separation Method

YanFei Jia[1,2], Xiaodong Yang[2]

[1]College of Electrical and Information Engineering
Beihua University
No.1 Xinshan Street, Longtan District, Jilin City, China

[2]College of Information and Communication Engineering
Harbin Engineering University
No.145 Nantong Street, Nangang District, Harbin City, China
jia_yanfei@163.com

Received March, 2016; revised December, 2016

Abstract. *This paper deals with adaptive blind source separation of noncircular signals from a complex-valued mixing signal. By making separated signals have the same second order statistics as the noncircular original signals, some adaptive complex-valued blind source separation methods are proposed. They achieve some advantages over the batch method based on second order statistics, but their convergence speed is slower. By incorporating the restriction that the separating matrix of white mixing signals should be a unitary matrix in the gradient cost function, an adaptive complex-valued blind source separation method with fast convergence speed is proposed. In the proposed methods, the separating matrix of whitened signals approximates a unitary matrix through the optimization process, without additional operations that force the matrix to be a unitary matrix. Validity of the proposed methods is demonstrated by simulations with Gaussian noncircular signals and communication signals. Simulation results prove the proposed methods have faster convergence speed and more stable performance for different mixing signals than other methods based on second order statistics.*
Keywords: Blind source separation; Noncircular signal; Unitary matrix; Complex valued-mixing signal.

1. **Introduction.** Blind source separation (BSS) is a signal processing method that can separate original signals from observed signals received by a sensor array without prior information about the system. Complex-valued blind source separation is used to separate an original signal from a complex-valued mixing signal. It has been widely used in biomedical signal processing, power system analysis, communication signal processing, and natural image processing [1-5]. Research on complex-valued BSS can be divided into three categories. The first category is methods based on a nonlinear function [6-12], which use a nonlinear function to approximate entropy, higher order statistics, or probability density. The second is methods based on kurtosis, higher order cumulant or moment [1,13-18], which directly use kurtosis or cumulant as the cost function to measure the non-Gaussian character of separated signals. The last is methods based on second order statistics, which use the characteristics of noncircular signals to realize blind source separation [19-25]. Each category has its own merits and demerits. The method based

on second order statistics is only suitable for noncircular signals, but has a simple structure, lower computation complexity, and is appropriate for complex Gaussian noncircular signals.

The strong-uncorrelating transform (SUT) [19] method is an earlier complex-valued BSS method based on second order statistics. It utilizes Takagis factorization to deal with a whitened signal to separate original signals and whenever applicable, remains perhaps the simplest and most accessible approach [26]. However, it is a batch method that is not suitable for real time processing. To overcome this problem, adaptive complex-valued BSS methods based on second order statistics are proposed. Scott proposed an equivariant adaptive method [21] which directly gives the update formula of the separating matrix and proves its convergence. Yang used an adaptive method to realize the SUT method [23]. Cong simultaneously used covariance and pseudo-covariance to propose an adaptive complex-valued BSS method [24]. Hao, also inspired by the SUT method, proposed a batch method based on the pseudo-uncorrelating transform (PUT)[25], which supposes that the covariance of the original signal is diagonal and its pseudo-covariance is a unit matrix. In recent years, the performance and separability of SUT have also been researched [26-27]. Adaptive complex-valued BSS methods based on second order statistics are designed to separate signals in real time; convergence speed directly affects real time. In this paper, to accelerate convergence rate, three adaptive complex-valued BSS methods based on second order statistics are proposed with fast convergence rates.

2. Model and Property.

2.1. Complex-valued BSS model.

Generally, a noise-free linear complex-valued BSS model can be expressed as

$$x = As \tag{1}$$

where $x = [x_1, x_2, \cdots, x_m]^T$ is the complex-valued mixed signal of n statistically independent zero mean signals $s = [s_1, s_2, \cdots, s_n]^T$. A is a nonsingular complex-valued mixing matrix with size $m \times n$. In this paper, we suppose that the number of mixed signals equals the number of original signals. The aim of BSS is to estimate both the mixing matrix A and original signals s from observed signals without using information about the system. In non-underdetermined BSS, it is realized by searching the optimal separating matrix W

$$y = Wx \tag{2}$$

where $W = [w_1, w_2, \cdots, w_n]^{\mathrm{T}}$ and $y = [y_1, y_2, \cdots, y_n]^T$ is the estimated signal of the original signal s . If $WA = I$, then $y = s$. Since complex-valued BSS does not utilize information about the mixing system, it has some indeterminacy in amplitude, sequence, and phase. This indeterminacy does not affect the shape of the estimated source signal waveform, which contains most information about the source signals.

2.2. Second order statistics property of complex-valued vector.

Define a complex-valued random vector $s = s_R + js_I$, where s_R and s_I are the real and imaginary parts, respectively, and $j = \sqrt{-1}$ is the imaginary unit. The expectation $E[.]$ of random vector s is defined as

$$E[s] = E[s_R] + jE[s_I] \tag{3}$$

and its covariance is defined as

$$\mathrm{cov}(s) = E[(s - E(s))(s - E(s))^H] \tag{4}$$

where $()^H$ denotes Hermitian transpose. Its pseudo-covariance is defined as

$$p\mathrm{cov}(s) = E[(s - E(s))(s - E(s))^T] \tag{5}$$

where $()^T$ denotes transpose. The covariance, together with pseudo-covariance, is the full expression of second order statistics for random vector s . If the pseudo-covariance equals zero, the random vector is called circular or proper. If both the covariance and pseudo-covariance of the random vector are nonzero, the random vector is called noncircular or improper. The complex-valued noncircular signal is frequently used in digital communications, such as with BSPK, UQPSK, MASK, and ASK signals.

3. **Related Work.** The adaptive BSS methods based on the second order statistics of complex-valued signals all use the second order statistic properties of noncircular to separate the original signal. They can be described as follows.

Suppose the original signal is $s = [s_1, s_2, \cdots, s_n]^{\mathrm{T}}$ and that s_i is a noncircular signal with unit covariance and zero mean. The complex-valued mixing signal x is

$$x = As \tag{6}$$

and the estimated signal $y = [y_1, y_2, \cdots, y_n]^T$ of the original signal is

$$y = Wx \tag{7}$$

If we successfully estimate the original signal, then the characteristics of the estimated signal must be the same as the original signal. So, the pseudo-covariance matrix of the estimated signal y is a nonzero diagonal matrix and its covariance matrix is a unit matrix. They can be expressed as

$$\begin{cases} E[yy^H] = E[wxx^H w^H] = I \\ E[yy^T] = E[wxx^T w^T] = \Lambda \end{cases} \tag{8}$$

Adaptive complex-valued BSS methods based on second order statistics use the above characteristic to construct a cost function or to directly propose an update formula.

Scott directly proposes the separating matrix update formula[21]

$$w(k+1) = w(k) + \mu(I - w(k)Rw(k)^H - tri[w(k)Pw(k)^T])w(k) \tag{9}$$

where k is the number of iterations, μ is the learning rate, $R{=}E[xx^H]$, and $P{=}E[xx^T]$.

Cong uses the above characteristic to construct a cost function [24]

$$\begin{aligned} \Phi = &\tfrac{1}{4}tr[(w(k)Rw(k)^H - \Lambda_H)(w(k)Rw(k)^H - \Lambda_H)^H] + \\ &\tfrac{1}{4}tr[(w(k)Pw(k)^T - \Lambda_T)(w(k)Pw(k)^T - \Lambda_T)^H] \end{aligned} \tag{10}$$

and deduces the update formula

$$w(k+1) = w(k) - \mu(w(k)Rw(k)^H - \Lambda_H)w(k)R - \mu(w(k)^* P^H w(k)^H - \Lambda_T^H)w(k)P \tag{11}$$

Yang also uses second order statistics of noncircular signals to construct a cost function. The method proposed by Yang contains two parts [23]. The first part is adaptive whitening with the cost function

$$J_1(\mathbf{B}) = \frac{1}{2}||I - BRB^H||_F^2 \tag{12}$$

The update formula of the whitened matrix B is

$$B_{k+1} = B_k + \mu(I - B_k R B_k^H)B_k \tag{13}$$

The second part makes the pseudo-covariance of the whitened signal diagonal. The second part's cost function is

$$J_2(w) = \frac{1}{2}||\Lambda - wE[zz^T]w^T||_F^2 \tag{14}$$

where z=Bx is the whitened signal. Its update formula is

$$w_{k+1} = w_k + \mu(\Lambda - w_k BPB^T w_k^T)(w_k BPB^T w_k^T)^* w_k \tag{15}$$

$$w_{k+1} = w_{k+1}(w_{k+1}^H w_{k+1})^{-1/2} \tag{16}$$

Equation (16) is used to force the separating matrix to be a unitary matrix so that the covariance of the separating matrix signal will be a unit matrix.

From (9) and (11), we see that the methods converge to equilibrium only if the covariance and pseudo-covariance matrix are simultaneously diagonal. If only one matrix is diagonal, they converge to equilibrium until the other is diagonal. This has a drawback: if one covariance matrix becomes diagonal, we cant guarantee the matrix will still be diagonal in the following iteration that makes the other covariance matrix diagonal; this negatively affects convergence rate. Yang uses a serial update method to realize both covariance matrices as diagonal, which avoids the drawback of the Scott and Cong methods. However, equation (16) changes gradient direction in iterations, which affects convergence rate and increases computational complexity. To overcome this problem, based on the research of Yang, this paper proposes an adaptive BSS method with a fast convergence rate.

4. **Proposed method.** In order to accelerate the convergence rate of adaptive complex-valued BSS based on second order statistics, we also use a serial update method to update the separating matrix. The first step is adaptive whitening and we directly use Yang's method. The whitening matrix is rewritten as

$$B_{k+1} = B_k + \mu(I - B_k R B_k^H) B_k \tag{17}$$

where k is the number of iterations, B is the whitening matrix, μ is the learning rate, I is a unit matrix, R=$E[xx^H]$, and x is the mixed signal or observed signal. The second step is to diagonalize the pseudo-covariance of the whitened signal while keeping the covariance matrix diagonal. We use (14) as a cost function and its ordinary gradient with respect to w^* is

$$dw = (wPw^T - \Lambda)w^* P^H \tag{18}$$

where w is the separating matrix of the whitened signal, P=$BE[xx^T]B^T$, and Λ is a diagonal matrix comprised of diagonal elements of wPw^T . According to the supposition of a noncircular signal, the covariance matrix of separated signals should be a unit matrix

$$E[yy^H] = wE[BRB^H]w^H = I \tag{19}$$

where $y = wBx$ is the estimation of the original signal. After pre-whitening, $E[BRB^H]$=I . From (19) we have

$$ww^H - I = 0 \tag{20}$$

Differentiate (20) to obtain

$$d(ww^H - I) = dww^H + wdw^H = 0 \tag{21}$$

where dww^H is a skew symmetric matrix. The steepest direction keeping w unitary is

$$\Delta w = dww^H w - wdw^H w \tag{22}$$

The update formula for the separating matrix can be expressed as

$$w_{k+1} = w_k - \mu\Delta w_k = w_k - \mu[dw_k w_k^H w_k - w_k (dw_k)^{\mathrm{H}} w_k] \tag{23}$$

If w_k is a unitary matrix, w_{k+1} is approximately a unitary matrix. This can be proven with supposition $w_k w_k^H = I$ (w_k is a unitary matrix). The proof procedure is as following:

$$
\begin{aligned}
&w_{k+1}w_{k+1}^H \\
&= \{w_k - \mu[dw_k w_k^H w_k - w_k(dw_k)^H w_k]\}\{w_k - \mu[dw_k w_k^H w_k - w_k(dw_k)^H w_k]\}^H \\
&= w_k w_k^H - \mu[dw_k w_k^H w_k w_k^H - w_k(dw_k)^H w_k w_k^H + w_k w_k^H w_k(dw_k)^H - \\
&\quad w_k w_k^H dw_k w_k^H] + o(\mu^2) \\
&= w_k w_k^H - \mu[(dw_k)w_k^H - w_k(dw_k)^H + w_k(dw_k)^H - (dw_k)w_k^H] + o(\mu^2) \\
&= I + o(\mu^2)
\end{aligned} \tag{24}
$$

In (24), μ is the learning rate, which is a small value, and so the value of μ^2 is even smaller. If we omit $o(\mu^2)$, w_{k+1} is a unitary matrix. Substitution of (18) into (23) gives

$$
\begin{aligned}
w_{k+1} &= w_k - \mu[(w_k P w_k^T - \Lambda)w_k^* P^H w_k^H - w_k P w_k^T(w_k^* P^H w_k^H - \Lambda^H)]w_k \\
&= w_k - \mu[(D-\Lambda)D^H - D(D^H - \Lambda^H)]w_k \\
&= w_k - \mu(C - C^H)w_k
\end{aligned} \tag{25}
$$

where $C=(D-\Lambda)D^H$ and $D=w_k P w_k^T$.

If w is a unitary matrix, then $ww^H ww^H$ is a unit matrix. Consider if w is just approximately a unitary matrix, then an error of $o(\mu^2)$ exists in (24) between ww^H and the unit matrix. The error has a bigger effect on $ww^H ww^H$ than ww^H :

$$|ww^H ww^H - I| > |ww^H - I| \tag{26}$$

This can be proven by the following based on (24)

$$w_{k+1}w_{k+1}^H w_{k+1}w_{k+1}^H = (I + o(\mu^2))(I + o(\mu^2)) = I + 2o(\mu^2) + o(\mu^2)^2 \tag{27}$$

The term $o(\mu^2)$ and $o(\mu^2)^2$ are both bigger than zero, so

$$w_{k+1}w_{k+1}^H w_{k+1}w_{k+1}^H - I = 2o(\mu^2) + o(\mu^2)^2 > o(\mu^2) \tag{28}$$

Based on (24) ($w_{k+1}w_{k+1}^H - I = o(\mu^2)$) and (28), we can prove (26) that is true. Based on the above analysis, $ww^H ww^H$ is more sensitive than ww^H to error $o(\mu^2)$, so we use $ww^H ww^H - I$ to measure the distance between w and be a unitary matrix. Using ww^H instead of w in (20) obtains

$$ww^H ww^H - I = 0 \tag{29}$$

Differentiate (29) to obtain

$$wdw^H ww^H + ww^H dww^H = 0 \tag{30}$$

where $wdw^H ww^H$ is a skew symmetric matrix. The steepest direction keeping w unitary is

$$\Delta w = ww^H dww^H w - wdw^H ww^H w \tag{31}$$

The update formula for the separating matrix can be expressed as

$$
\begin{aligned}
w_{k+1} &= w_k - \mu\Delta w_k \\
&= w_k - \mu[w_k w_k^H dw_k w_k^H w_k - w_k(dw_k)^{\mathrm{H}} w_k w_k^H w_k] \\
&= w_k - \mu[w_k w_k^H (D-\Lambda)D^H - D(D^H - \Lambda^H)w_k w_k^H]w_k \\
&= w_k - \mu[F - F^H]w_k
\end{aligned} \tag{32}
$$

where $F=w_k w_k^H (D-\Lambda)D^H$ and $D=w_k P w_k^T$. The difference between (25) and (32) is that in (31) the term $w_k w_k^H$ is added. If $w_k w_k^H = I$, (32) can be simplified to (25). However, when w_k is just approximately a unitary matrix, (32) is more sensitive than (25) to the error ($w_k w_k^H - I$). This helps to modify the update formula of the separating matrix to make it move toward the optimal direction. The higher the order of ww^H, the more sensitive is its error with a unitary matrix. If we use ww^Hww^H instead of w in (21), we can use the same method to deduce the update formula for the separating matrix. For simplicity, we directly give its expression

$$\begin{aligned} w_{k+1} &= w_k - \mu[w_k w_k^H w_k w_k^H (D-\Lambda)D^H - D(D^H-\Lambda^H)w_k w_k^H w_k w_k^H]w_k \\ &= w_k - \mu[G - G^H]w_k \end{aligned} \tag{33}$$

where $G=w_k w_k^H w_k w_k^H (D-\Lambda)D^H$ and $D=w_k P w_k^T$. Although ww^H and ww^Hww^H increase the computational complexity of (32) and (33), compared with (25), they have the same number of dimensions as w . So, the increased computational complexity is just used to compute ww^H or ww^Hww^H , and it is not very large.

Based on analysis of (24), even if the learning rate is small and a higher order ww^H is used to reduce error in (32) and (33), an error will still exist between $w_{k+1}w_{k+1}^H$ and a unit matrix. To further reduce error, we insert a compensation term into (25), (32), and (33). The new update formulas for the separating matrix corresponding to (25), (32), and (33) are

$$w_{k+1}=w_k - \mu[(D-\Lambda)D^H - D(D^H-\Lambda^H)+w_k w_k^H - I]w_k \tag{34}$$

$$w_{k+1}=w_k - \mu[w_k w_k^H (D-\Lambda)D^H - D(D^H-\Lambda^H)w_k w_k^H + w_k w_k^H - I]w_k \tag{35}$$

$$w_{k+1}=w_k - \mu[w_k w_k^H w_k w_k^H (D-\Lambda)D^H - D(D^H-\Lambda^H)w_k w_k^H w_k w_k^H + w_k w_k^H - I]w_k \tag{36}$$

The initial matrix of w_k is a unit matrix, so the compensation term $w_k w_k^H - I$ equals zero in the first iteration. In the next iteration, w_k is not unitary and the error term $w_k w_k^H - I$ will revise the gradient of w_{k+1} to make w_{k+1} approximate a unitary matrix. When the proposed methods converge to equilibrium, $w_k w_k^H - I=0$, $w_{k+1}=w_k$, and w_{k+1} is a unitary matrix. Equation (17) with (34), (35), and (36) define proposed methods I, II, and III, respectively.

5. **Experimental results and analysis.** In order to test the algorithms, we use a synthesized signal and a digital communication signal as source signals, respectively. For simplicity, we directly use signal expectation instead of instantaneous value. Quality of separation is assessed by using performance index (PI), which is widely used in blind source separation. PI can be expressed as [28]

$$PI(H)=\frac{1}{n(n-1)}\left\{\sum_{i=1}^{n}\left(\sum_{j=1}^{n}\frac{|h_{ij}|}{\max_l |h_{il}|}-1\right)+\sum_{j=1}^{n}\left(\sum_{i=1}^{n}\frac{|h_{ij}|}{\max_l |h_{lj}|}-1\right)\right\} \tag{37}$$

where h_{ij} is the (i,j) element of the global system matrix $H = wBA$, A is the mixing matrix, B is the whitening matrix, w is the separating matrix of whitened signal, and $\max_l |h_{il}|$ and $\max_l |h_{lj}|$ are the maximum absolute value among elements in the ith row and jth column vector of H , respectively. When perfect separation is achieved, the performance index is zero. In practice, the value of the performance index is around 10^{-2} [28], which gives good performance. The deviation of w_k from a unitary matrix is measured by

$$f(k)= ||w_k w_k^{\mathrm{H}} - \mathrm{I}|_{\mathrm{F}}^2 \tag{38}$$

In the first experiment, every source is generated independently and the real and imaginary components of the sources are also generated independently from a Gaussian distribution with zero mean and unit variance. They can be expressed as

$$s_k = N(0,k) + j * N(0,1) \qquad k = 0,1,2 \tag{39}$$

where $N(0, k)$ is a random Gaussian distribution between 0 and k. The first source is circular and the other two are noncircular with distinct spectral coefficients. A mixing matrix is randomly generated from a uniform distribution (0,1). All algorithms are run with the same learning rate of 0.01 and 100 times. In each run, the sources and mixing matrix are regenerated.

Figure 1 shows convergence curves of the SUT [19], Yang [23], Scott [21], and proposed methods I, II, and III. SUT is a batch method, so its convergence curve is a straight line. Every subgraph contains 100 convergence curves corresponding to 100 different runs. From Figure 1, we see that every convergence curve of the proposed method is closer to one another than in the other methods, which means that the proposed method has more stable performance for different mixed sources than the other methods, especially proposed method III. Yang has the worst stability, where some curves deviate heavily from the others.

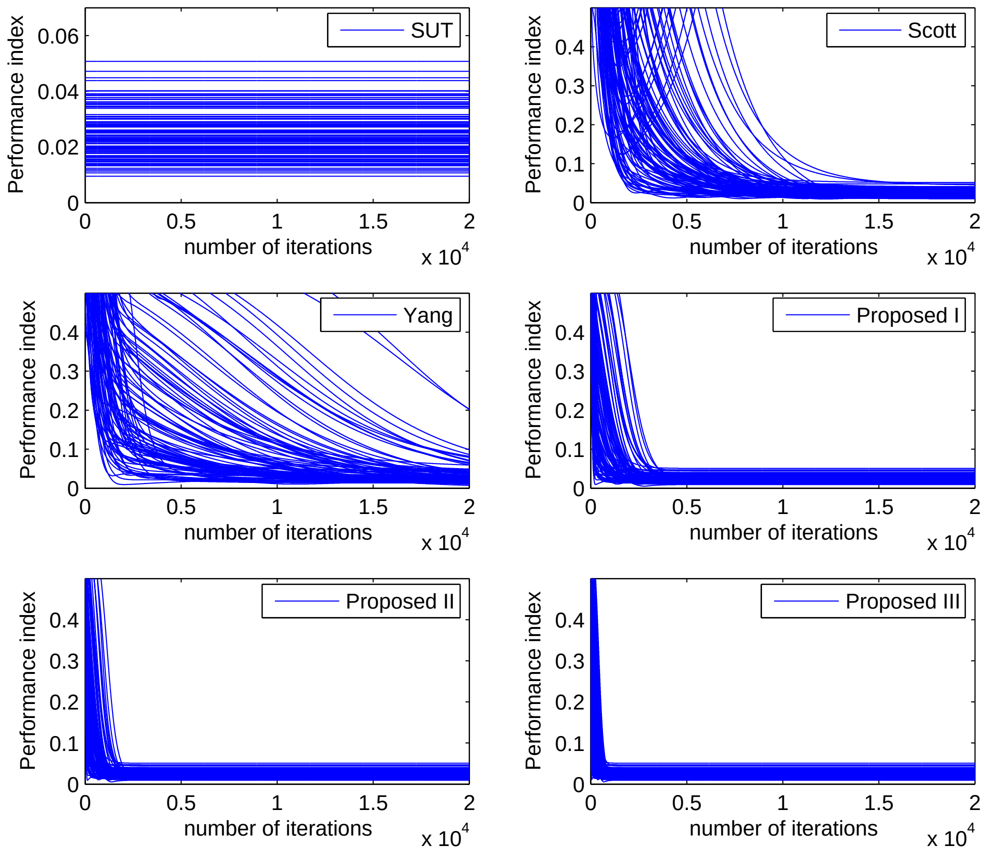

FIGURE 1. Convergence curves with Gaussian signals

Average convergence curves are shown in Figure 2. From Figure 2, we see that the number of iterations required for convergence is more than 20000, 12000, 4000, 2000, and 1000 for Yang, Scott, and proposed methods I, II, and III, respectively. The method with the fastest convergence is proposed method III, followed by proposed methods II and I. The slowest convergence is with the Yang method. At the stable point, we see that the performance index of Scott is bigger than SUT, and the proposed methods have the same performance index as SUT. The bigger the performance index, the bigger the error. So, the proposed method has the same error as SUT, while a smaller error than Scott. The deviations of from a unitary matrix for the proposed methods are shown in Figure 3. As we can see, the separating matrix of the proposed method has very small deviation from a unitary matrix in the whole iterative process, and can almost be omitted.

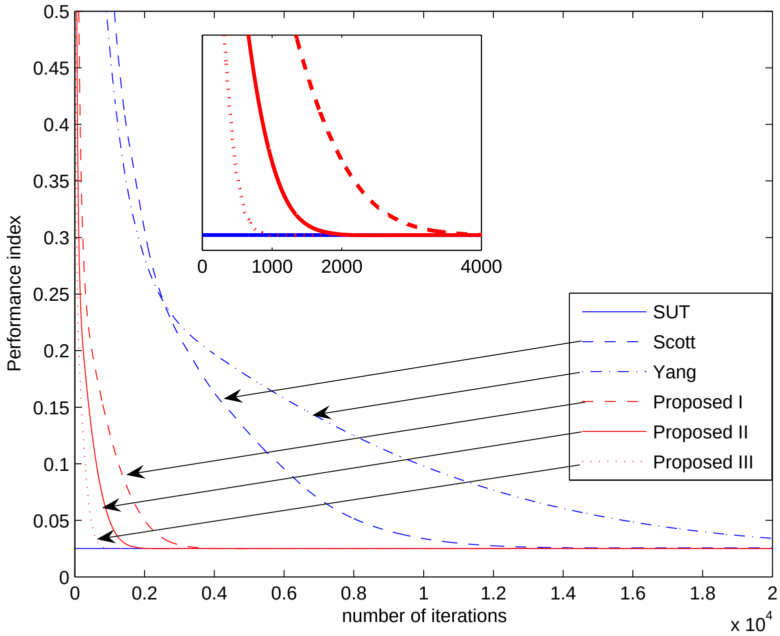

Figure 2. Average convergence curves with Gaussian signals

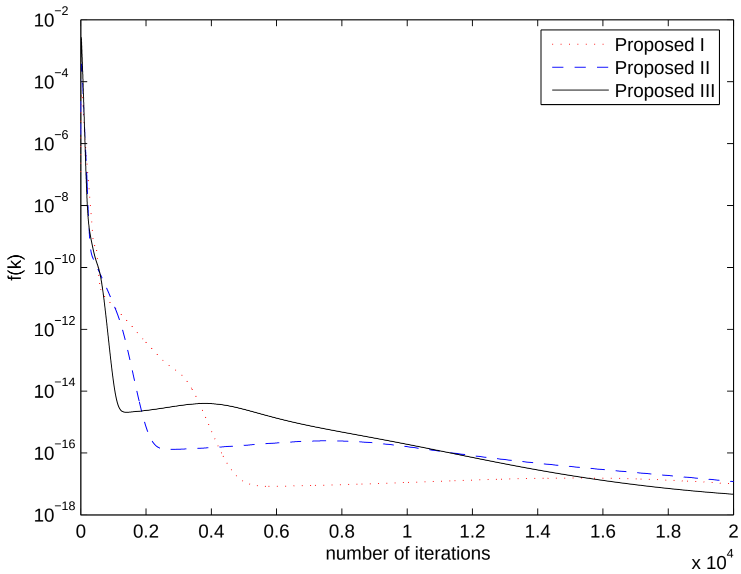

Figure 3. Deviation of separating matrix from a unitary matrix

In the second experiment, we suppose that three digital communication signals, two BPSK and one QAM, impinge on a uniform linear antenna array with three elements from directions of 10^0,25^0, and 30^0. In Figure 4, the first row is unknown source signals that we want to estimate from the mixing signals. Separated signals obtained using the proposed methods are shown in Figure 5. In Figure 5, the first, second, and third rows are the separated signals separated from the mixing signal by using proposed methods I, II, and III, respectively. Comparing the original signals in Figure 4 with the separated signals in Figure 5, we can easily see that that the fist and second column signals in every row in Figure 5 is almost the same with the fist and second column signals in the first row in Figure 4. The third column signal in every row in Figure 5 rotates some angle relatively the third column signal in the first row in Figure 4 which is an inherent indetermination for complex-valued blind source separation, but they have the same waveform. The separated signals have almost the same waveform as the source signals, so this shows that the proposed method is valid for communication signals.

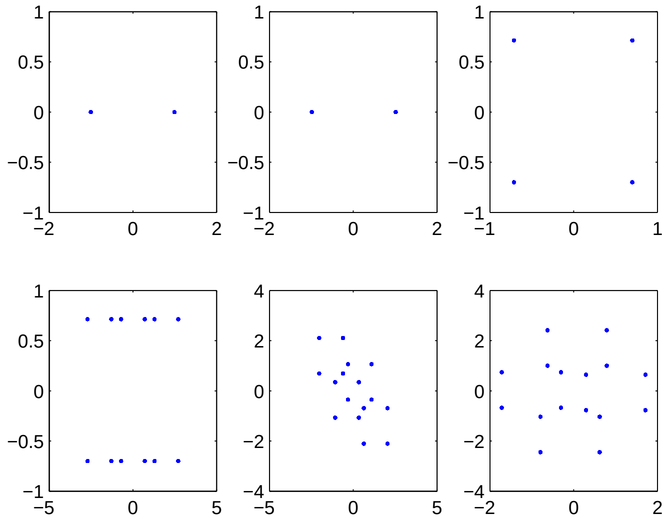

FIGURE 4. Original signals and mixing signals

Average convergence curves for the three methods without noise and with difference scale of noises are respectively shown in Figure 6 and Figure 7 from an average of 100 different simulation runs with a learning rate of 0.001. From Figure 6, we see that proposed method III starts to converge after 500 iterations, proposed method II after 1000 iterations, and proposed method I after 2000 iterations, while Yang starts to converge after 10000 iterations. Performance indices of SUT and Scott are both bigger than 0.1, which means they fail or have bad performance in separating mixing signals according to experience [28]. In this experiment, they can not separate the mixing signals successfully. This experiment shows that the proposed method is valid for communication signals and has a faster convergence speed than the other methods. To test the performance on noise signals, we add complex circular Gaussian noise with variance 0.05, 0.07 and 0.1. In Figure 6, the SUT and Scott failed in separating mixing signals, so we did not test the two methods in the following experiments. The convergence curves are shown in Figure 7. The first row, second and third row are the convergence curves with additive noise variance 0.05, 0.07 and 0.1 respectively in Figure 7. From the Figure 7, we can see that as the increasing of additive noise variance, the performance index at stationary point becomes large that the error is become large. Their performance index at stationary point is still smaller than 0.1, this shows that the proposed method is still validity for smaller noise signal.

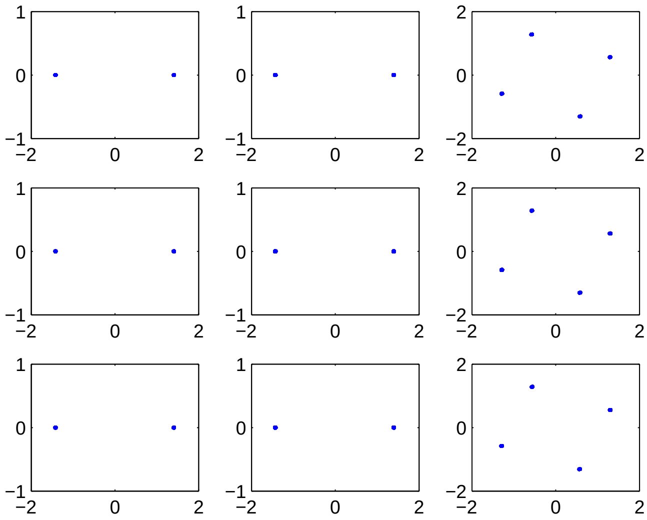

FIGURE 5. Separated signals

The proposed methods have two stages. The convergence curves shown in all figures are the convergence curves only for the second stage. For the above two experiments, in the first stage, after about 100 iterations, the whitening signal converges to unit matrix. Compared with other methods, the total iteration of proposed methods is still far less than other methods.

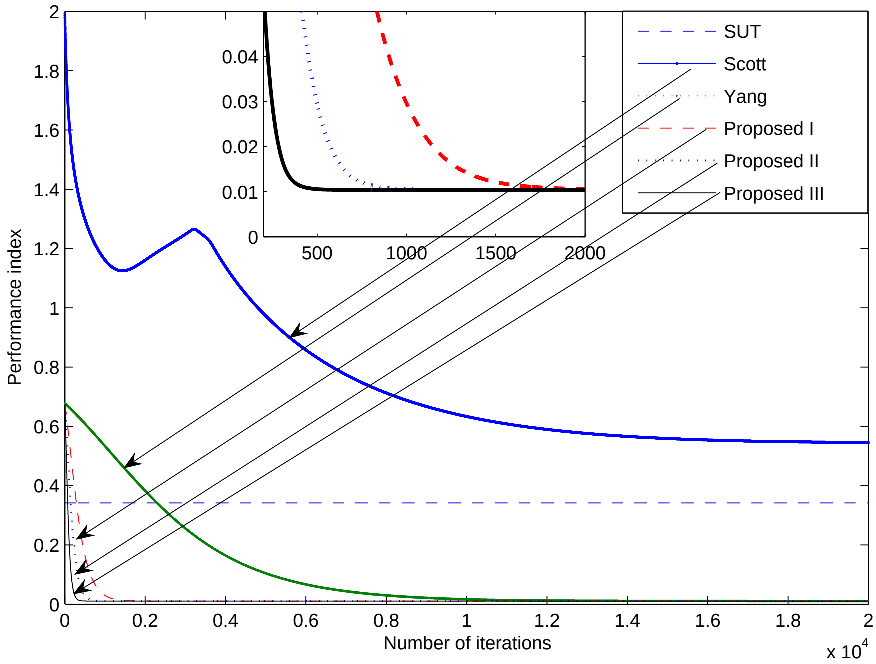

FIGURE 6. Average convergence curves with digital communication signals without noise

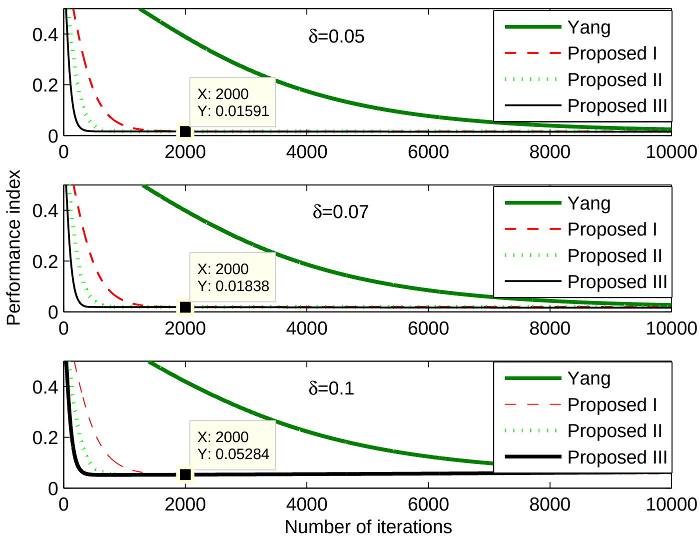

FIGURE 7. Average convergence curves with digital communication signals with different scale noises

6. **Conclusions.** This paper proposed a kind of adaptive complex-valued ICA method for noncircular signals based on covariance and pseudo-covariance matrices of noncircular original signals. The proposed methods have faster convergence speed and smaller error than other adaptive methods based on second order statistics. For different mixing source signals, the proposed methods still have better performance and faster convergence than the Scott method which is equivariant to the mixing system. For the digital communication signals used in this paper, the proposed methods are validity and still have faster convergence rate, while the SUT and Scott methods fail to recover the original signal.

REFERENCES

[1] J. F. Cardoso, and A souloumiac, Blind Beamforming for non-Gaussian signals, *Radar signal processing IEE proceedings F*, vol.140, no.6, pp.362–370,1994.

[2] H. Mehrabian, R. Chopra, and A. L. Martel, Calculation of intravascular signal in Dynamic contrast Enhanced- MRI using adaptive complex independent component analysis, *IEEE Transaction on medical imaging*, vol. 32, no.4, pp. 699–710,2013.

[3] F. Karimzadeh, S. Esmaeili, and S. H. Hosseinian, A novel method for noninvasive estimation of utility harmonic impedance based on complex independent component analysis, *IEEE Transactions on power delivery*, vol.30, no.4,pp.1843–1852,2015.

[4] G. B. Qian, L. P. Li,and M. G. Luo, On the blind channel identifiably of MIMO-STBC systems using noncircular complex FastICA algorithm, *Circuits, systems and signal processing*, vol.33, no.6, pp.1859–1881,2014.

[5] V.Laparra, U. Michael, G. J. Malo ,and A. Hyvarinen, Complex valued Independent component analysis of Natural Image, *Proceedings of the 21st international conference on Artificial neural networks*, Espoo, Finland,pp.213–220,2011.

[6] E. Bingham, and A. Hyv?rinen, 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.

[7] M. Novey, and T. Adali, On extending the complex FastICA algorithm to noncircular Sources, *IEEE Transactions on Signal Processing*, vol.56, no.5,pp.2148–2154,2008.

[8] M. Novey and T Adali, Complex ICA by negentropy maximization, *IEEE Trans. Neural Networks*, vol.19,no.4,pp.596–609,2008.

[9] G. S. Fu, R. Phlypo, M. Anderson, X. L. Li and T. Adali, An efficient entropy rate estimator for complex-valued signal processing: application to ICA, *2014 IEEE international conference on acoustic, speech and signal processing(ICASSP)*, Florence,pp.6216–6220,2014.

[10] G. S. Fu, R. Phlypo, M. Anderson and T. Adali, Complex independent component analysis using three types of diversity: Non-Gaussianity, Nonwhitenesss, and Noncircularity, *IEEE Transaction on signals processing*, vol.63,no.3,pp.794–805,2015.

[11] G. B. Qian, P. Wei and H. S. Liao, Efficient variant of noncircular complex fastICA algorithm for the blind source separation of digital communication signals, *Circuits systems signal processing*, vol.35,no.2, pp.1–13,2016.

[12] W. Zhao, Y. Shen, Z. G. Yuan, D. W. Liu, P. Xu, Y. Wei, W. Jian and N. Sha, A novel method for complex-valued signals in independent component analysis framework, *Circuits systems Signal processing*, vol.34, no.6, pp.1893–1913,2015.

[13] H. L. Li, and T. Adali, A class of complex ICA algorithms based on the kurtosis cost function, *IEEE transactions on neural networks*,vol.19, no.3,pp.408–420,2008.

[14] S. Javidi, D. P. Mandic, C. C. Took and A. Cichocki, Kurtosis-Based Blind Source Extraction of Complex Non-Circular Signals with Application in EEG Artifact Removal in Real-Time, *Frontiers in Neuroscience*, vol.5, Article 105, pp.1–18,2011.

[15] J.Voss, M. Belkin, and L. Rademacher, A pseudo-Euclidean Iteration for optimal recovery in noisy ICA, *eprint arXiv:1502.04148*,pp.1–17,2015.

[16] W. Rongjie, Z. Haifeng, Z. Yiju and C. Meiqian, A algorithm for adaptive complex blind source separation based on Newton update, *Acta electronic sinica*, vol.42, no.6, pp.1125–1131,2014.

[17] J. Ce, W. Yanru, and W. Xiaoyu, A new complex blind source separation algorithm based on standard kurtosis, *Journal of Northeastern University(Natural Science)*, vol.36,no.5,pp.614–617,2015.

[18] W. Zhao, W. Yimin, S. Yuehong, C. Yufan, Z. Yuan, X. Pengcheng and W. Jian, A efficient algorithm by kurtosis maximization in reference based framework, *Radioengineering*, vol.24, no.2,pp.544–551,2015.

[19] J. Eriksson, and V. Koivunen, Complex-valued ICA using second order statistics, *2004 IEEE workshop on machine learning for signal processing*,Sao Luis,Brazil, pp.183–192,2004.

[20] J. Eriksson, and V. Koivunen, Complex random vectors and ICA models: Identifiablility, Uniqueness and Separability, *IEEE transactions on information theory*, vol.52,no.3, pp.1–13,2006.

[21] C. D. Scott, J. Eriksson, and V. Koivunen, Equivariant algorithm for estimating the strong-uncorrelating transform in complex independent component analysis, *ICA2006*, Charleston,USA,pp.57–65,2006.

[22] C. D. Scott, J. Eriksson, and V. Koivunen, Adaptive estimation of the strong uncorrelating transform with applications to subspace tracking, *2006 IEEE International conference on acoustics, speech and signal processing*,Toulouse,vol.4,pp.IV 2006.

[23] S. Y.Yang, Z. Liquan, Z. Chaozhu, and J. Yanfei, Research on adaptive real-time blind separation algorithm of complex mixed signals, *Systems Engineering and Electronics*, vol.31, no.5, pp.1018–1021,2009.

[24] F.Cong, Q. H. Lin, P. Jia, X Shi, and T Ristaniemi, Second order impropriety based complex-valued algorithm for frequency-domain blind separation of convolutive speech mixtures, *2011 IEEE International workshop on machine learning for signal processing*, Santander, pp.1–6,2011

[25] H. shen, and M kleinsteuber, Algebraic solutions to complex blind source separation, *10th International Conference, LVA/ICA*, Tel Aviv, Israel,pp.74-81,2012.

[26] A. Yeredor, Performance analysis of the strong uncorrelating transformation in blind separation of complex valued sources, *IEEE transactions on signal processing*, vol.61, no.1, pp.478–483,2012.

[27] D. Ramirez, p. J. Schreier, J. Via, and I. Santamaria, Testing blind separability of complex Gaussian mixtures, *Signal Processing*,vol.96, pp.49–57,2014.

[28] A. Cichocki, and S. Amari, Adaptive blind signals and image processing, England:John Wiely and Sons, 2002.