PAPER Special Section on Blind Signal Processing: Independent Component Analysis and Signal Separation

Polar Coordinate based Nonlinear Function for Frequency-Domain Blind Source Separation

Hiroshi SAWADA[†], **Ryo MUKAI**[†], ***Members***, **Shoko ARAKI**[†], ***Nonmember***, ***and*** **Shoji MAKINO**[†], ***Member***

SUMMARY This paper discusses a nonlinear function for independent component analysis to process complex-valued signals in frequency-domain blind source separation. Conventionally, nonlinear functions based on the Cartesian coordinates are widely used. However, such functions have a convergence problem. In this paper, we propose a more appropriate nonlinear function that is based on the polar coordinates of a complex number. In addition, we show that the difference between the two types of functions arises from the assumed densities of independent components. Our discussion is supported by several experimental results for separating speech signals, which show that the polar type nonlinear functions behave better than the Cartesian type.

key words: *independent component analysis, blind source separation, frequency domain, complex-valued signal, polar coordinate, Cartesian coordinate, probability density function*

1. Introduction

Blind source separation (BSS) is a technique for estimating original source signals using only sensor observations that are mixtures of the original signals. If source signals are mutually independent and non-Gaussian (or non-stationary), we can employ techniques of independent component analysis (ICA) [1]–[6]. If the mixture is instantaneous, the situation is rather simple. In a real-world situation, however, signals are mixed in a convolutive manner with delay and reverberations, and longer reverberations make the BSS problem more difficult. In the convolutive case, a separating system typically consists of a matrix of filters, not just a matrix of scalars.

Methods for constructing such separating filters can be classified into two approaches. The first one is a time-domain approach, where the coefficients of the separating filters are calculated directly in the convolutive mixture model. The other is a frequency-domain approach [7]–[12], where the frequency responses of the separating filters are first calculated, and then the time-domain representation of the separating filters is obtained by applying an inverse DFT (discrete Fourier transform) to them.

This paper discusses the frequency-domain approach. It has an advantage in that ICA is applied to instantaneous mixtures, which are easier to solve than convolutive ones in the time domain. In frequency-domain BSS, instead, we have to deal with complex-valued signals. For this purpose, Smaragdis [8] proposed a complex-valued ICA algorithm, which was an extension of the information maximization approach [2]. The nonlinear function used in the extension was based on the Cartesian coordinates of a complex number: nonlinearities are applied to the real and imaginary parts separately. This type of nonlinearity has been widely used for complex-valued neural networks [13]–[15]. The BSS method proposed by Smaragdis actually works and is widely used by other researchers [7], [9], [11], [12]. However, no appropriate interpretation of the Cartesian type nonlinear function has been described. Moreover, it imposes an additional constraint that prevents a learning algorithm from converging unless a non-holonomic algorithm [16] is employed.

In this paper, we propose that a more appropriate nonlinear function for frequency-domain BSS is based on the polar coordinates of a complex number: nonlinearities are applied only to the amplitude. This type of nonlinearity has also been used for complex-valued neural networks [17], [18]. We derive the nonlinear function from the probability density function of frequency-domain signals that are assumed to be independent of the phase. When using the polar coordinate based function, there is no additional constraint mentioned above. We also provide an interpretation of the Cartesian coordinate based function. With experimental results for separating speech signals in a reverberant environment, we compare the behaviors of these two types of nonlinear function, and discuss the differences between them.

2. Blind Source Separation for Convolutive Mixtures

2.1 Problem Formulation

Suppose that P source signals $s_p(t)$ are mixed in an environment and observed at Q sensors

$$x_q(t) = \sum_{p=1}^{P} \sum_{k} h_{qp}(k) s_p(t-k),$$

where $h_{qp}(k)$ represents the impulse response from

Manuscript received June 28, 2002.
Manuscript revised October 8, 2002.

[†]The authors are with the NTT Communication Science Laboratories, NTT Corporation, 2-4 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0237, Japan

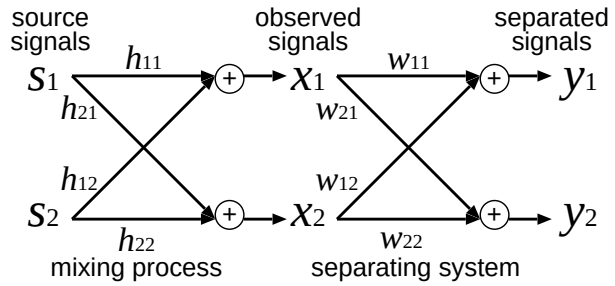


Fig. 1 BSS model

source p to sensor q. The set of impulse responses $h_{qp}(k)$ represents the mixing process. The goal of BSS is to obtain a separating system and also separated signals $y_1(t), \ldots, y_P(t)$ that are estimates of the source signals $s_1(t), \ldots, s_P(t)$. The separating system typically consists of a set of FIR filters $w_{rq}(k)$ that produces separated signals

$$y_r(t) = \sum_{q=1}^{Q} \sum_{k} w_{rq}(k) x_q(t-k).$$

The separation has to be accomplished without knowing the impulse responses $h_{qp}(k)$ or the information of the original source signals $s_p(t)$. If the source signals $s_p(t)$ are mutually independent, we can apply independent component analysis (ICA) to construct the separating system. Figure 1 shows a BSS model where $P = Q = 2$.

2.2 Frequency-Domain Approach

In the frequency-domain approach to constructing separating filters, the frequency responses $W_{rq}(\omega)$ of the separating filters are first calculated, and then the time-domain representation $w_{rq}(k)$ of the separating filters is obtained by applying an inverse DFT to them. The time and frequency representations of an FIR filter can be mutually converted by DFT and inverse DFT. The length L of an FIR filter $w_{rq}(k)$ corresponds to the resolution of the frequency response $W_{rq}(\omega)$.

By L-point windowed short time DFTs, time-domain signals $x_q(t)$ are converted into frequency-domain time-series signals

$$X_q(\omega, m) = \sum_{\tau=0}^{L-1} x_q(\tau + mS)\, w(\tau)\, e^{-j\omega\tau}, \qquad (1)$$

where $w(\tau)$ denotes a window function, S is a shifting interval of the window, and $\omega = 0, \frac{1}{L}2\pi, \ldots, \frac{L-1}{L}2\pi$. Now, we have $\mathbf{X}(\omega, m) = [X_1(\omega, m), \ldots, X_Q(\omega, m)]^T$ for each frequency ω. To obtain the frequency responses $W_{rq}(\omega)$, we solve an ICA problem

$$\mathbf{Y}(\omega, m) = \mathbf{W}(\omega)\mathbf{X}(\omega, m),$$

where $\mathbf{Y}(\omega, m) = [Y_1(\omega, m), \ldots, Y_P(\omega, m)]^T$, and $\mathbf{W}(\omega)$ is a $P \times Q$ matrix whose elements are $W_{rq}(\omega)$. $Y_r(\omega, m)$ is a frequency-domain representation of $y_r(t)$.

2.3 ICA Algorithm

Before explaining a complex-valued ICA, we review an ordinary real-valued ICA algorithm. Based on the information maximization approach [2], [3] combined with the natural gradient [4], a separating matrix $\mathbf{W}$ is gradually improved by the learning rule:

$$\Delta\mathbf{W} = \mu\,[\mathbf{I} - \langle\varphi(\mathbf{Y})\mathbf{Y}^T\rangle]\,\mathbf{W}.$$

In this formula, μ is a step size parameter that has an effect on the speed of convergence, $\langle\cdot\rangle$ denotes the averaging operator, and $\varphi(\cdot)$ is a nonlinear function defined as:

$$\begin{aligned} \varphi(\mathbf{Y}) &= [\varphi(Y_1), \ldots, \varphi(Y_N)]^T \\ \varphi(Y_i) &= -\frac{\partial}{\partial Y_i} \log p(Y_i) \end{aligned} \qquad (2)$$

where $p(Y_i)$ is the probability density function (pdf) of Y_i. If we assume $p(Y_i) = \alpha/\cosh^2(Y_i)$, then the function is hyperbolic tangent $\varphi(Y_i) = 2\tanh(Y_i)$, which is widely used for super-gaussian distributions [2], [3].

In frequency-domain BSS, signals obtained by DFT are complex. To deal with complex signals in ICA at each frequency, the calculation of $\Delta\mathbf{W}$ and the nonlinear function were extended [8]:

$$\begin{aligned} \Delta\mathbf{W} &= \mu\,[\mathbf{I} - \langle\Phi(\mathbf{Y})\mathbf{Y}^H\rangle]\,\mathbf{W} \\ \Phi(Y_i) &= \tanh[re(Y_i)] + j\cdot\tanh[im(Y_i)] \end{aligned} \qquad (3)$$

where $\mathbf{Y}^H$ represents the conjugate transpose of $\mathbf{Y}$, and $re(Y_i)$ and $im(Y_i)$ are the real and imaginary parts of Y_i, respectively. In the nonlinear function $\Phi(Y_i)$, $\tanh(\cdot)$ is applied separately in the real and imaginary parts. We call this type of function a Cartesian coordinate based function.

Although function (3) actually works, no appropriate interpretation of this function has been presented. Moreover, it sometimes has a convergence problem. Looking into the diagonal elements of $[\mathbf{I} - \langle\Phi(\mathbf{Y})\mathbf{Y}^H\rangle]$, we see that $\mathbf{W}$ converges to a point that satisfies

$$\langle\Phi(Y_i)Y_i^*\rangle = 1 \qquad (4)$$

where Y_i^* is the complex conjugate of Y_i. Extracting the real and imaginary parts of this equation, we have

$$\langle\tanh[re(Y_i)]re(Y_i)\rangle + \langle\tanh[im(Y_i)]im(Y_i)\rangle = 1, \quad (5)$$

$$\langle\tanh[im(Y_i)]re(Y_i)\rangle - \langle\tanh[re(Y_i)]im(Y_i)\rangle = 0, \quad (6)$$

respectively. Equation (5) makes the average amplitude of Y_i converge to some value. By contrast, Eq. (6) imposes an additional constraint that is satisfied when the two nonlinear correlations become zero or exactly

the same. We cannot find any implication useful for ICA of this constraint, and there are some cases where $\mathbf{W}$ does not converge well because of it. We show such a case in Sec. 5.

If we use the non-holonomic algorithm [16]:

$$\Delta\mathbf{W} = \mu\left[\mathrm{diag}(\langle\Phi(\mathbf{Y})\mathbf{Y}^H\rangle) - \langle\Phi(\mathbf{Y})\mathbf{Y}^H\rangle\right]\mathbf{W},$$

we can avoid constraint (6). Some researchers [9], [11] used this algorithm combined with a Cartesian coordinate based function (3). However, there is still another convergence problem, which is also shown in Sec. 5.

3. Polar Coordinate based Nonlinear Function

In this section, we derive an appropriate nonlinear function for frequency-domain BSS from the complex counterpart of Eq. (2):

$$\Phi(Y_i) = -\frac{\partial}{\partial Y_i}\log p(Y_i). \tag{7}$$

First, we make an assumption as regards the density $p(Y)$ of a complex-valued signal Y in the frequency domain.

Assumption 1: Let $Y = |Y|\,e^{j\cdot\theta(Y)}$ be a complex-valued signal. The pdf $p(Y)$ of Y is independent of the phase: $p(Y) = \alpha\cdot p(|Y|)$, where $p(|Y|)$ is the pdf of $|Y|$ and α is a constant.

This assumption is natural for a frequency-domain signal, since the phase of Y depends on the positions of the windows $w(\tau)$ of a windowed DFT (1) and the windows can be shifted arbitrarily.

Then, let us consider the derivative of a real-valued function $\log p(Y)$. Generally speaking, a real-valued function whose argument is a complex is not analytic: the derivative is not well-defined. Throughout this paper, we use the following definition of the derivative of a real-valued function.

Definition 1: Let $Y = Y_R + jY_I$ be a complex and $f(Y)$ be a real-valued function: $\mathbb{C}\to\mathbb{R}$. We define the derivative as:

$$\frac{\partial f(Y)}{\partial Y} \stackrel{\mathrm{def}}{=} \frac{\partial f(Y)}{\partial Y_R} + j\frac{\partial f(Y)}{\partial Y_I}.$$

The relevance of this definition lies in the fact that the result points to a direction in which $f(Y)$ increases.

Now, we are ready to derive an appropriate nonlinear function for frequency-domain BSS.

Theorem 1: Taking Assumption 1, $\Phi(Y)$ in (7) is

$$\begin{aligned}\Phi(Y) &= \varphi(|Y|)\,e^{j\cdot\theta(Y)},\\ \text{where } \varphi(|Y|) &= -\frac{\partial}{\partial|Y|}\log p(|Y|)\end{aligned}$$

Proof: From (7) and Assumption 1, we have

$$\begin{aligned}\Phi(Y) &= -\frac{\partial}{\partial Y}\log\alpha\cdot p(|Y|)\\ &= -\frac{\partial}{\partial|Y|}\log p(|Y|)\frac{\partial|Y|}{\partial Y} = \varphi(|Y|)\frac{\partial|Y|}{\partial Y}.\end{aligned}$$

And from Definition 1, we have

$$\begin{aligned}\frac{\partial|Y|}{\partial Y} &= (\frac{\partial}{\partial Y_R} + j\frac{\partial}{\partial Y_I})\sqrt{Y_R^2+Y_I^2}\\ &= \frac{1}{2}\frac{2Y_R}{\sqrt{Y_R^2+Y_I^2}} + j\frac{1}{2}\frac{2Y_I}{\sqrt{Y_R^2+Y_I^2}} = e^{j\cdot\theta(Y)}.\end{aligned}$$

This proves the theorem.

Here, we have a nonlinear function based on the polar coordinates of a complex number. In this type of nonlinear function, nonlinearity is applied only to the amplitude and the phase is preserved. By using this type, constraint (6) does not appear. Since Y_i^* is a complex conjugate of Y_i,

$$\Phi(Y_i)Y_i^* = \varphi(|Y_i|)\,e^{j\cdot\theta(Y_i)}\,|Y_i|\,e^{-j\cdot\theta(Y_i)} = \varphi(|Y_i|)\,|Y_i|.$$

Hence, the imaginary part of (4) becomes 0.

If we assume a super-gaussian distribution $p(|Y_i|) = \alpha/\cosh(|Y_i|)$, the corresponding nonlinear function becomes

$$\Phi(Y_i) = \tanh(|Y_i|)\,e^{j\cdot\theta(Y_i)}. \tag{8}$$

Based on the discussion so far, the nonlinear function (8) is more appropriate than (3) for separating super-gaussian signals in the frequency domain.

4. Interpretation of the Cartesian Coordinate based Function

Then, what kind of distribution leads to the Cartesian coordinate based function (3)? The next theorem provides an interpretation of the Cartesian type nonlinear function.

Theorem 2: If the density of a complex-valued signal Y is assumed to be $p(Y) = p(Y_R)\cdot p(Y_I)$, then Eq. (7) becomes

$$\begin{aligned}\Phi(Y) &= \varphi(Y_R) + j\cdot\varphi(Y_I),\\ \text{where } \varphi(Y_R) &= -\frac{\partial}{\partial Y_R}\log p(Y_R),\\ \varphi(Y_I) &= -\frac{\partial}{\partial Y_I}\log p(Y_I).\end{aligned}$$

Proof:

$$\begin{aligned}\Phi(Y) &= -\frac{\partial}{\partial Y_R}\log p(Y) - j\cdot\frac{\partial}{\partial Y_I}\log p(Y)\\ &= -\frac{\partial}{\partial Y_R}\log p(Y_R) - j\cdot\frac{\partial}{\partial Y_I}\log p(Y_I).\end{aligned}$$

Table 1 Experimental conditions

Direction of sources	$-30°$ and $40°$ (two sources)
Distance of two microphones	4 cm
Length of source signal	3 seconds, 6 seconds
Reverberation time	$T_R = 300$ ms
Sampling rate	8kHz
Window function	$w(\tau)$: Hanning
Window length	$L = 2048$ points (256 ms)
Shifting interval	$S = 1024 \sim 32$ points
Step size	$\mu = 0.2$
Gain parameter	$\eta = 100$
Number of iterations	100

This theorem states that a Cartesian coordinate based function is appropriate if Y_R and Y_I are mutually independent. We notice that if the real and imaginary parts of the same complex signal are mutually independent, the additional constraint (6) is satisfied since both nonlinear correlations become zero. By this theorem, one of the assumed densities of the nonlinearity (3) is $p(Y) = \alpha_R / \cosh(Y_R) \cdot \alpha_I / \cosh(Y_I)$, which contradicts the Assumption 1.

5. Experiments and Discussions

We have discussed the theoretical aspects of the two types of nonlinear function in Secs. 3 and 4. Now, in this section, we discuss their practical aspects with some experimental results.

5.1 Experimental Conditions

To compare the two types of nonlinear function, we conducted experiments to separate speech signals. We used the following two nonlinear functions:

$$\begin{array}{ll} \mathsf{Polar} & \Phi(Y) = \tanh(\eta|Y|)\, e^{j\cdot\theta(Y)} \\ \mathsf{Cartesian} & \Phi(Y) = \tanh(\eta Y_R) + j \cdot \tanh(\eta Y_I), \end{array}$$

where η is a gain parameter to control the nonlinearity. And we used the following two gradients of $\mathbf{W}$ to examine the effect of the additional constraint (6):

$$\begin{array}{ll} \mathsf{I} & \Delta\mathbf{W} = \mu\,[\mathbf{I} - \langle\Phi(\mathbf{Y})\mathbf{Y}^H\rangle]\,\mathbf{W} \\ \mathsf{Diag} & \Delta\mathbf{W} = \mu\,[\mathrm{diag}(\langle\Phi(\mathbf{Y})\mathbf{Y}^H\rangle) - \langle\Phi(\mathbf{Y})\mathbf{Y}^H\rangle]\,\mathbf{W}. \end{array}$$

The other conditions are summarized in Table 1.

Before applying ICA, frequency-domain observed signals $\mathbf{X}$ are sphered so that they become uncorrelated and have unit variances. This pre-process was very important in terms of making the ICA algorithm stable especially for the Diag cases, where Eq. (4) is not concerned. Without this process, the Diag cases could have exhibited irregular convergence speeds among the different frequency bins.

We measured the BSS performance from the average of SNRs (signal-to-noise ratios) at two outputs. Since we generated mixed signals by convolving source signals with impulse responses, we were able to decompose a mixed signal by $x_q(t) = \sum_{p=1}^{P} x_{qp}(t)$, where

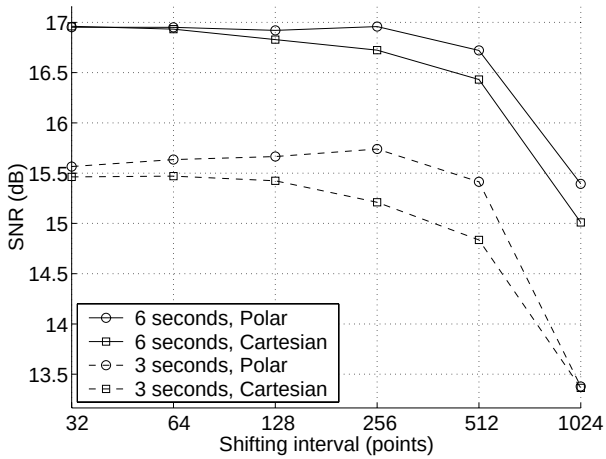


Fig. 2 Average SNRs with different shifting intervals

$x_{qp}(t) = \sum_k h_{qp}(k) s_p(t-k)$.

Using this decomposition, we could also decompose a separated signal by $y_r(t) = \sum_{p=1}^{P} y_{rp}(t)$, where

$y_{rp}(t) = \sum_{q=1}^{Q} \sum_k w_{rq}(k) x_{qp}(t-k)$.

In the SNR calculation at output r, we treated $y_{rr}(t)$ as a signal and $y_r(t) - y_{rr}(t)$ as a noise. Therefore, the SNR was calculated by

$10\log[\sum_t y_{rr}(t)^2] - 10\log[\sum_t \{y_r(t) - y_{rr}(t)\}^2]$.

To compare the results as accurately as possible, we avoided the influence of the permutation problem [9], [11], [12] of frequency-domain BSS. We selected the best permutation by calculating the SNR in each frequency bin. Therefore, the results are ideal under the condition that the permutation problem is solved perfectly. The SNR calculation was performed in the same manner as that described above except that it was in the frequency domain. Let $X_{qp}(\omega, m)$ be the result of a windowed short time DFT for $x_{qp}(t)$. Then, we were able to decompose a separated signal by $Y_r(\omega, m) = \sum_{p=1}^{P} Y_{rp}(\omega, m)$, where

$Y_{rp}(\omega, m) = \sum_{q=1}^{Q} W_{rq}(\omega) X_{qp}(\omega, m)$.

5.2 Comparison of Separation Performance

We have performed experiments under various conditions, and discovered that the difference between Polar and Cartesian performance depends on the shifting interval of the window in the short time DFT (1). Figure 2 shows the situation. Each plot represents the average SNR for 24 combinations (12 pairs of speech signals and 2 gradients). Generally, a shorter shifting interval improves the separation performance because the number of data samples used in ICA increases. However,

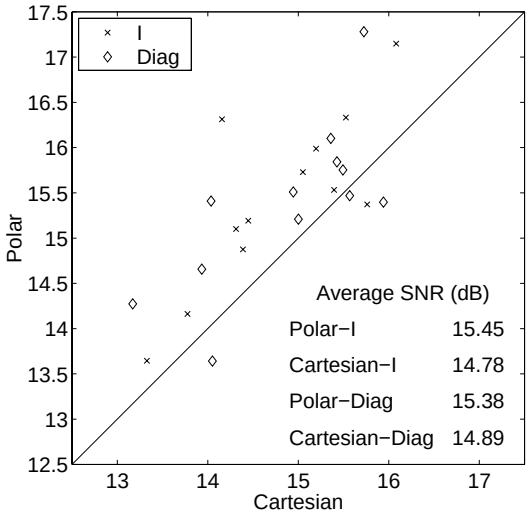


Fig. 3 Comparison of Polar and Cartesian SNRs (source signal length: 3 seconds, shifting interval: 512 points)

the improvements are saturated at some point since too short a shifting interval simply results in redundant data samples. We see that improvements are saturated rapidly (at 256 points) in the Polar case, but slowly (at around 64 points) in the Cartesian case. We also observe that the advantage of Polar becomes less significant as the shifting interval decreases.

Figure 3 shows the results when the source signals were 3 seconds long and the shifting interval was 512 points. This was a case where the difference between the Polar and Cartesian performance was fairly large. Each plot represents the Polar and Cartesian results with the same gradient and with the same combination of speech signals. We see that the Polar result is better than the Cartesian result in most cases, and the selection of gradients (I or Diag) does not greatly affect the separation performance.

5.3 Comparison of Convergence

This subsection discusses the convergence when the source signals were 3 seconds long and the shifting interval was 512 points. First, we discuss the additional constraint (6) imposed in the Cartesian–I case. Figure 4 shows the values of $[\mathbf{I} - \langle \Phi(\mathbf{Y})\mathbf{Y}^H \rangle]$ at some frequency bin. The horizontal axis corresponds to the number of iterations. The first graph shows the absolute values of each element for the Cartesian–I case. We see oscillations that hinder convergence. They come from the imaginary parts of the diagonals as shown in the second graph. If we use a polar coordinate based function, we can eliminate such oscillations as discussed in Sec. 3. The third graph shows the Polar–I case. We can see a smooth convergence. Clearly, the mutual information among $\mathbf{Y}$ is well minimized in this case unlike in the Cartesian–I case.

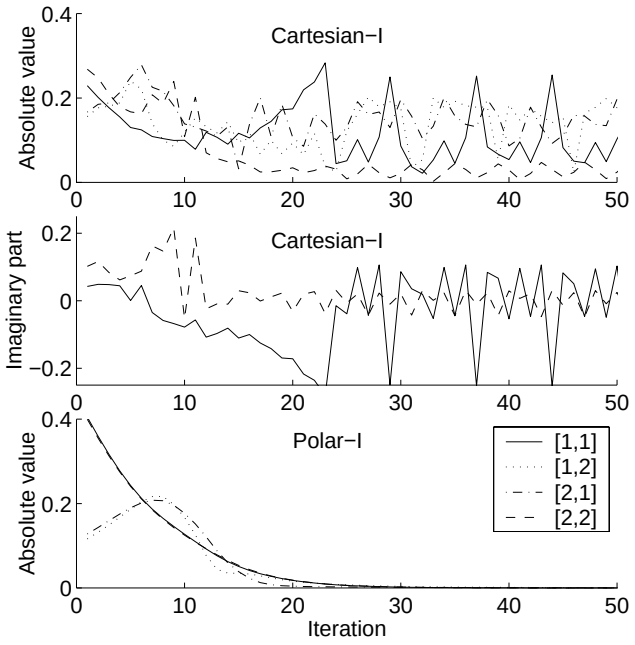


Fig. 4 Values of $[\mathbf{I} - \langle \Phi(\mathbf{Y})\mathbf{Y}^H \rangle]$

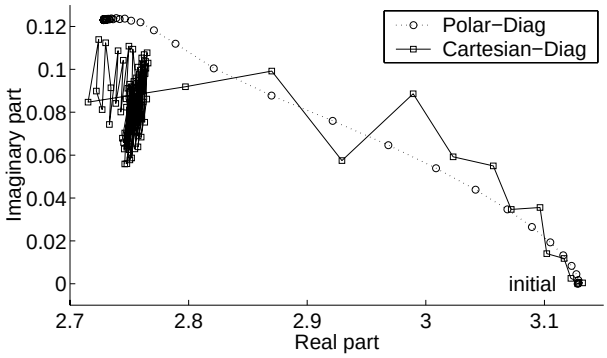


Fig. 5 Trajectory of W_{11} on complex plane

If we use Diag as the calculation of $\Delta\mathbf{W}$, we can eliminate the additional constraint (6) even in the Cartesian case. Accordingly, by investigating the Diag results, we can see the differences that arise purely from the difference between the Polar and Cartesian nonlinearities. In fact, we found another convergence problem in a Cartesian–Diag case. Figure 5 shows the trajectory of element W_{11} of $\mathbf{W}$ at some frequency bin. We see that the direction of the movement changes gradually in Polar–Diag, whereas it changes sharply and frequently in Cartesian–Diag. The difference comes from the assumed densities, as discussed in Secs. 3 and 4. Figure 6 shows the contour and gradient of $-\log p(Y)$, being $p(Y) = \alpha/\cosh(\eta\,|Y|)$ in Polar, and $p(Y) = \alpha/[\cosh(\eta\,Y_R) \cdot \cosh(\eta\,Y_I)]$ in Cartesian. The gradient corresponds to $\Phi(Y)$. We see that the direction of the gradient changes smoothly in the Polar case, whereas it changes steeply near the vertical and hori-

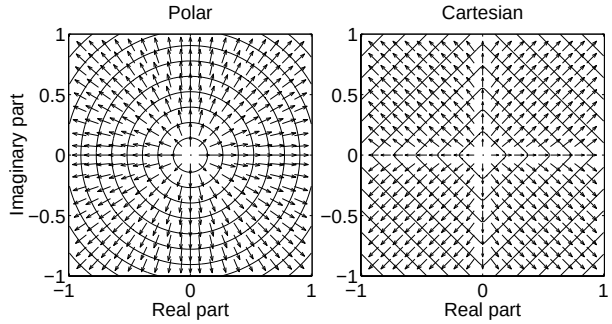


Fig. 6 Contour and gradient of $-\log p(Y)$

zontal axes in the Cartesian case. We consider that the jag in Fig. 5 comes from this steepness. However, it may be smoothed out by the averaging operator $\langle \Phi(Y)Y^* \rangle$ if we increase the number of samples. This is partly why the advantage of Polar becomes less significant as the shifting interval decreases as shown in Fig. 2.

6. Conclusions

We proposed a polar coordinate based nonlinear function to process complex-valued signals in ICA. Compared with the Cartesian coordinate based function, the main difference arises from the assumed densities of the independent signals. The assumption for the polar coordinate based function is that the densities are phase independent. It is more natural than the assumption for the Cartesian coordinate based function since signals are produced by a windowed DFT. With several experiments, we examined the advantages of the Polar type in a practical situation. If data samples were not so redundant, the difference between the Polar and Cartesian performance was fairly large and there were some convergence problems in the Cartesian case.

Acknowledgement

We thank Dr. Hiroshi Saruwatari for valuable discussions and providing us with a set of impulse responses, and Dr. Shigeru Katagiri for continuous encouragement.

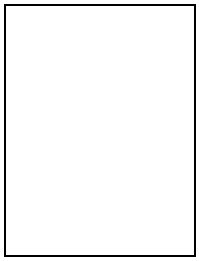
References

[1] P. Comon, "Independent component analysis, a new concept?," *Signal Processing*, vol. 36, pp. 287–314, 1994.
[2] A. Bell and T. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution," *Neural Computation*, vol. 7, no. 6, pp. 1129–1159, 1995.
[3] T. W. Lee, *Independent Component Analysis - Theory and Applications*, Kluwer Academic Publishers, 1998.
[4] S. Amari, "Natural gradient works efficiently in learning," *Neural Computation*, vol. 10, no. 2, pp. 251–276, 1998.
[5] S. Haykin, Ed., *Unsupervised Adaptive Filtering*, John Wiley & Sons, 2000.
[6] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, John Wiley & Sons, 2001.
[7] A. D. Back and A. C. Tsoi, "Blind deconvolution of signals using a complex recurrent network," in *Proc. Neural Networks for Signal Processing*, 1994, pp. 565–574.
[8] P. Smaragdis, "Blind separation of convolved mixtures in the frequency domain," *Neurocomputing*, vol. 22, pp. 21–34, 1998.
[9] S. Ikeda and N. Murata, "A method of ICA in time–frequency domain," in *Proc. ICA '99*, Jan. 1999, pp. 365–370.
[10] L. Parra and C. Spence, "Convolutive blind separation of non-stationary sources," *IEEE Trans. Speech Audio Processing*, vol. 8, no. 3, pp. 320–327, May 2000.
[11] S. Kurita, H. Saruwatari, S. Kajita, K. Takeda, and F. Itakura, "Evaluation of blind signal separation method using directivity pattern under reverberant conditions," in *Proc. ICASSP 2000*, June 2000, pp. 3140–3143.
[12] F. Asano, S. Ikeda, M. Ogawa, H. Asoh, and N. Kitawaki, "A combined approach of array processing and independent component analysis for blind separation of acoustic signals," in *Proc. ICASSP 2001*, May 2001, pp. 2729–2732.
[13] H. Leung and S. Haykin, "The complex backpropagation algorithm," *IEEE Trans. Signal Processing*, vol. 39, no. 9, pp. 2101–2104, 1991.
[14] N. Benvenuto and F. Piazza, "On the complex backpropagation algorithm," *IEEE Trans. Signal Processing*, vol. 40, no. 4, pp. 967–969, 1992.
[15] T. Nitta, "An extension of the back-propagation algorithm to complex numbers," *Neural Networks*, vol. 10, no. 8, pp. 1392–1415, 1997.
[16] S. Amari, T. P. Chen, and A. Cichocki, "Nonholonomic orthogonal learning algorithm for blind source separation," *Neural Computation*, vol. 12, no. 6, pp. 1463–1484, 2000.
[17] G. M. Georgiou and C. Koutsougeras, "Complex domain backpropagation," *IEEE Trans. Circuits and Systems II*, vol. 39, no. 5, pp. 330–334, 1992.
[18] A. Hirose, "Continuous complex-valued back-propagation learning," *Electronics Letters*, vol. 28, no. 20, pp. 1854–1855, 1992.

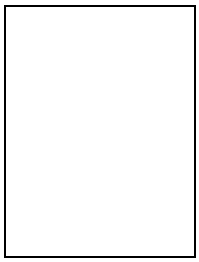
Hiroshi Sawada received the B.E., M.E. and Ph.D. degrees in information science from Kyoto University, Kyoto, Japan, in 1991, 1993 and 2001, respectively. In 1993, he joined NTT Communication Science Laboratories. From 1993 to 2000, he was engaged in research on the computer aided design of digital systems, logic synthesis, and computer architecture. Since 2000, he has been engaged in research on signal processing and independent component analysis for blind source separation. He is a member of the IEEE and the Acoustical Society of Japan (ASJ).

Ryo Mukai received the B.S. and M.S. degrees in information science from the University of Tokyo, Tokyo, Japan, in 1990 and 1992, respectively. He joined NTT in 1992. From 1992 to 2000, he was engaged in the research and develop-

ment of processor architecture for network service systems and distributed network systems. Since 2000, he has been with NTT Communication Science Laboratories, where he is engaged in research on blind source separation. His current research interests include digital signal processing and its applications. He is a member of the IEEE, ACM, the Acoustical Society of Japan (ASJ), and the Information Processing Society of Japan (IPSJ).



Shoko Araki received the B.E. and M.E. degrees in mathematical engineering and information physics from the University of Tokyo, Tokyo, Japan, in 1998 and 2000, respectively. In 2000, she joined NTT Communication Science Laboratories. Her research interests include digital signal processing, array signal processing and blind source separation for speech. She is a member of the IEEE and the Acoustical Society of Japan (ASJ).



Shoji Makino received the B.E., M.E. and Ph.D. degrees from Tohoku University, Sendai, Japan, in 1979, 1981 and 1993, respectively. He joined the Electrical Communication Laboratory of NTT in 1981. Since then, he has been engaged in the research and development of acoustic echo cancellation and adaptive algorithms. He is now a Senior Research Scientist, Supervisor, and Group Leader at the Speech Open Laboratory of the NTT Communication Science Laboratories. His research interests include blind source separation of convolutive mixtures of speech, acoustic signal processing, and adaptive filtering and its applications. He is a Senior Member of the IEEE, and a member the Acoustical Society of Japan (ASJ). He received the IEICE Paper Award in 2002, the ASJ Paper Award in 2002, the IEICE Achievement Award in 1997, and the ASJ Outstanding Technological Development Award in 1995. He is the author or co-author of more than 170 articles in journals and conference proceedings and has been responsible for more than 140 patents.