IEICE Proceeding Series

# Performance of Nonlinear BSS by PSO structure

Takuya Kurihara, Kenya Jin'no

Vol. 2 pp. 244-247
Publication Date: 2014/03/18
Online ISSN: 2188-5079

2013 International Symposium on Nonlinear Theory and its Applications
NOLTA2013, Santa Fe, USA, September 8-11, 2013

# Performance of Nonlinear BSS by PSO structure

Takuya Kurihara† and Kenya Jin'no‡

†Graduate school of Electrical Engineering, Nippon Institute of Technology
‡Department of Electrical and Electronics Engineering, Nippon Institute of Technology
4–1 Gakuendai, Minami-Saitama, Saitama 345-8501, Japan
Email: e1072215@estu.nit.ac.jp, jinno@nit.ac.jp

**Abstract**—Blind source separation (BSS) is a technique for recovering an original source signal from mixing signals without the aid of information of the source signal. In this study, we consider the case where the original signals are nonlinearly mixed. In order to solve such problem, we apply a radial basis function (RBF) network to the nonlinear BSS system. The inverse mapping of the nonlinear mixture system is approximated by the RBF network. For the system to be able to approximate the inverse mapping, it is necessary to learn the parameter of the RBF network. We suppose the original source signals are independent. In this case, if the mixture signals can be separated, the higher order cross-moment of the output signals are decreased. Particle swarm optimization is used for the learning algorithm. By using a numerical simulation, we conform the performance of the signal separation ability of the proposed system. Simulation results indicate that the proposed approach has good performance.

## 1. Introduction

Blind source separation (BSS) is a technique for recovering an original source signal from mixing signals without the aid of information of the source signal. Almost of the BSS are realized by using an independent component analysis (ICA) method[1]. The ICA supposes that each signal is independent and the signal has non-Gaussianity property. The BSS technique is applied to the sound signal processing, EEG, MEG[2], and so on.

In general, the BSS supposes that the mixture signal is composed with a linear combination of unknown independent signals. However, we can consider the nonlinear mixture signal case. Such problems are called as a nonlinear BSS. There are some processing techniques to solve the nonlinear BSS[3][4][5]. The gradient method is generally applied to the learning procedure of these methods. On the other hand, we apply a particle swarm optimization algorithm to the learning algorithm of the nonlinear BSS by using an RBF network.

## 2. Nonlinear mixing system

A nonlinear mixture model can be described as

$$\boldsymbol{x}(t) = \mathbf{f}(\boldsymbol{s}(t)) \tag{1}$$

where $\boldsymbol{x}(t) = (x_1(t), \cdots, x_n(t))^T$ is an observed signal vector, $\boldsymbol{s}(t) = (s_1(t), \cdots, s_n(t))^T$ is a source signal vector generated from an independent source signal, superscript $T$ denotes the transposition, and $\mathbf{f}$ is unknown mixing function mapped from $\mathbb{R}^n$ to $\mathbb{R}^n$.

A nonlinear separating system can be written as

$$\boldsymbol{y}(t) = \mathbf{g}(\boldsymbol{x}(t), \boldsymbol{\theta}) \tag{2}$$

where $\boldsymbol{y}(t) = (y_1(t), \cdots, y_n(t))^T$ is a separated signal vector, $\mathbf{g}$ is a separating function mapped from $\mathbb{R}^n$ to $\mathbb{R}^n$, and $\boldsymbol{\theta}$ is a parameter vector of the separating function $\mathbf{g}$.

The purpose of the nonlinear BSS is the searching of the map $\mathbf{g}$. A typical solution is the inverse of the mixing function $\mathbf{f}(\cdot)$. However, we cannot calculate the inverse function because the mixing function is unknown. In order to approximate the inverse function, the ICA algorithm is one of the solutions. The ICA algorithm supposes each source signal is independent. It is the problem of finding a mapping $\mathbf{g}$ that is independent of the output signals $\boldsymbol{y}(t)$.

If the mapping is a linear transformation, the ICA can solve the problem because it has the uniqueness of solutions. On the other hand, the nonlinear mapping does not identify if the solution space has a restriction[6].

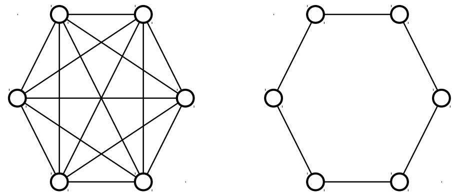
## 3. Radial Basis Function Network

Radial basis function (RBF) network approximates a nonlinear mapping[7]. Therefore, the inverse mapping of the nonlinear mixture system can be approximated by the RBF network. In this paper, we consider the nonlinear BSS system by using the RBF network. The RBF network consists of three layers; an input layer, a hidden layer, and an output layer. The unit of the hidden layer employs an RBF unit. The output of the output layer is a linear combination of the RBF units.

The RBF network is described by

$$y_i = \beta_i + \sum_{j=1}^{m} \alpha_{ij} \mathrm{K}_j(\boldsymbol{x}), \ \mathrm{K}_j(\boldsymbol{x}) = \exp\left(\frac{\|\boldsymbol{x} - \boldsymbol{\mu}_j\|^2}{\sigma_j^2}\right) \tag{3}$$

where, $\mathrm{K}_j(\boldsymbol{x})$ denotes a radial function and $\boldsymbol{x}$ represents the distance from the origin. $\alpha_{ij}$ is an weight coefficient between the $j$-th radial function and the $i$-th output, and $\beta_i$ is the $i$-th offset.



(a) full connection (b) ring connection

Figure 1: Network structures

We can select $K_j(\boldsymbol{x})$ from some functions. In this paper, we apply Gaussian function as the function $K_j(\boldsymbol{x})$. $\boldsymbol{\mu}_j$ denotes the $j$-th center vector, $\sigma_j$ is a parameter of the $j$-th RBF expanse.

## 4. Particle Swarm Optimization

Particle swarm optimization (PSO) is one of the metaheuristic algorithms for optimization problems[8][9].

The dynamics of PSO algorithm is described as

$$\begin{cases} \boldsymbol{v}_j^{t+1} = w\boldsymbol{v}_j^t + c_1 r_1 \left(\boldsymbol{pbest}_j^t - \boldsymbol{x}_j^t\right) \\ \qquad\qquad +c_2 r_2 \left(\boldsymbol{lbest}_j^t - \boldsymbol{x}_j^t\right) \\ \boldsymbol{x}_j^{t+1} = \boldsymbol{x}_j^t + \boldsymbol{v}_j^{t+1} \end{cases} \tag{4}$$

$x_j^t$ denotes a location vector of the $j$-th particle on the $t$-th iteration in the $N$-dimensional space, and $v_j^t$ denotes a velocity vector of the $j$-th particle on the $t$-th iteration. The PSO has three parameters, $w$, $c_1$, and $c_2$. $w$ is an inertia weight coefficient, $c_1$, and $c_2$ are acceleration coefficients. $r_1$, and $r_2$ are two separately generated uniformly distributed random numbers in the range $[0, 1]$. $\boldsymbol{pbest}_j^t$ means the location vector that gives the best value of the evaluation function of the $t$-th iteration. $\boldsymbol{lbest}_j^t$ means the location vector that gives the best value of the evaluation function on the $t$-th iteration in neighborhood of the $j$-th particle. The neighborhood is determined by the network topology of the particles. Figure 1 shows the two network topologies. Figure 1(a) is full connection topology which corresponds to the conventional PSO. Figure 1(b) is ring connection topology. Comparing with these structures, the full connection exhibits quick convergence. On the other hand, the ring structure exhibits remarkable search performance for multi-modal function[9].

## 5. Nonlinear BSS System

The our proposed nonlinear BSS system is shown in Fig. 2. The system consists of two parts; a whitening operation and the RBF network. The parameters $\sigma$, $\mu$, and $\alpha$ of the RBF network are adjusted by using the PSO algorithm. $\beta$ is set as the average of the separation signal becomes 0.

The whitening operation is that the variance of the signal becomes 1 and the covariance becomes 0. If the source signals are independent, the cross correlation is 0. Therefore, the whitening operation is very important for the ICA.

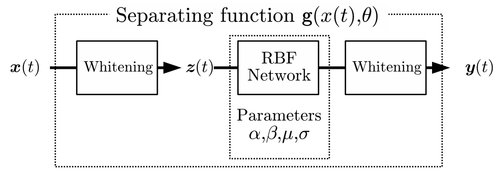



Figure 2: The separation system

The cross moment is applied to evaluate the learning. The cross moment measures the independence of the output signals. We apply the following evaluation function based on the cross moment.

$$C(\boldsymbol{y};\theta) = \sum_{i_1=1}^{k}\cdots\sum_{i_n=1}^{k}\left(E\left[y_1^{i_1}\cdots y_n^{i_n}\right] - E\left[y_1^{i_1}\right]\cdots\left[y_n^{i_n}\right]\right)^2 \tag{5}$$

where $\boldsymbol{y};\theta$ means the generated signal $\boldsymbol{y}$ depending on the parameter vector $\theta$. $k$ denotes the maximum order for the evaluation. If the stochastic variables are independent, the value of Eq. (5) becomes 0.

## 6. Simulation

In this section, we carry out some numerical simulations to separate the blind sources by using the proposed procedure. In order to evaluate the separation ability, a mean square error is applied. The mean square error is derived from two signals, $\boldsymbol{s}(t)$ and $\boldsymbol{y}(t)$.

$$mse = \frac{1}{n}\sum_{i=1}^{n}\min_j e_{i,j},\ e_{i,j} = \mathrm{E}[s_i(t)^2] - \frac{\mathrm{E}[s_i(t)y_j(t)]^2}{\mathrm{E}[y_j(t)^2]} \tag{6}$$

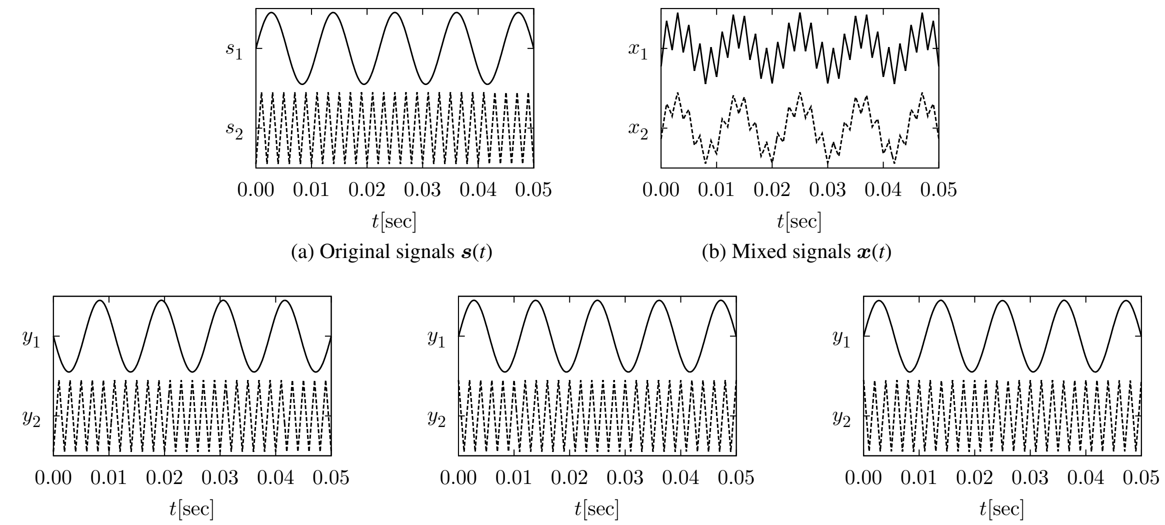
*Example 1*: Linear Mixture Case

We consider a two-channel linear mixture system. The following linear mixture system is applied.

$$\boldsymbol{x}(t) = B_0\boldsymbol{s}(t),\ B_0 = \begin{bmatrix} 0.5 & 0.5 \\ 0.7 & 0.3 \end{bmatrix} \tag{7}$$

The source signals consist of a sinusoidal signal and a triangle waveform; i.e., $\boldsymbol{s}(t) = [\sin(2\pi 90t),\ \mathrm{tri}(t,\tau)]^T$. where $\mathrm{tri}(t,\tau)$ denotes a periodic triangle wave with period $\tau = 0.002$. Figure 3(a) illustrates the source signals $\boldsymbol{s}(t)$, and Fig. 3(b) illustrates the mixture signals $\boldsymbol{x}(t)$.

In order to compare the performance, we carry out the numerical simulations by using the standard PSO method, the Ring-PSO method and the steepest descent method. The standard PSO has the full-connection structure. For the numerical simulations, the same initial values are applied. The simulation conditions of the PSO method are shown in Table 1. Moreover, the number of particles is 40. Figures 3(c), 3(d) and 3(e) show the signals separated by the standard PSO, the Ring-PSO and the steepest descent method. The evaluation value and its mean squared error of the signals $\boldsymbol{s}(t)$ and $\boldsymbol{x}(t)$ are shown in Table 2. The evaluation value and its mean squared error at each simulation




(a) Original signals $\boldsymbol{s}(t)$

(b) Mixed signals $\boldsymbol{x}(t)$

(c) Separated signals $\boldsymbol{y}(t)$ by the standard PSO

(d) Separated signals $\boldsymbol{y}(t)$ by the Ring-PSO

(e) Separated signals $\boldsymbol{y}(t)$ by the steepest descent method

Figure 3: The signals in linear mixture case

Table 1: Simulation parameters for linear mixture case

| | |
|---|---|
| The number of signal | 2 |
| The number of hidden neuron | 9 |
| The number of maximum iteration | 2000 |
| The number of trial | 50 |
| The inertia weight coefficient $w$ | $w = 0.7298$ |
| The acceleration coefficient $c_1$, $c_2$ | $c_1 = c_2 = 1.49$ |

Table 2: Evaluation value and its mean square error in linear mixture case

| | fitness | *mse* |
|---|---|---|
| The original signals $\boldsymbol{s}(t)$ | 0.000 | 0 |
| The mixed signals $\boldsymbol{x}(t)$ | 0.089 | 0.129 |

Table 3: Simulation results in linear mixture case

| | fitness | | *mse* | |
|---|---|---|---|---|
| | average | min | average | min |
| standard PSO | 0.000 | 0.000 | 0.004 | 0.000 |
| Ring-PSO | 0.000 | 0.000 | 0.011 | 0.000 |
| gradient method | 20.36 | 0.000 | 0.163 | 0.000 |

result are shown in Table 3. In these cases, all trials are success to separate the signals. These results indicate that the proposed method has the ability to separate the linear mixed signals.

*Example 2*: Nonlinear Mixture Case

Next, we consider a two-channel nonlinear mixture with sigmoid function. The following nonlinear mixture system is applied.

$$\boldsymbol{x}(t) = B_2 \tanh\left(B_1 \boldsymbol{s}(t)\right) \tag{8}$$

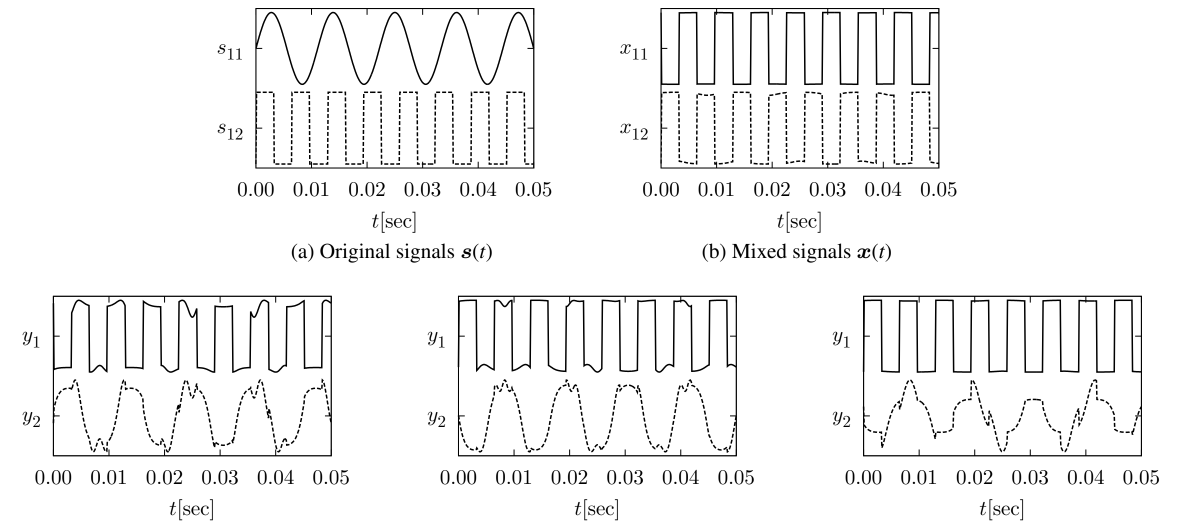
$$B_1 = \begin{bmatrix} 0.8 & 2.2 \\ 0.4 & 2.6 \end{bmatrix},\ B_2 = \begin{bmatrix} 0.1 & -0.9 \\ 0.7 & 0.3 \end{bmatrix} \tag{9}$$

We apply the following sinusoidal and rectangle waveform signals to the source signals; i.e., $\boldsymbol{s}(t) = [\sin(2\pi 90t),\ \mathrm{sgn}(\sin(2\pi 155t))]^T$. By using such nonlinear mixture system, the mixture signals $\boldsymbol{x}(t)$ are generated. Figure 5(a) illustrates the source signals $\boldsymbol{s}(t)$, and Fig. 5(b) illustrates the mixture signals $\boldsymbol{x}(t)$.

In order to compare the performance, we carry out the numerical simulations by using the standard PSO method, the Ring-PSO method and the steepest descent method. These methods use the same initial value. The simulation conditions of the PSO method are shown in Table 1. Moreover, the number of particles is changed from 10 to 80 by 10 units. Figure 4 shows the relationship between the evaluation function value and the number of particles. The simulation results indicate that the search performance is improved when the number of particles is increased. In addition, Figure 4 indicates that the Ring-PSO exhibits small evaluation value. Figures 5(c), 5(d) and 5(e) show the separate signals by the standard PSO, the Ring-PSO and the steepest descent method. These are the result of the case of 40 particles. The evaluation value and its mean squared error of the signals $\boldsymbol{s}(t)$ and $\boldsymbol{x}(t)$ are shown in Table 4. The evaluation value and its mean squared error of each simulation results are shown in Table 5. These results indicate that the evaluation value of the Ring-PSO method is small comparing with the steepest descent method. Therefore, the Ring-PSO method exhibits high performance.

## 7. Conclusions

In this paper, we proposed the nonlinear BSS algorithm which consists with the RBF network. The parameters of the RBF network are adjusted by the PSO algorithm. Comparing with the conventional gradient method, we con-




(a) Original signals $\boldsymbol{s}(t)$

(b) Mixed signals $\boldsymbol{x}(t)$

(c) Separated signals $\boldsymbol{y}(t)$ by the standard PSO

(d) Separated signals $\boldsymbol{y}(t)$ by the Ring-PSO

(e) Separated signals $\boldsymbol{y}(t)$ by the steepest descent method

Figure 5: The signals in nonlinear mixture case: The number of particles was 40.

Table 4: Evaluation value and its mean square error in nonlinear mixture case

| | fitness | *mse* |
|---|---|---|
| The original signals $\boldsymbol{s}(t)$ | 0.000 | 0 |
| The mixed signals $\boldsymbol{x}(t)$ | 1.542 | 0.250 |

Table 5: Simulation results in nonlinear mixture case: The number of particles was 40.

| | fitness | | *mse* | |
|---|---|---|---|---|
| | average | min | average | min |
| standard PSO | 0.096 | 0.000 | 0.172 | 0.021 |
| Ring-PSO | 0.050 | 0.000 | 0.201 | 0.011 |
| gradient method | 18.551 | 0.000 | 0.250 | 0.034 |

firmed the proposed algorithm exhibits higher performance when the identical initial values are used. However, the parameters used in this experiment are not optimal. Thus, analysis of the parameters is required.

In addition, the system falls into local solutions frequently in the numerical simulations. One of the reasons why the system is caught in such local solution is the characteristics of the evaluation function. Therefore, to improve the evaluation function is one of our future problems.

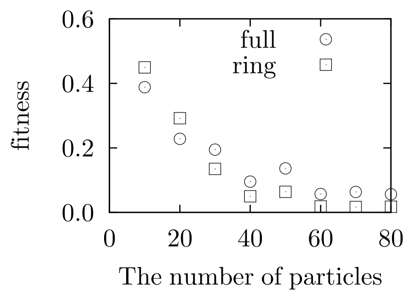



Figure 4: Search performance and the number of particles

## References

[1] P. Comon, IEEE Signal Processing, vol. 36, pp.287-314, 1994.

[2] R. Vigário, J. Särekä, V. Jousmäki, M. Hämäläinen, and E. Oja, IEEE Trans. Biomed. Eng. , vol. 47, pp. 589-593, 2000.

[3] P. Pajunen, A. Hyvärinen, and J. Karhunen, Proc. Int. Conf. on Neural Information Processing, pp. 1207-1210, Hong Kong, 1996.

[4] H. Lappalainen, and X. Giannakopoulos, Proc. Int. Conf. Artificial Neural Networks, pp. 19-24, 1999

[5] Y. Tan, J. Wang, and J. M. Zurada, IEEE Trans. Neural Networks, vol. 12, pp. 124-134, 2001.

[6] A. Hyvarinen and P. Pajunen, Neural Networks, vol. 12, pp. 429-439, 1999.

[7] S. Sirayanone and R. L. Hardy, Journal of Applied Sciences and Computations Vol. 1, pp. 437-475, 1995.

[8] J. Kennedy, Proc. of IEEE International Conference on Evolutionary Computation, pp. 303-308, 1997.

[9] T. Tsujimoto, T. Shindo, T. Kimura and K. Jin'no, Proc. 2012 IEEE World Congress on Evolutionary Computation, pp. 1526-1531, 2012.