

## Research Article

# Sampling Adaptive Learning Algorithm for Mobile Blind Source Separation

Jingwen Huang<sup>1</sup> and Jianshan Sun <sup>2</sup>

<sup>1</sup>Beijing University of Chemical Technology, Beijing 100029, China

<sup>2</sup>School of Management, Hefei University of Technology, Hefei 230009, China

Correspondence should be addressed to Jianshan Sun; sunjs9413@hfut.edu.cn

Received 23 November 2017; Accepted 26 December 2017; Published 18 March 2018

Academic Editor: Yin Zhang

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

Learning rate plays an important role in separating a set of mixed signals through the training of an unmixing matrix, to recover an approximation of the source signals in blind source separation (BSS). To improve the algorithm in speed and exactness, a sampling adaptive learning algorithm is proposed to calculate the adaptive learning rate in a sampling way. The connection for the sampled optimal points is described through a smoothing equation. The simulation result shows that the performance of the proposed algorithm has similar Mean Square Error (MSE) to that of adaptive learning algorithm but is less time consuming.

## 1. Introduction

With the fast development of the information and computation technologies, the big data analysis and cognitive computing have been widely used in many research areas such as medical treatment [1], transportation [2], and wireless communication [3, 4]. Blind Source Separation (BSS) is a popular research topic in the area of wireless communication. With the fast development of mobile computing, BSS has been widely used in the mobile signal analysis. BSS is an integration of artificial neural network, statistical signal processing, and information theory. The core of BSS is its ability to extract independent components from an observed mixture signal, without requiring a prior knowledge. Such flexibility has made BSS popular in many applications [5–7] especially in mobile intelligence [8, 9].

Artificial neural network based Independent Component Analysis (ICA) is the widely used method in BSS, because it provides powerful tools to capture the structure in data by learning. Based on this theory, Natural Gradient Algorithm (NGA) is employed to find the appropriate coefficient vector of artificial neural network [10]. Nonholonomic Natural Gradient Algorithm (NNGA) is addressed and applied in the BSS [11, 12]. In its application, learning rate for training coefficient vector plays an important role on the performance of the

algorithm, which has relationship with not only the update times but also the speed of convergence. This attracts many researchers' attention to the learning algorithms [13, 14].

Most well-known traditional learning algorithms assume that the learning rate is a small positive constant. Inappropriate constant will lead to relative slow convergence speed or big steady state error. There are lots of studies on the learning rate which aim at the better performance and higher convergence speed. von Hoff and Lindgren [15] developed adaptive step size control algorithm for gradient-based BSS. They used the coefficients of the estimating function to provide an appropriate "measure of error" and serve as the basis for a self-adjusting time-varying step-size. Hai proposed a conjugate gradient procedure for second-order gradient-based BSS [16]. The second-order gradient-based BSS was formulated as a constrained optimization problem. A conjugate gradient procedure with the step size derived optimally at each iteration was proposed to solve the optimization problem. In these algorithms, the step size is updated in iteration, whose value is adjusted according to the time-varying dynamics of the signals. These approaches lead to better performance. The real time search for the step size, however, requires more online calculations as well as recursion, resulting in an increase in computational complexity. On the other hand, in the recursion for optimal step size, there is still

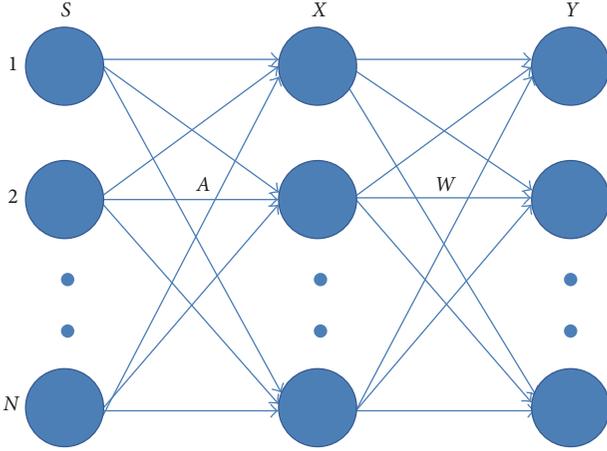


FIGURE 1: The system model.

a constant left to be estimated, which leads to an endless loop.

The objective of this paper is to find appropriate learning algorithm, to provide better performance as well as less computation time. The proposed sampling learning rate is based on adaptive learning algorithm, which only calculates and samples a few appropriate points. These selected points are connected by the proposed normalized smooth equation.

In the following, we first review the principle of blind signal separation. Then, we discuss the adaptive learning algorithm and propose the sampling adaptive learning algorithm. Finally, we present two typical examples in mobile voice signal. Different constant learning rates are compared firstly to analyze the relationship between the convergence speed and steady state error. Then the comparison between the adaptive learning algorithm and the sampling adaptive learning algorithm is made, illustrating that the proposed algorithm has similar Mean Square Error (MSE) to that of the adaptive learning algorithm but consumes less computational time.

## 2. Blind Signal Separation

The model considered in this paper is described by Figure 1. A set of individual source signals  $\mathbf{s}(k)$  is mixed with  $A$  matrix to produce a set of mixed signals  $\mathbf{x}(k)$ .

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

where  $A = [a_{ij}] \in \mathbb{R}$  is the an unknown mixing matrix independent of time,  $\mathbf{s}(k) = [s_1(k), s_2(k), \dots, s_N(k)]^T$  is the vector of source signals, and  $\mathbf{x}(k) = [x_1(k), x_2(k), \dots, x_N(k)]^T$  is the vector containing the observed signals.

BSS separates the mixed signals, through the determination of an unmixing matrix  $W(k)$  to recover the approximation of the original signals. The recovered output signal is described by

$$\mathbf{y}(k) = W(k)\mathbf{x}(k), \quad (2)$$

where  $W$  is the  $N \times N$  matrix to be adjusted such that  $WA = I$ . With the NGA, the matrix  $W$  is updated from  $W_k$  at time  $k$  to  $W_{k+1}$  by using the following adaptation rule.

$$W(k+1) = W(k) + \partial(k) [I - f(y(k))y^T(k)]W(k), \quad (3)$$

where  $\partial(k)$  is the learning rate.  $f(y)$  is the score function defined by

$$f_i(y_i) = -\frac{p_i(y_i)'}{p_i(y_i)}, \quad (4)$$

where  $p_i(y_i)$  is the probability density function of the  $i$ th source signal. It is assumed that the source signals are zero mean. Hence, we have

$$E(s_i(t)) = 0, \quad (5)$$

where  $E$  denotes the expectation.

According to the reference [17],  $f(y)$  can be set to be  $f(y) = y^3$  when  $\mathbf{s}(k)$  is the sub-Gaussian signal;  $f(y) = \tan(y)$  when  $\mathbf{s}(k)$  is the super-Gaussian signal. Most mobile voice signals are super-Gaussian signals and most mobile image signals are sub-Gaussian signals. The function is accordingly selected based on the mobile signal type.

The learning rate is a very important factor for the performance of BSS in controlling the magnitudes of the updates of the estimated parameters. It can be constant or variable. The constant learning rate means that the adaptation in (3) is based on fixed step-size parameter. Because the step size is proportional to convergence speed and the steady state error, the big step size leads to high convergence speed and big steady state error. However, what we expect is the high convergence speed and small steady state error. In conclusion, the main problem of the constant learning rate focuses on the incompatibility between the convergence speed and the steady state error [18, 19]. Therefore, many researchers were aiming at adaptiveness by using the variable step-size approaches. We will discuss the adaptive learning algorithm and propose a sampling adaptive learning algorithm in the following section.

## 3. Sampling Adaptive Learning Algorithm

**3.1. Adaptive Learning Algorithm for BSS.** The idea of adaptively changing the step size of the learning rate is called adaptive learning algorithm. The adaptive learning algorithm can balance the convergence speed and the steady error performance. In this case, we discuss the adaptive learning algorithm which updates the learning rate step size through the estimate function. Although the distance between the estimated parameter and its optimal value is not directly available to control the step size, evaluation function can be developed to estimate the distance so that the step size can be determined by the following recursion:

$$\partial(k+1) = \partial(k) + \beta \bar{\varphi}(H_k), \quad (6)$$

TABLE I: Performance of CLA and ALA.

	Convergence time (iterative time)	Mean steady state error	Computation time
CLA	681	2.1	55.6 s
ALA	482	1.9	96.6 s

where  $\beta$  is a constant and  $H_k$  is an estimate function.  $\varphi$  is the evaluation function. In this adaptive learning algorithm, the estimate function is set to be

$$H_k = I - f_k(y) y_k^T. \quad (7)$$

A smoothed version of the evaluation function, denoted by  $\widehat{H}_{ijk}$ , can be set to be

$$\begin{aligned} \widehat{H}_{ijk} &= E(H_{ijk}) \\ &= -E(f'(s_i))E(s_j^2)W_{ijk}A - E(f(s_i)s_i)W_{jik}A \end{aligned} \quad (8)$$

and the evaluation function  $\overline{\varphi}(H_k)$  is set to be

$$\overline{\varphi}(H_k) = \max(|\widehat{H}_{ijk}|). \quad (9)$$

This idea is similar to Least Mean Square Adaptive Filter (LMS) and Reinforcement Learning (RL). It can be concluded from (6)–(9) that the step size depends on the evaluation of the estimate function. The principle of this adaptive algorithm implies that the step size is small when the errors are small and the step size is large when the errors are large. For time-invariant systems the step size systematically decreases when the learning rate is close to their optimum.

However, we must notice that, in the process to obtain a better learning rate, the adaptive step-size control algorithm actually introduces another recursion. In this recursion shown in (6), there is still an unknown constant  $\beta$  to be determined. If we want to find the optimal  $\beta$ , we also need to find another recursion like (6), which leads to endless iteration.

On the other side, the recursion as described in (6) introduces another calculation recycle which adds additional computation. In some case, calculation with adaptive learning rate even consumes more time than the calculation with constant learning rate. Figure 2 is the ideal adaptive step size curve for noise-free signal. Every point on the curve is calculated with (6)–(9). We did the simulation for BSS based on the adaptive step-size and compared it with the selected ideal constant step size. Table 1 shows the performance index resulting from the application of the two methods and illustrates that the convergence speed of ALA is considerably higher when the adaptive step-size algorithm is employed. The steady state error of these two algorithms is similar when the random possibility is considered. As for the computation time, it is obvious that ALA consumes more. We can conclude from the analysis that although the adaptive algorithm balances the convergence speed and the steady state error, it consumes more computation time as shown in Table 1.

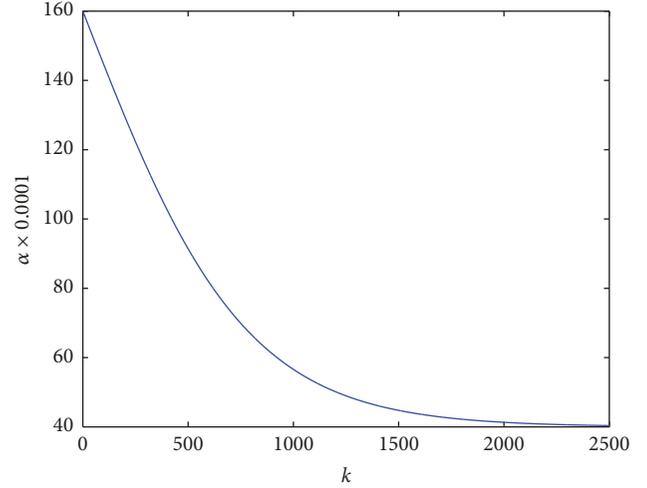
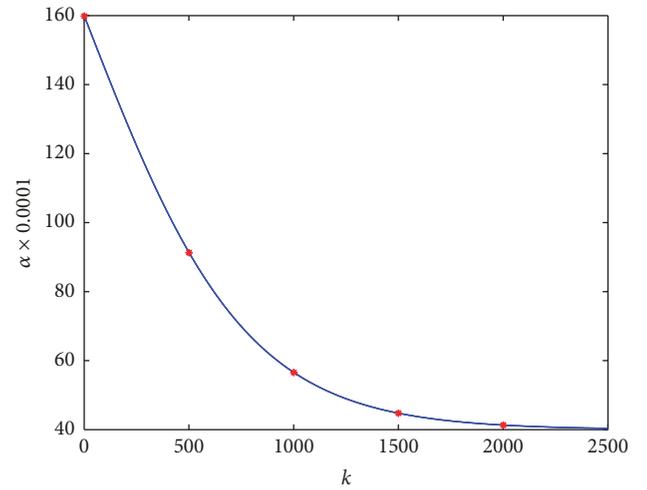


FIGURE 2: The ideal adaptive step size curve for noise-free signal.


 FIGURE 3: Effect of sample learning rate on updating the matrix  $W$ .

**3.2. Adaptive Learning Algorithm Based on Sampling.** An alternative way to use the adaptive strategy is to implement the adaptive step size on sampling. In this method, only several points need the adaptive calculation. For training the  $W$  matrix, it is not necessary to have the variable step size all the time. As shown in Figure 3, several sampling points are enough for keeping  $W$  updated. Let  $\Delta K$  be the sampling interval; the time variable  $k'$  becomes

$$k' = \left\lceil \frac{k}{\Delta K} \right\rceil, \quad (10)$$

where  $\lceil \cdot \rceil$  means rounding.

Then the learning rate can be represented as

$$\partial(k' + 1) = \partial(k') + \beta \overline{\varphi}(H_{k'}). \quad (11)$$

Equations ((7)–(9)) therefore become

$$\begin{aligned}
 H_{k'} &= I - f_{k'}(y) y_{k'}^T \\
 \widehat{H}_{ijk'} &= E(H_{ijk'}) \\
 &= -E(f'(s_i)) E(s_j^2) W_{ijk'} A \\
 &\quad - E(f(s_i) s_i) W_{jik'} A \\
 \bar{\varphi}(H_{k'}) &= \max(|\widehat{H}_{ijk'}|).
 \end{aligned} \tag{12}$$

The new equations reduce the times of iteration by dividing the sampling interval  $\Delta K$ . We could choose the  $\Delta K$  according to the required accuracy and speed. We also provide the connection for the sampled optimal points, which smooth the curve between two optimal points. Based on this analysis, the learning rate between two optimal points can be expressed as

$$\partial(k) = (1 - \tanh(\zeta k)) \partial_{1s} + \tanh(\zeta k) \partial_{2s}, \tag{13}$$

where  $\partial_{1s}$  and  $\partial_{2s}$  are two sampling points which can be obtained at times  $k'$  and  $k' + 1$  with sampling interval. When  $\zeta k = 5.3$ ,  $\tanh(\zeta k) \approx 1$  can be added as the condition used to determine where to switch the step size. In (13),  $k$  is the only variable that keeps changing. This algorithm will not bring another recursion for the system but will still have the optimal value choice.

For the convenience of application, the normalized form for  $\partial(k)$  is given by

$$\begin{aligned}
 \partial(k) &= \left(1 - \tanh\left(100\zeta_N \frac{k - k_{s1}}{k_{s1} - k_{s2}}\right)\right) \partial_{1s} \\
 &\quad + \tanh\left(100\zeta_N \frac{k - k_{s1}}{k_{s1} - k_{s2}}\right) \partial_{2s},
 \end{aligned} \tag{14}$$

where  $\zeta_N$  is set to be 0.053.  $k_{s1}$  and  $k_{s2}$  are the first and second sampling points, respectively.

## 4. Experiment and Result

*4.1. Case 1.* To test the algorithm, five sub-Gaussian source signals commonly studied in the mobile system are employed. The source signals are

$$\begin{aligned}
 s(1) &= \sin(2\pi 800t) \\
 s(2) &= \cos(2\pi 500t) \\
 s(3) &= \sin(2\pi 100t + 6 \cos(2\pi 40t)) \\
 s(4) &= \sin(2\pi 50t) \\
 s(5) &= \text{rand}(1, \text{length}(t)).
 \end{aligned} \tag{15}$$

These source signals are shown in Figure 5. A mixing coefficient was assigned to be independent white Gaussian noise, which is with zero mean. These sources were mixed producing mixtures as shown in Figure 6. The mixtures were

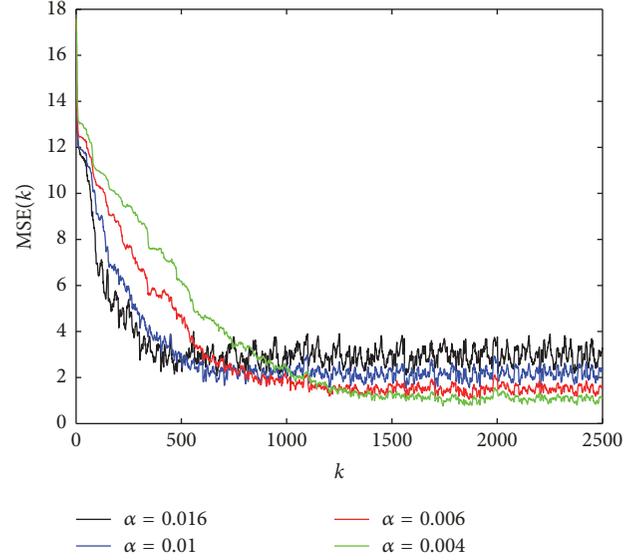


FIGURE 4: Mean Square Error (MSE) for different fixed learning rate.

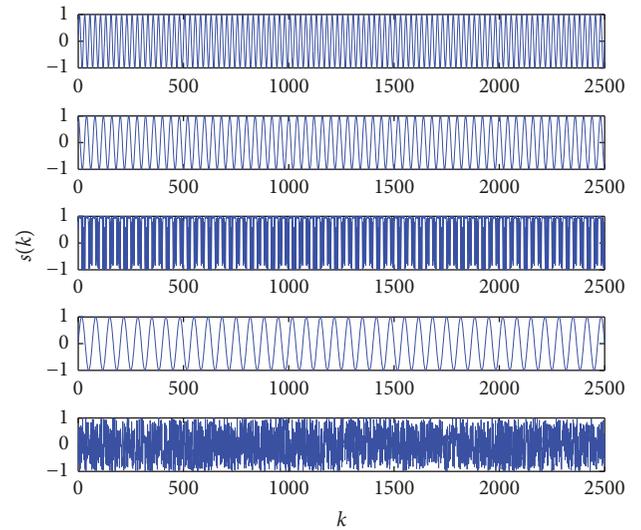


FIGURE 5: The raw source signals.

separated using fixed learning rate, variable learning rate, and sampled learning rate which are named Constant Learning Algorithm (CLA), Adaptive Learning Algorithm (ALA), and sampling adaptive learning algorithm (SALA).

Firstly, we employ four fixed learning rate  $\partial = [0.016, 0.010, 0.006, 0.004]$  to check the performance. Figure 4 depicts the Mean Square Error (MSE) for these four fixed learning rates. In this case the MSE [20] is defined as

$$\begin{aligned}
 \text{MSE}(k) &= \sum_{i=1}^n \left\{ \left( \frac{\sum_{k=1}^n |W_{ik} A|}{\max_j |W_{ij} A|} - 1 \right) \right. \\
 &\quad \left. + \left( \frac{\sum_{k=1}^n |W_{ki} A|}{\max_j |W_{ji} A|} - 1 \right) \right\}.
 \end{aligned} \tag{16}$$

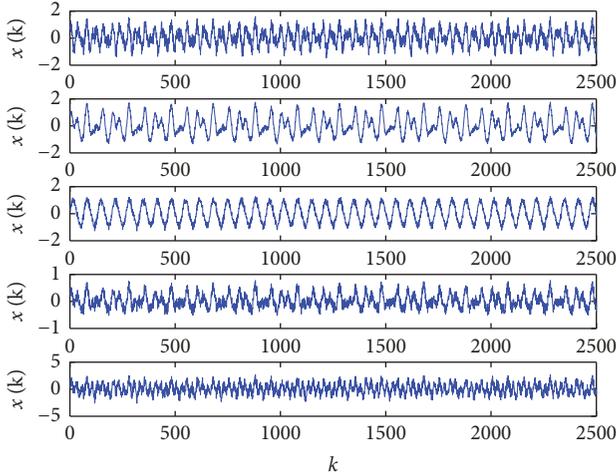


FIGURE 6: The mixed signals.

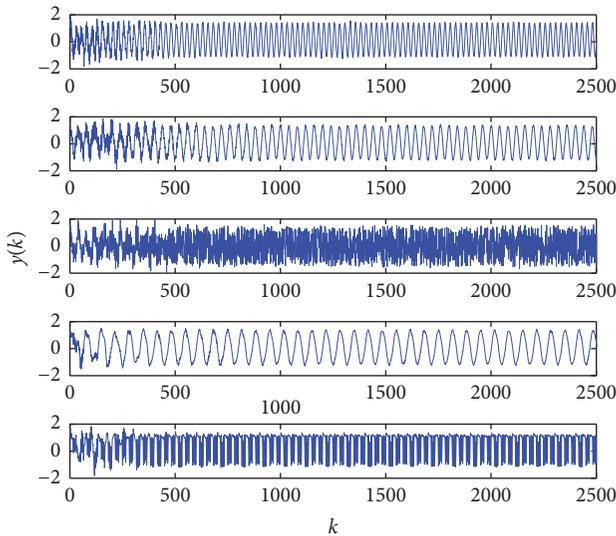


FIGURE 7: The separated signals.

As can be seen from Figure 4, MSE varied according to different fixed learning rate. The bigger learning rate has fast convergence speed at the beginning while the smaller learning rate leads to better stability at the end.

Adaptive Learning Algorithm (ALA) can balance the requirement of convergence speed and steady state error. However, ALA requires more computation time in iteration. To solve this problem, we use the sampling adaptive learning algorithm (SALA), by using the sampling interval in the adaptive learning algorithm. Figure 7 is the output signal with SALA. To compare SALA with ALA, (6) and (11) are adopted, respectively, but all with the initial learning rate  $\partial = 0.016$  and negative constant  $\beta$ . Figure 8 depicts the MSE of ALA and SALA.

Figures 7 and 8 show that the MSE of ALA and SALA are similar. The high MSE in Figure 8 before the 500th iterative time leads to inaccurate estimated curve in Figure 7

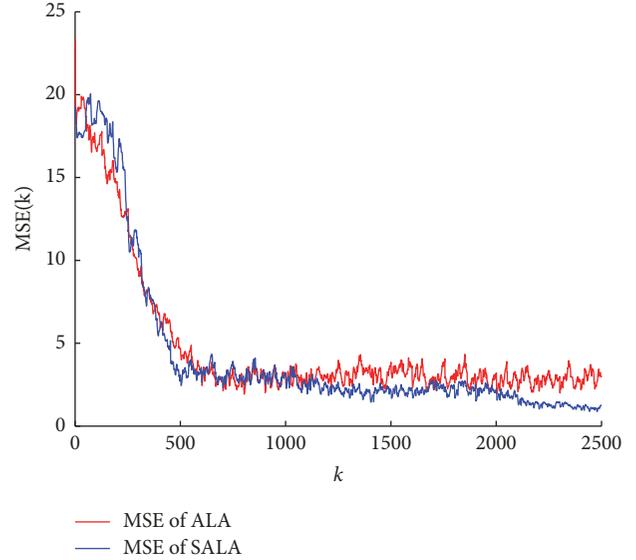


FIGURE 8: The MSEs of different algorithms.

TABLE 2: Performance of ALA and SALA for mobile signals.

	Convergence time (samples)	Mean steady state error	Computation time
ALA	582	2.3	96.6 s
SALA	586	2.0	64.2 s

correspondingly while the output estimated curve in Figure 7 becomes accurate when the steady state error in Figure 7 gets small. The results were evaluated using the performance index which is listed in Table 2. The convergence time and the mean steady state error of ALA and SALA are on the same level considering the random factor, but the computation time of SALA is obviously less than ALA. We can conclude from the analysis that the proposed SALA has advantage over ALA in the computation time.

**4.2. Case 2.** In order to verify the effectiveness of the proposed SALA, two music sources from the real environment are tested through the simulation. These music sources were mixed by random Gaussian noise matrix  $A$  of full rank  $m = 2$ . The mixtures were separated using SALA with the sampled variable step size. Figure 9 confirms the accurate and fast estimate as observed in Case 1.

## 5. Conclusions

Based on the discussion of fixed step-size algorithm and adaptive step-size algorithm for the blind separation of sources, a sampling adaptive step-size algorithm has been proposed. The algorithm has similar MSEs with adaptive step-size algorithm, but less computational time. By a smooth connection between two optimal points, the sampling method also has smooth curve and does not bring more recursion.

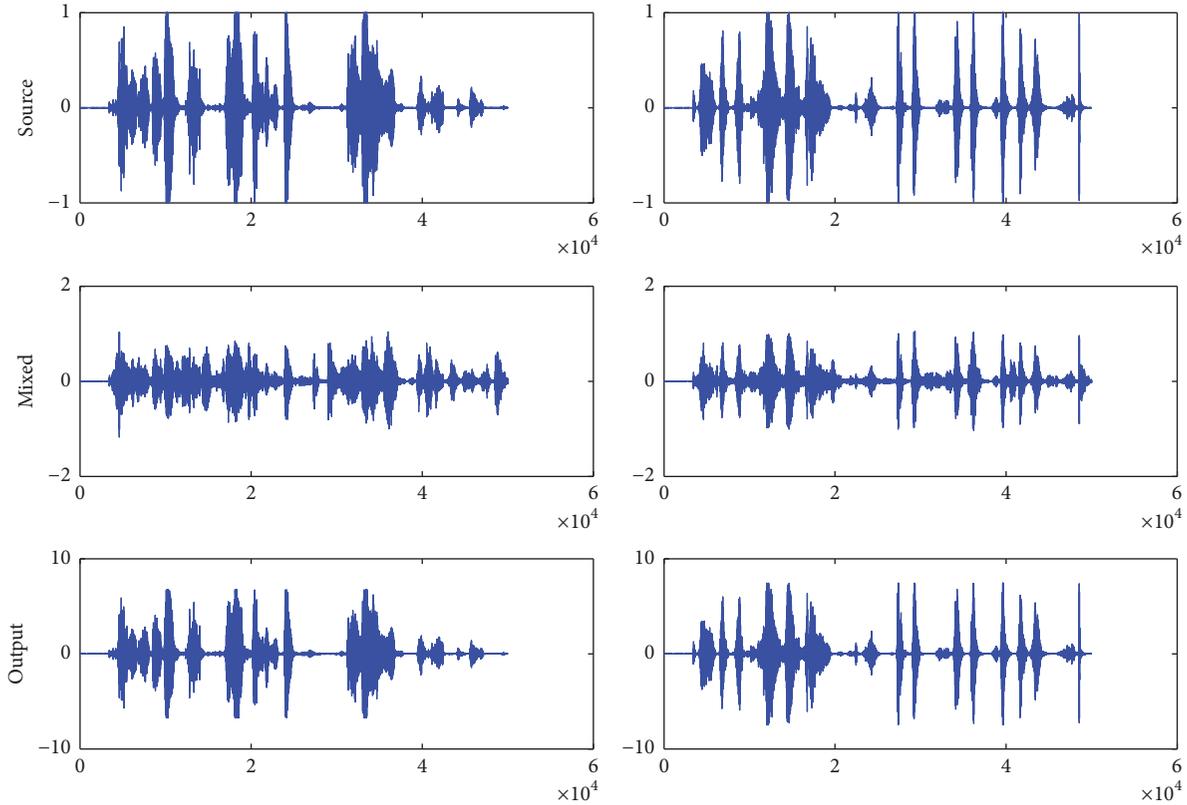


FIGURE 9: SALA for the real signal.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work is supported by the Natural Science Foundation of China through the Grant 11702016.

## References

- [1] L. Ogiela, R. Tadeusiewicz, and M. R. Ogiela, "Cognitive computing in analysis of 2D/3D medical images," in *Proceedings of the 2007 International Conference on Intelligent Pervasive Computing, IPC 2007*, pp. 15–18, October 2007.
- [2] Y. Zhang, M. Chen, N. Guizani, D. Wu, and V. C. Leung, "SOVCAN: safety-oriented vehicular controller area network," *IEEE Communications Magazine*, vol. 55, no. 8, pp. 94–99, 2017.
- [3] M. Chen, Y. Zhang, L. Hu, T. Taleb, and Z. Sheng, "Cloud-based wireless network: virtualized, reconfigurable, smart wireless network to enable 5G technologies," *Mobile Networks and Applications*, vol. 20, no. 6, pp. 704–712, 2015.
- [4] Y. Zhang, "GroRec: a group-centric intelligent recommender system integrating social, mobile and big data technologies," *IEEE Transactions on Services Computing*, vol. 9, no. 5, pp. 786–795, 2016.
- [5] D. N. Levin, "Model-independent method of nonlinear blind source separation," in *Latent Variable Analysis and Signal Separation. LVA/ICA, 2017., Lecture Notes in Computer Science*, P. Tichavský, M. Babaie-Zadeh, O. Michel, N. Thirion-Moreau, and P. Tichavský, Eds., vol. 10169, pp. 310–319, Springer, 2017.
- [6] C. Hu, Q. Yang, M. Huang, and W. Yan, "Sparse component analysis-based under-determined blind source separation for bearing fault feature extraction in wind turbine gearbox," *IET Renewable Power Generation*, vol. 11, no. 3, pp. 330–337, 2017.
- [7] N. D. Stein, "Nonnegative tensor factorization for directional blind audio source separation, Computer Science".
- [8] E. Visser and L. Tewon, "Blind source separation in mobile environment using a priori knowledge," in *Proceedings of IEEE International Conference on Acoustics*, vol. 3, pp. 893–896, Quebec, Canada, 2004.
- [9] S. M. Naqvi, Y. Zhang, and J. A. Chambers, "Multimodal blind source separation for moving sources," in *Proceedings of the ICASSP 2009 - 2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 125–128, Taipei, Taiwan, April 2009.
- [10] C. Ji, B. Tang, and J. Wang, "Blind source separation based on variable step length natural gradient algorithm," in *Proceedings of the 2011 3rd International Conference on Awareness Science and Technology, iCAST 2011*, pp. 261–264, September 2011.
- [11] Y. L. Niu, J. C. Ma, and Y. Wang, "Blind separation for blurred images based on the adaptive nonholonomic natural gradient algorithm," in *Proceedings of the 2008 2nd International Symposium on Systems and Control in Aerospace and Astronautics, ISSCAA 2008*, December 2008.
- [12] C. Ji, K. Yang, Y.-R. Wang, and M.-D. Liu, "Variable step-size nonholonomic natural gradient algorithm based on sign operator," *Moshi Shibie yu Rengong Zhineng/Pattern Recognition and Artificial Intelligence*, vol. 27, no. 11, pp. 1026–1031, 2014.

- [13] M. R. Bastian, J. H. Gunther, and T. K. Moon, "A simplified natural gradient learning algorithm," *Advances in Artificial Neural Systems*, vol. 2011, pp. 3–13, 2011.
- [14] J. Nan, "Natural gradient reinforcement learning algorithm with TD ( $\lambda$ )," *Computer Science*, vol. 37, no. 12, pp. 186–189, 2010.
- [15] T. P. von Hoff and A. G. Lindgren, "Adaptive step-size control in blind source separation," *Neurocomputing*, vol. 49, pp. 119–138, 2002.
- [16] H. H. Dam, D. Rimantho, and S. Nordholm, "Second-order blind signal separation with optimal step size," *Speech Communication*, vol. 55, no. 4, pp. 535–543, 2013.
- [17] S. L. Gay and J. Benesty, *Acoustic Signal Processing for Telecommunication*, Springer, Boston, MA, USA, 2000.
- [18] M. G. Jafari, J. A. Chambers, and D. P. Mandic, "A novel adaptive learning rate sequential blind source separation algorithm," *Signal Processing*, vol. 84, no. 4, pp. 801–804, 2004.
- [19] J. Ce, Y. Peng, and Y. Yang, "Blind source separation based on improved natural gradient algorithm," in *Proceedings of the 2010 8th World Congress on Intelligent Control and Automation, WCICA 2010*, pp. 6804–6807, Jinan, China, July 2010.
- [20] Y. Zhang, S. Lou, W. Zhang, and H. Chang, "Blind source separation algorithm of natural gradient based on estimation of score function," *Shuju Caiji Yu Chuli/Journal of Data Acquisition and Processing*, vol. 26, no. 2, pp. 167–171, 2011.

