

## Research Article

# Variable Step-Size Method Based on a Reference Separation System for Source Separation

Pengcheng Xu, Zhigang Yuan, Wei Jian, and Wei Zhao

*College of Communications, PLA University of Science and Technology, Nanjing 210007, China*

Correspondence should be addressed to Pengcheng Xu; [pchxu2005@gmail.com](mailto:pchxu2005@gmail.com)

Received 9 December 2014; Revised 7 April 2015; Accepted 9 April 2015

Academic Editor: Jian-Nong Cao

Copyright © 2015 Pengcheng Xu et al. 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.

Traditional variable step-size methods are effective to solve the problem of choosing step-size in adaptive blind source separation process. But the initial setting of learning rate is vital, and the convergence speed is still low. This paper proposes a novel variable step-size method based on reference separation system for online blind source separation. The correlation between the estimated source signals and original source signals increases along with iteration. Therefore, we introduce a reference separation system to approximately estimate the correlation in terms of mean square error (MSE), which is utilized to update the step-size. The use of “minibatches” for the computation of MSE can reduce the complexity of the algorithm to some extent. Moreover, simulations demonstrate that the proposed method exhibits superior convergence and better steady-state performance over the fixed step-size method in the noise-free case, while converging faster than classical variable step-size methods in both stationary and nonstationary environments.

## 1. Introduction

Blind source separation (BSS) aims at extracting the latent unknown source signals from their observed mixtures by an array of sensors without a priori knowledge of the original source signals and the mixing coefficients. In the separating process, nothing can be used except for the observation sequences and the statistical characteristic assumptions of the sources. This makes BSS become a versatile tool used in many multisensor systems such as antenna arrays in acoustics or electromagnetism, chemical sensor arrays, and electrode arrays in electroencephalography [1].

Several optimization algorithms have been proposed for BSS [2] and can be generally categorized into batch-based algorithms and adaptive (sequential) algorithms. Batch-based algorithms are block-wise and will not work until a block of data samples is received, such as the fast fixed-point algorithm [3]. In this paper, we consider the latter, which have particular practical advantages due to their computational simplicity and latent ability in tracing a nonstationary environment [4].

However, the traditional adaptive BSS algorithms such as equivariant adaptive separation via independence algorithm

(EASI) [5] and natural gradient algorithm (NGA) [6] usually assume that the step-size is a small positive constant, leading to an inevitable conflict between the learning rate and stability performance, that is, slow convergence speed or large steady state error. A simple way to solve the conflict is reducing the learning rate as the iteration goes on [7, 8], but it brings about another new problem: if the learning rate decreases to be too small before source components are extracted, the separation system will fail to separate sources properly. To improve the learning rate and stability performance, variable step-size algorithms have been proposed. The variable step-size algorithms can exploit the online measurements of the state of the separation system from the outputs and the parameter updates. In [4, 9, 10], variable step-size algorithms have been derived according to the gradient of different contrasts, that is, NGA, EASI, and S-NGA algorithms. Zhang et al. put forward a grading learning algorithm based on the measurements of correlation of the separating signals, whose learning rate is updated by the state of separating [11]. Hsieh et al. proposed an effective learning rate adjustment method based on an improved particle swarm optimizer [12]. But the separating performance of these variable step-size algorithms is usually sensitive to the initial parameter settings.

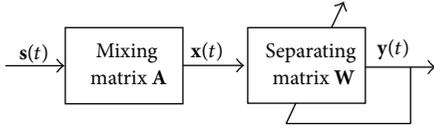


FIGURE 1: Adaptive BSS model.

As a result, the convergence is still slow and improper initial value of learning rate results in large steady state error or even divergence. Ou et al. proposed a variable step-size algorithm based on an auxiliary separation system [13]. The step-size is updated by estimating a pseudo-performance index in the light of the index descending in an exponential form. Compared to classical variable step-size methods, the separation performance of Ou's method is less sensitive to the initial settings.

In order to improve the initial convergence and stability performance, we consider using a reference separation system based on MSE of the instantaneous outputs to update the step-size. This technique is shown to improve the convergence speed and the steady-state performance. Moreover, the use of "minibatches" can reduce the whole computational load of the algorithm. The remainder of this paper is organized as follows. In Section 2, the principle of adaptive source separation methods is briefly summarized. Our algorithm is proposed in Section 3. Numerical stimulation results and discussion are provided in Section 4. At the end of the paper, a concise conclusion is given. What is more, this paper can be regarded as an important complement for Ou's method in [13].

## 2. Adaptive Algorithms for BSS

In the noise-free instantaneous case, we assume that  $n$  unknown statistically independent zero mean source signals, with at most one having a Gaussian distribution, contained within  $\mathbf{s}(t) = [s_1(t), \dots, s_n(t)]^T$  pass through an unknown mixing system  $\mathbf{A} \in \mathfrak{R}^{m \times n}$  ( $m \geq n$ ); therefore  $m$  mixed signals  $\mathbf{x}(t) = [x_1(t), \dots, x_m(t)]^T$  can be modeled as

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

where  $t$  is the time index and  $\mathbf{T}$  is the vector transpose operator. To simplify the problem, we further assume that the number of sources matches the number of mixtures, that is,  $m = n$ , an exactly determined problem.

The blind separation problem is then to recover original source signals  $\mathbf{s}(t)$  from observations  $\mathbf{x}(t)$ , which is equivalent to estimate an separating matrix  $\mathbf{W} \in \mathfrak{R}^{n \times n}$  that performs the inverse operation of the mixing process, as subsequently used in separation model. Figure 1 shows a block diagram of adaptive BSS model. Then the output signal vector is obtained:

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

where  $\mathbf{y}(t) = [y_1(t), \dots, y_n(t)]^T$  is an estimate of  $\mathbf{s}(t)$  to within the well-known permutation and scaling ambiguities.

Based on classical contrast such as mutual information contrast, maximum likelihood contrast, and informax principle, many adaptive algorithms have been proposed to estimate  $\mathbf{s}(t)$ . Amari proved the NGA algorithm is the fastest least-mean-square (LMS) type BSS algorithm [6]. The natural gradient BSS algorithm based on the mutual information contrast, maximum likelihood contrast, and informax principle have the same form:

$$\mathbf{W}(t+1) = \mathbf{W}(t) + \mu \{ \mathbf{I} - \phi[\mathbf{y}(t)]\mathbf{y}^T(t) \} \mathbf{W}(t), \quad (3)$$

where  $\mathbf{I}$  is identity matrix, and  $\phi(\cdot) = [\phi_1(\cdot), \dots, \phi_n(\cdot)]^T$ ,  $\phi_i(\cdot)$ ,  $i = 1, 2, \dots, n$  are increasing odd functions, usually called activation functions.

Based on the fact that the separating matrix  $\mathbf{W}$  can be factorized into the product of an orthogonal matrix and the prewhitening matrix, via combining LMS-type updating formulas of these two matrixes above, using some reasonable approximation, the EASI algorithm is derived [5]:

$$\begin{aligned} \mathbf{W}(t+1) = \mathbf{W}(t) + \mu \{ \mathbf{I} - \mathbf{y}(t)\mathbf{y}^T(t) - \phi[\mathbf{y}(t)]\mathbf{y}^T(t) \\ + \mathbf{y}(t)\phi[\mathbf{y}^T(t)] \} \mathbf{W}(t). \end{aligned} \quad (4)$$

It has been shown that, as compared with using a fixed step-size, the algorithm with a variable step-size has an improved convergence rate. Yuan et al. derived a gradient variable step-size algorithm for the NGA algorithm [10], which adapts the step-size in the form of

$$\mu(t) = \mu(t-1) - \rho \nabla_{\mu} J(t) \Big|_{\mu=\mu(t-1)}, \quad (5)$$

where  $\rho$  is a small constant, and  $J(t)$  is an instantaneous estimate of the cost function from which the NGA algorithm is derived.

What should be noticed is that the activation functions  $\phi_i(\cdot)$  (i.e., the step-size update functions) can be identical, when sources are all sub-Gaussian or super-Gaussian signals. The distinct distributions of signals determine the different activation functions; that is, the separation of all sub-Gaussian sources usually utilizes the cubic function while the proper choice for super-Gaussian sources separation is hyperbolic tangent function.

## 3. The Proposed Algorithm

As is known to us, in the process of adaptive BSS the estimated signals  $\mathbf{y}(t)$  will approximate to source signals  $\mathbf{s}(t)$  as iteration goes on if the permutation and scaling ambiguities of the estimated signals  $\mathbf{y}(t)$  can be eliminated [14]. The correlation between the estimated signals  $y_i(t)$  and source signals  $s_j(t)$  can be evaluated by mean-square-error, which is defined as

$$\text{mse}_{ij} = \frac{1}{T} \sum_{t=1}^T [y_i(t) - s_j(t)]^2, \quad (6)$$

where  $T$  is the sample size, and  $y_i(t)$  as well as  $s_j(t)$  is normalized before the evaluation of  $\text{mse}_{ij}$ . When the separation

system is steady, the mean-square-error matrix MSE, whose  $(i, j)$  element is  $\text{mse}_{ij}$ , has one, and only one, zero entry in each row and column.

If we can calculate the matrix MSE at each update of the separating matrix  $\mathbf{W}$ , a rule for variable step-size algorithm is to adjust  $\mu$  adaptively in terms of MSE. However, since the source signals  $\mathbf{s}(t)$  are unknown, the matrix MSE at each update is not accessible in practice.

In this section, we propose to estimate MSE approximately by combining a reference separation system  $\mathbf{W}_a(t)$ , which follows the same optimization criteria and updating principle as  $\mathbf{W}(t)$  based on natural gradient algorithm (NGA), except for the initialization. Hence, we obtain that

$$\begin{aligned} \mathbf{W}_a(t+1) &= \mathbf{W}_a(t) \\ &+ \mu(t) \left\{ \mathbf{I} - \phi[\mathbf{y}_a(t)] \mathbf{y}_a^T(t) \right\} \mathbf{W}_a(t), \end{aligned} \quad (7)$$

which represents the reference signal. The correlation between  $\mathbf{y}(t)$  from the primary separation system and  $\mathbf{y}_a(t)$  from the reference system should increase as iteration goes on regardless of the ambiguities. Therefore, at every iterative, we replace mean-square-error in (6) by

$$\text{mse}_{ij} = \frac{1}{T} \sum_{t=1}^T \left[ \bar{y}_i(t) - \bar{y}_{aj}(t) \right]^2, \quad (9)$$

where

$$\begin{aligned} \bar{y}_i(t) &= \left| \frac{y_i(t)}{\text{norm}(y_i(t))} \right|, \\ \bar{y}_{aj}(t) &= \left| \frac{y_{aj}(t)}{\text{norm}(y_{aj}(t))} \right|, \end{aligned} \quad (10)$$

where  $\text{norm}(\cdot)$  denotes the root-mean-square value of the output vectors, and the operator  $|\cdot|$  takes the absolute value of the normalized vector. In this way, scaling ambiguity can be removed.

Online procedures use a given sample every time [6], whereas to appropriately evaluate mean-square-error one time requires some samples just as (9) indicates. Therefore, we consider updating the separating matrix once over a ‘‘minibatch,’’ that is, a small block of signal samples, while the observation window slides [15, 16]. Hence, the online updating equation of the separating matrix becomes

$$\begin{aligned} \mathbf{W}_a(k+1) &= \mathbf{W}_a(k) \\ &+ \mu(k) \left\{ \mathbf{I} - \frac{1}{T} \sum_{t=1}^T \phi[\mathbf{y}_a(t)] \mathbf{y}_a^T(t) \right\} \mathbf{W}_a(k), \end{aligned} \quad (11)$$

where  $k$  is the iteration number index (or the minibatch index), and the step-size parameter is updated by a nonlinear function in the form of

$$\mu(k) = \beta \left\{ 1 - \exp[-\alpha C(k)] \right\} \quad (12)$$

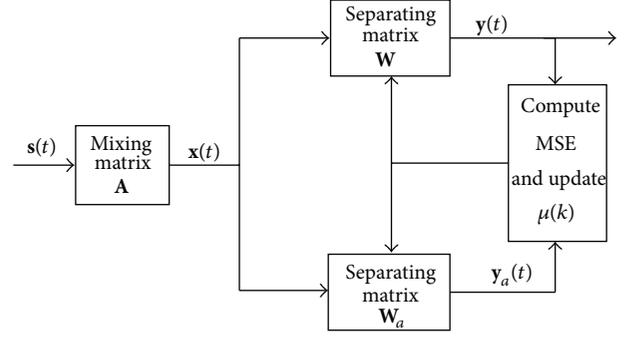


FIGURE 2: Scheme of the proposed algorithm.

which is a widely used rule in adaptive filtering algorithms [17]. The primary separation system follows the same updating rule above, where the parameters  $\alpha$  and  $\beta$  are two positive constants, which control the shape of the function curve and initial step-size, respectively. The effects of these two parameters on performance of the algorithm will be investigated in the next section. We define the correlation function as

$$C(k) = \sum_{l=k-D+1}^k \sum_{i=1}^n \frac{g^{k-l}}{n} \min_j (\text{mse}_{ij}(l)), \quad (13)$$

where  $\text{mse}_{ij}(l)$  denotes mean-square-error of  $l$ th minibatch,  $g$  represents a weighting factor  $0 < g < 1$ , and  $D$  is the number of weighting functions. Thus we introduce exponential weighting into past data, which is proper especially when the channel characteristics are time-variant [18, 19].

The separation procedure using ((9)–(13)) represents the proposed variable step-size algorithm. Figure 2 shows the scheme of our proposed algorithm.

Regarding the computation load, in the proposed algorithm  $2n^3 + 5n^2T + 3n^2 + (D^2 + D)/2 + 5$  products per iteration are required, but the total iteration number is  $N/T$ . Therefore,  $[2n^3/T + 5n^2 + 3n^2/T + (D^2 + D)/(2T) + 5/T]N$  products are calculated in a whole separation procedure, which has the same quantity level, that is,  $o(Nn^3)$  products, as the other algorithms in [5, 6, 10, 13]. If the sample size  $T$  of ‘‘minibatches’’ is large enough, the operation quantity will be much smaller than the others. But considering the tracking performance especially in nonstationary environment, the sample size of ‘‘minibatches’’ should be selected moderately.

## 4. Simulation Results and Discussion

Here, several sets of simulation results are provided to demonstrate the performance of the proposed algorithm. Generally speaking, comparisons among fixed step-size algorithms, classical variable step-size algorithms and the proposed algorithm in both stationary and nonstationary environments have been carried out.

*Experiment 1.* Comparison between the proposed algorithm and fixed step-size algorithms.

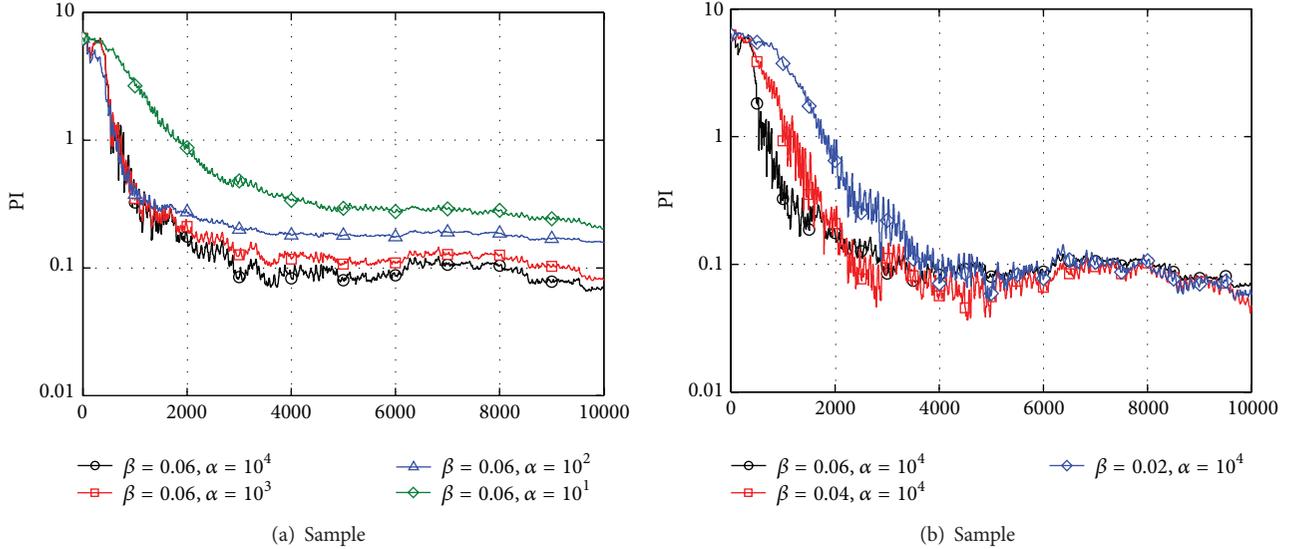


FIGURE 3: Effects of parameters  $\alpha$  and  $\beta$  on the performance of proposed algorithm.

In this experiment, we consider the separation of three zero mean sub-Gaussian sources in stationary environment:

$$\mathbf{s}(t) = [\sin(2\pi 200t), r(t), \text{sign}(\sin(2\pi 100t))], \quad (14)$$

where  $r(t)$  is a random source signal distributed uniformly in  $[-0.5, 0.5]$ . The mixing matrix  $\mathbf{A}$  is randomly generated subject to the normal distribution with mean 0 and standard deviation 1, and three receivers are used ( $m = n = 3$ ). The sampling period is set to 0.0001 s.

To evaluate the performance of the BSS algorithms, we use the cross-talking error as the performance index [5, 20–22]:

$$\text{PI} = \sum_{i=1}^n \left( \sum_{j=1}^n \frac{|g_{ij}|}{\max_k |g_{ik}|} - 1 \right) + \sum_{j=1}^n \left( \sum_{i=1}^n \frac{|g_{ij}|}{\max_k |g_{kj}|} - 1 \right), \quad (15)$$

where the  $n \times n$  matrix  $\mathbf{G} = \mathbf{W}\mathbf{A} = \{g_{ij}\}$  is the combined mixing-separating matrix. As  $\mathbf{W}$  converges to  $\text{PDA}^{-1}$ , the combined mixing-separating matrix  $\mathbf{G}$  will converge to PD, a generalized permutation matrix, and PI will converge to zero.

In the algorithms, activation function  $\phi(x) = x^3$  is applied. The step-size  $\mu = 0.001$  and 0.01 is taken in natural gradient algorithm [6] and optimized EASI algorithm [20], respectively. The parameters of the proposed algorithm are set to  $D = 5$ ,  $g = 0.9$ . Considering the balance between tracking performance and evaluation accuracy of the mean-square-error matrix MSE, the sample size of “minibatches”  $T = 5$  in all the experiments. The effects of crucial parameters  $\beta$  and  $\alpha$  on the performance of proposed algorithm are investigated in Figure 3. What should be noticed is that larger  $\beta$  and  $\alpha$ , respectively, lead to faster initial learning rate and better

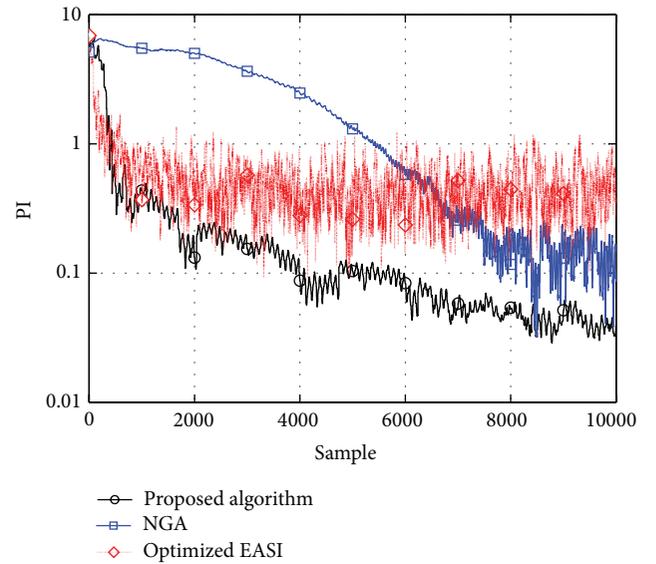


FIGURE 4: Average performance indexes over 500 independent runs of the three algorithms in noise-free case.

convergence performance, so the results provide reference for choosing appropriate parameters  $\beta$  and  $\alpha$ . Hence, we set the parameters  $\beta$  and  $\alpha$  to be 0.06 and  $10^4$ , respectively.

Besides, if sources include both sub-Gaussian and super-Gaussian signals, the activation functions should not be the same increasing odd function. The activation functions might be initialized by polynomials or kernel functions with some adjustable parameters, so that the optimal activation functions vector  $\phi(\cdot) = [\phi_1(\cdot), \dots, \phi_n(\cdot)]^T$  can be estimated adaptively along with the iteration [23]. However, further investigation on the activation functions might go beyond this paper.

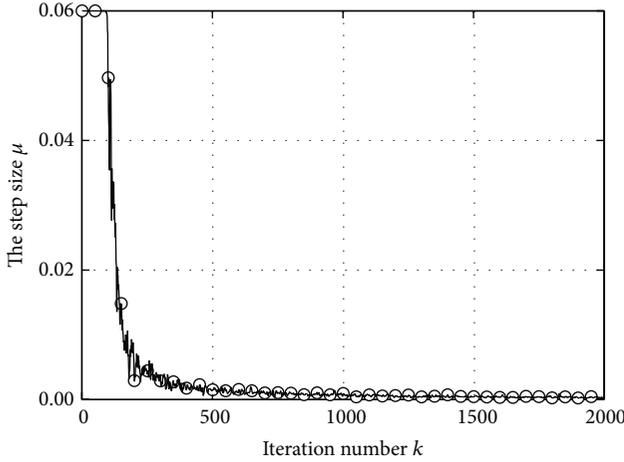


FIGURE 5: Evolution of the step-size  $\mu$  by the proposed algorithm.

Figure 4 plots the average PI value obtained from the simulations of three adaptive algorithms for 500 Monte Carlo trials. From the plots, we can see that the proposed algorithm provides the fastest convergence speed, while achieving lower steady state error than both the NGA and optimized EASI approaches. The step-size  $\mu$ , evolution of which is demonstrated in Figure 5, decreases generally in the exponential type as iterations. We observe that step-size maintains a constant 0.06 during about 100 iterations of the beginning. This is attributed to the choosing of parameter  $\alpha$ , which can benefit for high initial learning rate yet sensitive detection in the state of separating. As result, the separating performance is more robust to the setting of initial learning rate.

*Experiment 2.* Comparison between the proposed algorithm and variable step-size algorithms.

In this experiment, we firstly define the function  $\mathbf{Q}(t) = \phi(\mathbf{y}(t))\mathbf{y}^T(t)$  in (11). In order to allow fair comparison, the same function in [10] is used for the proposed algorithms; that is,

$$\begin{aligned} \mathbf{Q}(t) = & \text{diag} [\phi(\mathbf{y}(t)) \mathbf{y}^T(t)] \\ & + \text{off} [\phi(\mathbf{y}(t)) \text{sign}(\mathbf{y}^T(t))], \end{aligned} \quad (16)$$

where  $\text{diag}[\cdot]$  and  $\text{off}[\cdot]$  denote the operation of taking the diagonal elements and off-diagonal elements of a matrix, respectively, and two zero mean sub-Gaussian sources  $\mathbf{s}(t) = [\sin(2\pi 200t), r(t)]$  are mixed by a  $2 \times 2$  mixing matrix  $\mathbf{A}_0$ ; that is,

$$\mathbf{A}_0 = \begin{bmatrix} 1.0 & 0.5 \\ -0.3 & 0.8 \end{bmatrix}. \quad (17)$$

Zero-mean independent white Gaussian noise is added to the mixture with the signal-to-noise ratio being equal to 20 dB. The parameters including the initial step-size in LMS-type algorithms are manually tuned so each algorithm has nearly the same steady state performance. The initial value

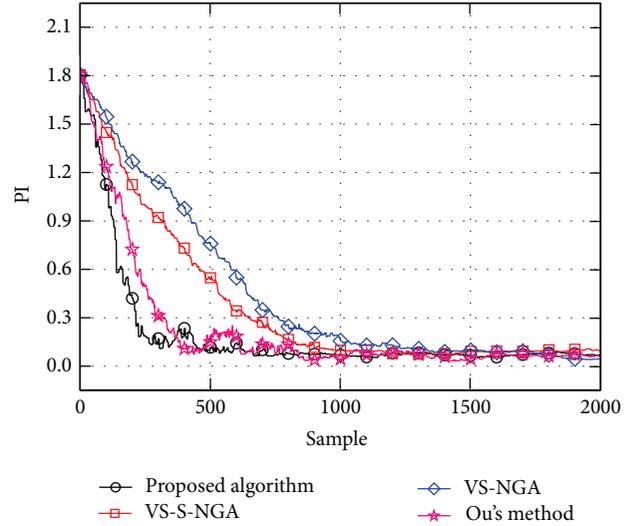


FIGURE 6: Average performance indexes over 500 independent runs of the three variable step-size algorithms in noisy case.

of  $\mu$  for classical variable step-size algorithms, that is, VS-NGA and VS-S-NGA in [10], is set to 0.004,  $\rho = 10^{-5}$ , and 500 Monte Carlo trials are run for averaged performance. The parameters of Ou's method in [13] are set to  $\beta = 0.08$ ,  $\alpha = 10$ . For the proposed algorithm, the parameters are, respectively,  $\beta = 0.08$ ,  $\alpha = 10^3$ ,  $D = 5$ , and  $g = 0.9$ . The parametric settings imply that the proposed algorithm can lead to a higher learning rate while maintaining appropriate steady state performance. The average values of PI resulting from three approaches are compared in Figure 6. The proposed algorithm only requires approximately 500 samples for convergence; however, the other three algorithms need 600 samples at least. Clearly, the performance of the proposed algorithm is considerably improved over the classical variable step-size algorithms in noisy case.

Figure 7 plots the average PI value of three approaches in a nonstationary environment. The mixing matrix to simulate the time-varying environments is chosen as

$$\mathbf{A} = \mathbf{A}_0 + \Xi, \quad (18)$$

where  $\Xi = \eta\Xi + \tau\text{randn}(n)$ ,  $\text{randn}(\cdot)$  is MATLAB built-in function [24], and the initial  $\Xi$  is set to a null matrix. Here,  $\eta = 0.9$  and  $\tau = 0.001$ . The initial parameter for classical variable step-size algorithms is the same as the experiment in noisy case. The parameters of the proposed algorithm are reset to  $\beta = 0.03$ ,  $\alpha = 10^3$ ,  $D = 5$ ,  $g = 0.9$ . Likewise, results are obtained over 500 Monte Carlo runs. From this figure, it is observed that the proposed algorithm converges faster than VS-NGA and VS-S-NGA algorithms in the nonstationary environment.

Finally, we checked the computational time and separation performance for different separation methods in noisy case. The Fast-ICA algorithm, a classical batch-based method, is also utilized for comparison. The sources, mixing matrix, and initial parameters are set as Experiment 2. The data length is set to 10000 samples, which is enough for

TABLE 1: Average PI and execution time for various methods under different level of noise.

Separation Methods	Average PI					Average Execution Time (s)
	Signal to Noise Ratios (dB)					
	0	5	10	20	30	
Optimized EASI ( $\mu = 0.001$ )	0.244	0.229	0.192	0.151	0.137	0.162
NGA ( $\mu = 0.001$ )	0.293	0.238	0.216	0.175	0.155	0.043
VS-NGA	0.305	0.203	0.184	0.148	0.143	0.073
VS-S-NGA	0.263	0.141	0.117	0.092	0.081	0.134
Ou's Method	0.269	0.132	0.114	0.071	0.066	0.241
<b>Proposed Algorithm</b>	<b>0.278</b>	<b>0.156</b>	<b>0.121</b>	<b>0.076</b>	<b>0.064</b>	<b>0.125</b>
Fast-ICA	0.458	0.183	0.129	0.056	0.025	0.551

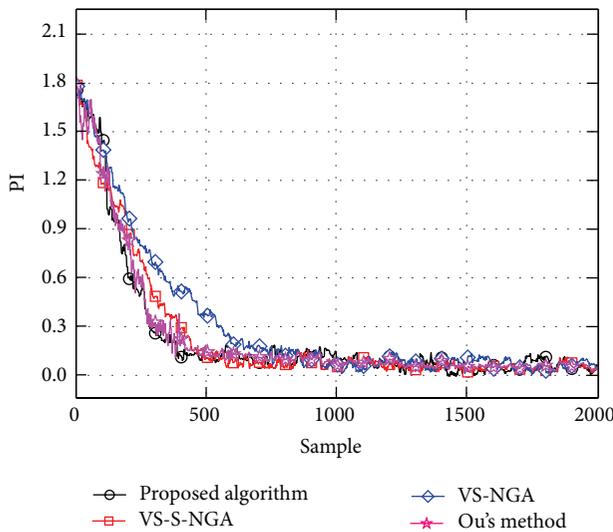


FIGURE 7: Average performance indexes over 500 independent runs of the three variable step-size algorithms in nonstationary environment.

achieving the convergence. The iteration number of Fast-ICA was set to 100. The results are provided in Table 1. It can be seen that Fast-ICA generally has better separation performance under high SNR (signal to noise ratios). However, it costs large computational time since it is a kind of batch-based algorithm, and it will not work until a large number of data samples are received. In contrast, though the proposed algorithm performs slightly worse than Fast-ICA, it behaves better when the noise power is increased (SNR = 0 dB, 5 dB, 10 dB). As a novel adaptive online algorithm, the proposed algorithm has particular advantage due to its computational simplicity and latent ability in tracing the noise and nonstationary environments. It also shows that the proposed algorithm performs even better than optimized EASI, NGA, and VS-NGA in terms of average PI. Similar separation performance with lower computational load is also obtained compared to VS-S-NGA and Ou's method. This demonstrates the complexity analysis in Section 3; that is, the utilization of "minibatches" would probably reduce the computational cost.

## 5. Conclusion

In this paper, we propose a new variable step-size algorithm for blind source separation. Reference separation system is utilized to acquire the mean-square-error matrix which is treated as the metric to update the step-size. As for performance comparison, fixed step-size algorithms, classical variable step-size algorithms, and the proposed algorithm have been carried out in both stationary and nonstationary environments. The performance of the abovementioned approaches is analyzed and compared in terms of cross-talking error. It is revealed that the proposed scheme has improved learning rate and stability performance over the fixed step-size algorithms and converges faster than classical variable step-size algorithms.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

The author would like to thank anonymous referees for constructive comments which are valuable for improving the paper. This work was supported by the National Natural Science Foundation of China under Grant 61172061 and the Natural Science Foundation of JiangSu Province in China under Grant BK2011117. This work was also supported by the National Natural Science Foundation of China under Grant 61201242 and Grant 60772083.

## References

- [1] P. Comon and C. Jutten, *Handbook of Blind Source Separation: Independent Component Analysis and Applications*, chapter 1, Elsevier Press, New York, NY, USA, 2010.
- [2] G. S. Fu, R. Phlypo, M. Anderson, X. L. Li, and T. Adali, "Blind source separation by entropy rate minimization," *IEEE Transactions on Signal Processing*, vol. 62, no. 16, pp. 4245–4255, 2014.
- [3] B. Loesch and B. Yang, "Cramér-Rao bound for circular and noncircular complex independent component analysis," *IEEE*

- Transactions on Signal Processing*, vol. 61, no. 2, pp. 365–379, 2013.
- [4] J.-T. Chien and H.-L. Hsieh, “Nonstationary source separation using sequential and variational Bayesian learning,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 24, no. 5, pp. 681–694, 2013.
  - [5] J.-F. Cardoso and B. H. Laheld, “Equivariant adaptive source separation,” *IEEE Transactions on Signal Processing*, vol. 44, no. 12, pp. 3017–3030, 1996.
  - [6] S.-I. Amari, “Natural gradient works efficiently in learning,” *Neural Computation*, vol. 10, no. 2, pp. 251–276, 1998.
  - [7] H. H. Yang, “Serial updating rule for blind separation derived from the method of scoring,” *IEEE Transactions on Signal Processing*, vol. 47, no. 8, pp. 2279–2285, 1999.
  - [8] S. Amari, A. Cichocki, and H. H. Yang, “A new learning algorithm for blind signal separation,” in *Proceedings of the Advances in Neural Information Processing Systems (NIPS '96)*, vol. 8, pp. 757–763, 1996.
  - [9] S. C. Douglas and A. Cichocki, “Adaptive step size techniques for decorrelation and blind source separation,” in *Proceedings of the 32nd Asilomar Conference on Signals, Systems & Computers*, vol. 2, pp. 1191–1195, November 1998.
  - [10] L. Yuan, W. Wang, and J. A. Chambers, “Variable step-size sign natural gradient algorithm for sequential blind source separation,” *IEEE Signal Processing Letters*, vol. 12, no. 8, pp. 589–592, 2005.
  - [11] X. D. Zhang, X. L. Zhu, and Z. Bao, “Grading learning for blind source separation,” *Science in China E*, vol. 32, no. 5, pp. 693–703, 2002.
  - [12] S.-T. Hsieh, T.-Y. Sun, C.-L. Lin, and C.-C. Liu, “Effective learning rate adjustment of blind source separation based on an improved particle swarm optimizer,” *IEEE Transactions on Evolutionary Computation*, vol. 12, no. 2, pp. 242–251, 2008.
  - [13] S. F. Ou, X. H. Zhao, and Y. Gao, “Variable step-size blind source separation algorithm with an auxiliary separation system,” *Acta Electronica Sinica*, vol. 37, no. 7, pp. 1588–1593, 2009 (Chinese).
  - [14] H. Nakajima, K. Nakadai, Y. Hasegawa, and H. Tsujino, “Blind source separation with parameter-free adaptive step-size method for robot audition,” *IEEE Transactions on Audio, Speech and Language Processing*, vol. 18, no. 6, pp. 1476–1485, 2010.
  - [15] F. Nesta, T. S. Wada, and B.-H. Juang, “Batch-online semi-blind source separation applied to multi-channel acoustic echo cancellation,” *IEEE Transactions on Audio, Speech and Language Processing*, vol. 19, no. 3, pp. 583–599, 2010.
  - [16] F. Nesta, T. S. Wada, S. Miyabe, and B.-H. Juang, “On the non-uniqueness problem and the semi-blind source separation,” in *Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA'09)*, pp. 101–104, New Paltz, NY, USA, October 2009.
  - [17] M. Z. Bhotto and A. Antoniou, “Robust set-membership affine-projection adaptive-filtering algorithm,” *IEEE Transactions on Signal Processing*, vol. 60, no. 1, pp. 73–81, 2012.
  - [18] X.-L. Zhu and X.-D. Zhang, “Adaptive RLS algorithm for blind source separation using a natural gradient,” *IEEE Signal Processing Letters*, vol. 9, no. 12, pp. 432–435, 2002.
  - [19] Y. X. Chen, L. N. Tho, B. Champagne, and C. J. Xu, “Recursive least squares constant modulus algorithm for blind adaptive array,” *IEEE Transactions on Signal Processing*, vol. 52, no. 5, pp. 1452–1456, 2004.
  - [20] J. M. Ye, H. H. Jin, S. T. Lou, and K. J. You, “An optimized EASI algorithm,” *Signal Processing*, vol. 89, no. 3, pp. 333–338, 2009.
  - [21] T. Adali, M. Anderson, and G. S. Fu, “Diversity in independent component and vector analyses: identifiability, algorithms, and applications in medical imaging,” *IEEE Signal Processing Magazine*, vol. 31, no. 3, pp. 18–33, 2014.
  - [22] P. W. Chen, H. Hung, O. Komori, S.-Y. Huang, and S. Eguchi, “Robust independent component analysis via minimum  $\gamma$ -divergence estimation,” *IEEE Journal on Selected Topics in Signal Processing*, vol. 7, no. 4, pp. 614–624, 2013.
  - [23] Y. Xue, F. Ju, Y. Wang, and J. Yang, “A source adaptive independent component analysis algorithm through solving the estimating equation,” *Expert Systems with Applications*, vol. 36, no. 10, pp. 12306–12313, 2009.
  - [24] G. Marsaglia and W. W. Tsang, “The Ziggurat method for generating random variables,” *Journal of Statistical Software*, vol. 5, no. 8, 2000, <http://www.jstatsoft.org/v05/i08/>.



**Hindawi**

Submit your manuscripts at  
<http://www.hindawi.com>

