Indian Journal of Science and Technology, *Vol 9(39), DOI: 10.17485/ijst/2016/v9i39/100365, October 2016*

ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645

A Novel Blind Source Separation Algorithm using Bussgang Criterion and Natural Gradient

Govind Murmu* **and S. Bhattacharya**

Department of Electronics Engineering, Indian School of Mines, Dhanbad - 826004 Jharkhand, India; govind.mu@gmail.com, subatabh@yahoo.com

Abstract

Objectives: In this paper, a novel algorithm based on Transform Domain Least Mean Square (TDLMS) is proposed for Blind Source Separation (BSS). The proposed algorithm is compared with several other BSS algorithms in detail and the results are discussed extensively. **Methods/Statistical Analysis:** To solve the problem of BSS employing the present approach, the ordinary gradient used in conventional LMS algorithm is replaced by natural gradient on the Stiefel manifold. The natural gradient is computed from a cost function based on Bussgang criterion. The proposed algorithm is compared with previously reported LMS type and Recursive Least Square (RLS) type algorithms for four different performance criteria – cross–talk error convergence, harmonic distortion in recovered signals, average deviation from orthogonality of demixing matrix and time complexity. **Findings**: Using simulations it is found that the proposed algorithm has best cross-talk error-convergence and least harmonic distortion as compared to other algorithms. However, the average deviation from orthogonality for demixing matrix and average simulation time for the proposed algorithm are comparable to LMS-type algorithms and estimates 35% as compared to RLS-type algorithms. **Application/Improvements:** The use of natural gradient and the pre-whitening process improves the performance of the algorithm. This algorithm is applied to separate signals from its mixture.

Keywords: Bussgang Criterion, Bind Source Separation (BSS), Cost-Function, LMS type BSS, Natural Gradient, RLS type BSS

1. Introduction

Blind Source Separation (BSS) consists of separating a set of unobserved source signals from a set of linear mixture of them when the mixing matrix coefficients are not known. A number of algorithms have been developed to solve the BSS problem. One of the classifications of the BSS algorithms is based on their choice of the cost function and minimization/maximization of this cost function. Two cost functions that have been widely used are based on Bussgang criterion[1, 2] and nonlinear PCA criterion[3, 4] Based on these cost functions, Least Mean Square (LMS) type and Recursive Least Square (RLS) type algorithms have been developed[1,4]. The superiority of RLS type algorithms over LMS types is well studied in literature. The better performance of RLS is due to an inherent whitening process of input samples in RLS type algorithm itself. In this paper, a transform domain LMS method is proposed that overcomes the limitations of LMS algorithm and gives better performance than RLS algorithms as well. The proposed algorithm incorporates Discrete Fourier Transform (DFT), for prewhitening of source mixtures and is formulated in the Riemannian space using natural gradient update on the Stiefel manifold which is reported to be superior to ordinary gradient method of BSS applications[5,6]. The proposed LMS-type algorithm has been compared with LMS-type and RLS-type algorithms reported in literature.

Section 2 introduces the problem formulation in BSS using Bussgang criterion and natural gradient over Stiefel manifold. The proposed algorithm is derived in Section 3. Simulations are presented in Section 4. The results of simulation are discussed in Section 5 and conclusions are made in Section 6.

**Author for correspondence*

2. Theory

2.1 Problem Formulation

Let $\mathbf{x}_t = [x_1(t), x_2(t), \ldots, x_n(t)]^T$ be the set of mixtures obtained from a set of source signals $\mathbf{s}_t = [s_1(t), s_2(t), \ldots, s_n(t)]^T$, through an unknown mixing system $\mathbf{H}$, i.e.

$$\mathbf{x}_t = \mathbf{H}\mathbf{s}_t \tag{1}$$

where $\mathbf{H}$ is $n \times n$ invertible matrix on $\mathbf{R}^n$. Let $\mathbf{W}$ be $n \times n$ separating matrix that is an estimate of the inverse of $\mathbf{H}$, so that

$$\mathbf{y}_t = \mathbf{W}\mathbf{x}_t \tag{2}$$

becomes an estimate of source vector $\mathbf{s}_t$. Then the learning problem for (2), with orthonormality constraint on the weight matrix, i.e. $\mathbf{W}\mathbf{W}^T = \mathbf{I}$ ($\mathbf{I}$ is $n \times n$ identity matrix) becomes[7]

$$J(\mathbf{W}_{opt}) = \min_{\mathbf{W} \in N^{n \times n}} J(\mathbf{W}) \tag{3}$$

where $J(\mathbf{W})$ is a cost function.. 2

2.2 Bussgang Criterion

In the Bussgang algorithm, the output signal, $\mathbf{y}_t$, from the tap delay filter is passed through a non-linearity $\mathbf{g}(\mathbf{y})$ as shown in Figure 1.

A Bussgang process has to satisfy[8]

$$E\{y(n)y(n-k)\} = E\{g(y(n))y(n-k)\} \tag{4}$$

where $g(\cdot)$ is a zero-memory nonlinearity and $E[\cdot]$ is the statistical expectation operator. The Bussgang property can be applied to the instantaneous mixing scenario by replacing the time series vector $\mathbf{y}_t$ by a vector of different outputs $y_i(n), i = 1, 2, \ldots, N$ obtained from N different sources. The Bussgang condition in

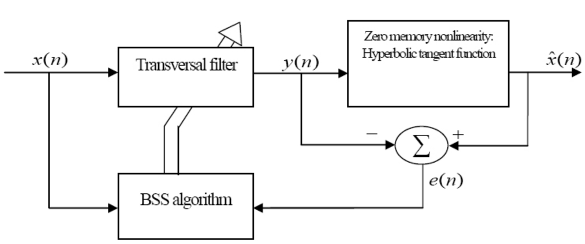


Figure 1. Block diagram of blind equalizer8.

such a case is applied spatially rather than time–wise as[1]

$$E\{g(\mathbf{y}_t)\mathbf{y}_t^T\} = E\{\mathbf{y}_t\, \mathbf{y}_t^T\} \tag{5}$$

2.3 Natural Gradient in the Stiefel Manifold

As mentioned in section 1, natural gradient works more efficiently than ordinary gradient when the underlying space of parameters is not Euclidean[5]. Bussgang cost function (the same is true for nonlinear PCA criterion as well) involves computation of the estimate of the source signal $\widehat{s_k}$, obtained as the output of a zero-memory nonlinearity $\mathbf{g}(\cdot)$. The input to the nonlinear system is the adaptive filter output $\mathbf{y}$ which combines mixed signals linearly through tap weights $\mathbf{W}$. The net result is a cost function that is a nonquadratic function of tap weights. This leads to use of natural gradient in our approach. The natural gradient in Stiefel manifold is given as[9]

$$\tilde{\nabla} J(\mathbf{W}) = \nabla J(\mathbf{W})\mathbf{W}^T\mathbf{W} - \mathbf{W}\{\nabla J(\mathbf{W})\}^T\mathbf{W} \tag{6}$$

where $\nabla J(\mathbf{W})$ is an ordinary gradient.

3. Proposed Algorithm

3.1 Cost function, Gradient and TDLMS Algorithm

The memoryless nonlinear condition of (5) can be rewritten as the cost function for the LMS type algorithm[1]

$$J(\mathbf{W}_t) = E\{\| g(\mathbf{y}_t) - \mathbf{y}_t \|^2\} \tag{7}$$

The cost function associated with nonlinear PCA[9] is

$$J_1(\mathbf{W}_t) = E\left\{\left\|\mathbf{v}_t - \mathbf{W}_t^T g(\mathbf{W}_t\mathbf{v}_t)\right\|^2\right\} \tag{8}$$

where $\mathbf{v}_t$ is the whitened input vector.

Comparison of (7) with (8) shows that (7) is simpler. Though there are other contrast functions like *maximum likelihood*, *negenotropy* etc., they are closely connected to the previous two cost functions under certain conditions. However (7) involves an expectation operator (E), which requires an infinite number of samples, which is not available in practical applications. It is then simplified as weighted norm

$$J(\mathbf{W}_t) = \sum_{k=1}^{t} \beta^{t-k} \| g(\mathbf{y}_k) - \mathbf{y}_k \|^2 \tag{9}$$

where $\|\cdot\|$ is the norm value and $0 < \beta < 1$ is forgetting factor. The cost function obtained in (9) can be compared to the RLS-type algorithm[9] which uses a non-linear PCA criterion. In the proposed algorithm, we have considered minimization of difference in the output explicitly, while this RLS type algorithm[9] considers minimization of the difference between the input and inverse transformation of the desired output, i.e. a nonlinear function of the actual output. The gradient of $J(W_t)$ with respect to W_t is

$$\nabla J(W_t) = \frac{\partial J(W_t)}{\partial W_t} = \sum_{k=1}^{t} \beta^{t-k}[\{[g(y)]_k) - y_k\}x_k^T] \quad (10)$$

Putting (10) into (6) we get

$$\tilde{\nabla} J(W_t) = \sum_{k=1}^{t} \beta^{t-k}\left[\{[g(y)]_k) - y_k\}x_k^T W_t^T W_t - W_t\left\{\{[g(y)]_k) - y_k\}x_k^T\right\}^T W_t\right]$$

Expanding it

$$\tilde{\nabla} J(W_t) = \sum_{k=1}^{t} \beta^{t-k}\{[g(y)]_k)x_k^T W_t^T W_t - W_t x_k g(y_k)^T W_t - y_k x_k^T W_t^T W_t + W_t x_k y_k^T W_t\}$$

$$= \sum_{k=1}^{t} \beta^{t-k}\{[g(y)]_k)y_k^T - y_k g(y_k)^T\}W_t \quad (11)$$

If $\beta = 0$ in (11) then the corresponding stochastic gradient learning algorithms is

$$W_t = W_{t-1} + \eta[[g(y)]_t)y_t^T - y_t [g(y)]_t)^T]W_{t-1} \quad (12)$$

where η is a positive learning rate or step size of the algorithm.

The n-dimensional vector x_t which corresponds to the mixed source signals is pre-whitened to reduce correlation among the observations. The whitened input vector v_t is thus expressed as

$$v_t = Vx_t \quad (13)$$

where V is the pre-whitening matrix.

Next the separating matrix W is applied to v_t to yield

$$y_t = Wv_t = WVx_t = WVH\, s_t \quad (14)$$

Here y_t, the output of BSS, is a permuted and scaled version of the source vector s_t,

$$WVH = \Phi D \quad (15)$$

where Φ is a $n \times n$ permutation matrix and D is $n \times n$ diagonal scaling matrix.

The optimum solution for W satisfies the relation[10,11]

$$W_{opt}W_{opt}^T = I \quad (16)$$

Hence the algorithms are developed in such a way that they adapt W_t to maintain the orthonormality of the rows of W_t.The proposed algorithm is compared with RLS-type algorithms which is detailed in next section for continuity.

3.2 RLS type Algorithm

An RLS type algorithm is formulated by considering $0 < \beta < 1$ in (11).Then from $\tilde{\nabla}[J(W]_t) = 0$ the optimal value of the filter weights are given as

$$W_{opt} = \left[\sum_{k=1}^{t} \beta^{t-k} y_k g(y_k)^T\right]^{-1}\left[\sum_{k=1}^{t} \beta^{t-k} g(y_k) x_k^T\right]$$

$$= R_t^{-1}\mathcal{H}_t \quad (17)$$

where,

$$R_t = \sum_{k=1}^{t} \beta^{t-k} y_k g(y_k)^T \quad \text{and} \quad \mathcal{H}_t = \sum_{k=1}^{t} \beta^{t-k} g(y_k) x_k^T$$

The orthonormality constraint on the weight matrix W has been used to arrive at the expression for W_{opt}, as given in (16). Now R_t and $\mathcal{H}_t$ can be calculated iteratively as

$$R_t = \beta R_{t-1} + y_t g(y_t)^T \quad (18)$$

$$\mathcal{H}_t = \beta \mathcal{H}_{t-1} + g(y_t) x_t^T \quad (19)$$

Using matrix inverse lemma and taking $P_t = R_t^{-1}$, the RLS algorithm for BSS with Bussgang cost function is given as

$$y_t = W_{t-1}\, x_t$$

$$W_t = W_{t-1} + [P_t[g(y]_t)y_t^T - Q_t y_t[g(y]_t)^T]W_{t-1} \quad (20)$$

where

$$P_t = \frac{1}{\beta}\left[P_{t-1} - \frac{P_{t-1}y_t[g(y]_t)^T P_{t-1}}{\beta + [g(y]_t)^T P_{t-1} y_t}\right]$$

and

$$\mathbf{Q}_t = \frac{\mathbf{P}_{t-1}}{\beta + [\mathbf{g}(\mathbf{y}]_t)^T \mathbf{P}_{t-1} \mathbf{y}_t}$$

4. Simulations

In order to study the performance of transform domain LMS type algorithm for BSS and to do its comparison with other related algorithms, simulations have been performed for signal-level comparison and content-level comparison. All the algorithms were compared for their computational-complexity. Then two performance indices are defined (Cross-talk error and deviation from orthogonality) and algorithms are compared for these indices.

4.1 Signal Level Comparison

The ability to separate signals from a mixture of them is tested with two different source signals.

$$\mathbf{s}_t = [\sin(2\pi 300t + 6\cos(2\pi 60t)), \mathrm{sign}(\cos[(2\pi 155t))]]^T$$

The source signals are among those signals which were taken by previous authors[9,11] for experimentation. These signals are sampled at $8KHz$ and then mixed by an orthogonal matrix which is randomly generated in each ensemble run. Use of orthogonal mixing matrix is justified from the fact that we have taken the weight matrix $\mathbf{W}_t$ as orthogonal (for simplicity) and the transformation used is also orthogonal. The step size is chosen as $\eta = 60 \times 10^{-4}$ and initial weight values of separating filter are $\mathbf{W}_0 = \mathbf{I}$. The nonlinear function $\mathbf{g}(\cdot) = \tanh(\cdot)$ is used in the algorithm.

4.2 Content Level Comparison

The performance of transform domain LMS type algorithm, related to recovery of source signals from the mixture, is studied in comparison to other algorithms. The proposed scheme (LMS-type algorithm with pre-whitened input) has been compared with the following algorithms

(i) RLS algorithm by [9],
(ii)the algorithm by [12],
(iii) the EASI algorithm[13] and
(iv) RLS algorithm by [14].

All algorithms have initial weight values set as $\mathbf{W}_0 = \mathbf{I}$. Step size $\eta = 60 \times 10^{-4}$ is taken for Douglas algorithm[12] and $\eta = 200 \times 10^{-4}$ is kept for EASI[13] and DFT-LMS type BSS algorithm. The forgetting factor $\beta = 0.983$ is same for all RLS-type algorithms. The nonlinear function $\mathbf{g}(\cdot) = \tanh(\cdot)$ is used in all the schemes. Two source signals, sine wave of $800Hz$ (s_1) and $90Hz$ (s_2), are sampled at a sampling frequency of $8KHz$. They are mixed by an orthogonal transform and each of the algorithms is used to separate these mixed signals. The spectrums of the recovered signal are observed to make a content level comparison among the algorithms.

4.3 Computational Level Comparison

The computational complexity of different algorithms is calculated. Assuming that all data are real, the proposed transform domain algorithm requires approximately $6n^2$ number of multiplication and $n(3n-2)$ number of signed addition. Since for higher values of *n*, the number of multiplication is much greater than the number of addition and also average time for multiplication is significantly larger than the average time for addition, a comparison based on only number of multiplication is undertaken. The RLS algorithm[9] requires $(9n^2 + 2n)$ and EASI algorithm[13] requires $(5n^2 + n)$ number of multiplications for one iteration. Thus the computational complexity of transform domain LMS type algorithm is slightly higher than algorithms of Douglas[12] and EASI algorithm[13] but it is much lower than the complexity of the RLS algorithm[9] .

4.4 Performance Indices

As mentioned in Section 2.2, the input to the demixing algorithm is $\mathbf{V}\mathbf{x}_t$ and consequently the total separating matrix is $\mathbf{W}_t\mathbf{V}$ where $\mathbf{W}_t$ is the weight matrix and $\mathbf{V}$ is the pre-whitening matrix. This leads to the demixed signal $\mathbf{y}_t$ being expressed as, $\mathbf{y}_t = \mathbf{W}_t\mathbf{V}\mathbf{x}_t = \mathbf{W}_t\mathbf{V}\mathbf{H}\mathbf{s}_t = \mathbf{G}_t\mathbf{s}_t$ where $\mathbf{G}_t = \mathbf{W}_t\mathbf{V}\mathbf{H}$. Now, the following tests are carried out for the matrices $\mathbf{G}$ and $\mathbf{W}$:

(i) Whether $\mathbf{G}$ is a diagonal matrix, and

(ii) Whether $\mathbf{W}$ is an orthogonal matrix.

The first test determines whether the original signals are exactly recovered and the second test is carried out to validate our assumption about orthogonality of the weight matrix. The extent to which the conditions stated in (i) and (ii) are satisfied is determined by finding out the following parameters:

4.4.1 Cross Talk Error (PI)

The performance index, considered as cross talk error, is defined as[9,15,16]

$$E = \sum_{i=1}^{n}\left(\sum_{j=1}^{n}\frac{|g_{ij}|}{\max_{k}|g_{ik}|}-1\right)+\sum_{j=1}^{n}\left(\sum_{i=1}^{n}\frac{|g_{ij}|}{\max_{k}|g_{kj}|}-1\right)$$

where $\mathbf{G} = \mathbf{WVH} = \{\mathbf{g}_{ij}\}$ is a mixture of mixing-transformation-separating matrix. For better performance, it is expected that PI should be as small as possible.

4.4.2 Average Deviation from Orthogonality

The average deviation of the separating matrix away from orthogonality is expressed as $\mathfrak{D} = \|\mathbf{WW}^{T} - \mathbf{I_D}\|_F$, where $\mathbf{I_D} = \mathbf{diag}(\mathbf{WW}^{T})$ is diagonal matrix composed by the diagonal elements of $\mathbf{WW}^{T}$, and $\|\cdot\|_F$ denotes the Frobenius norm[16-18] which is calculated as

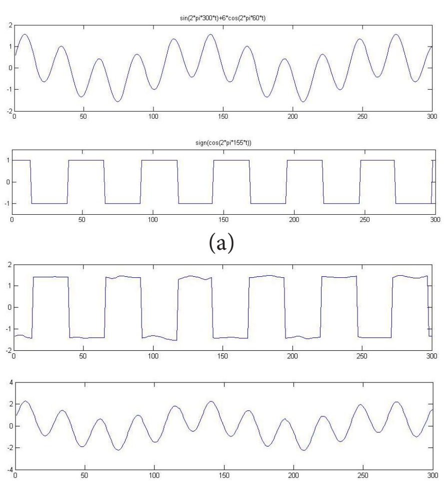
$$\mathfrak{D} = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}|d_{ij}|^{2}}$$

A salient feature of this norm is that it is invariant under rotations.

5. Results and Discussions

The original and recovered signals, obtained with the proposed scheme are shown in Figure 2. The observations from this figure tell that the recovered signals are almost identical to the original signal with the exception that they are scaled versions of original signals. The case of permutation also exists. Research to minimize the permutation and scaling problem is also an active area of research and currently do not cover the scope of this paper.

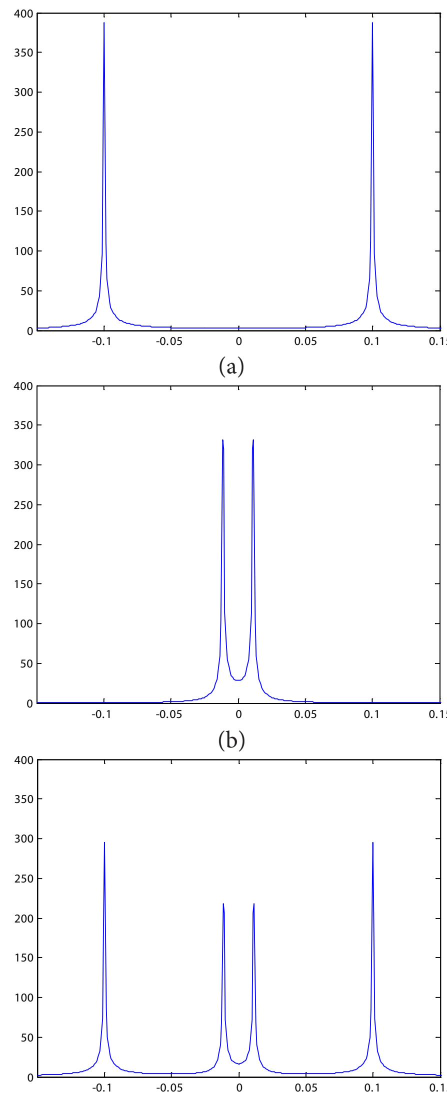
To show the relative separation performance of proposed algorithm with other algorithms the spectrums of recovered signals were compared. The spectrums are shown in Figure 3. Figure 3a and 3b are the spectrums of the original signals s_1 and s_2, respectively. Next in Figure 3 (c-l), the left column and right column show the spectrums of recovered s_1 signal and s_2 signal respectively, obtained using the above mentioned algorithms. It can be observed that recovered s_1 has some components of s_2 and vice versa. This distortion in recovery of



(a)

(b)

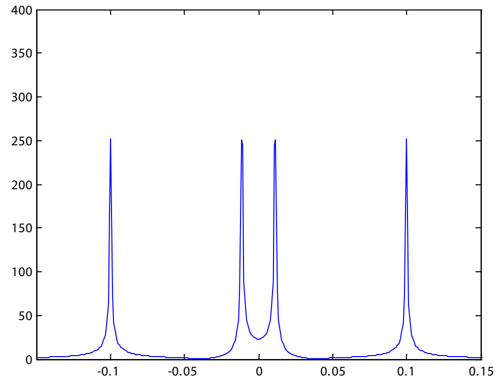
Figure 2. Original and recovered signal obtained in simulation.

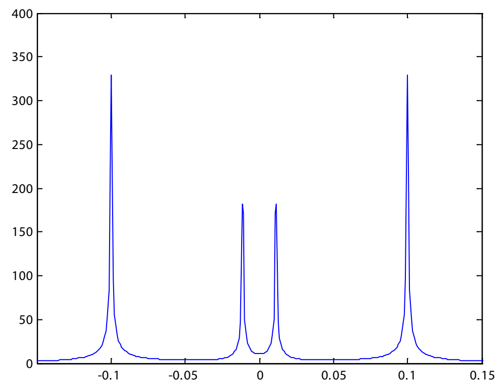


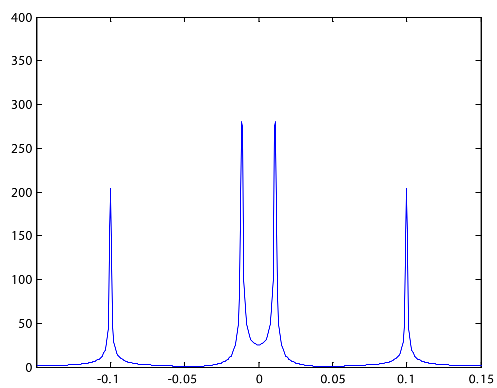
(a)

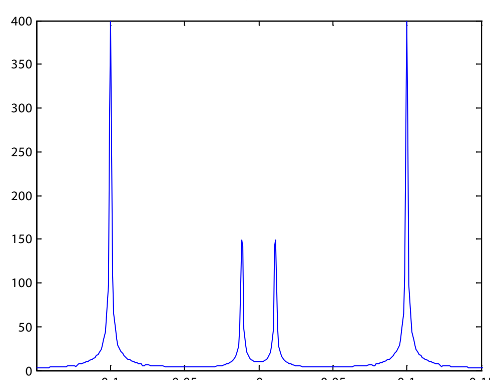
(b)

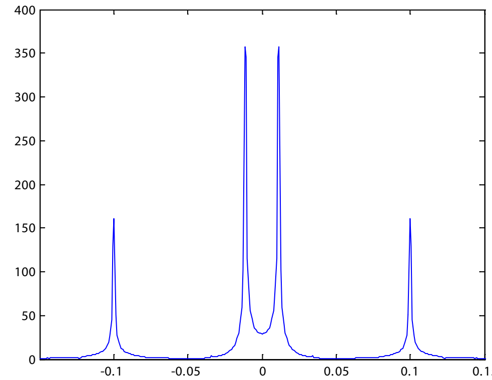
(c)

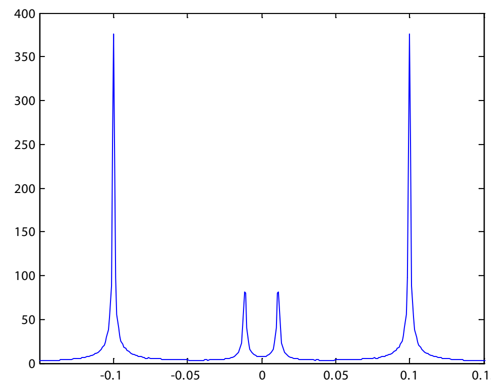

0 50 100 150 200 250 300 350 400
-0.1 -0.05 0 0.05 0.1 0.15
(d)

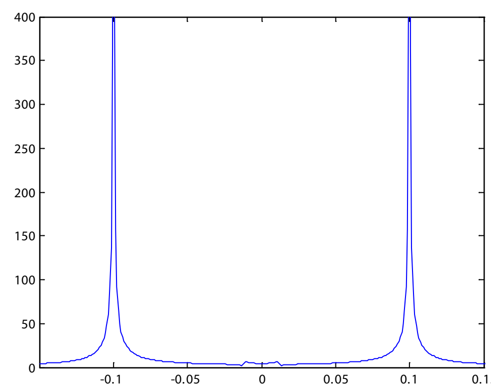

0 50 100 150 200 250 300 350 400
-0.1 -0.05 0 0.05 0.1 0.15
(e)

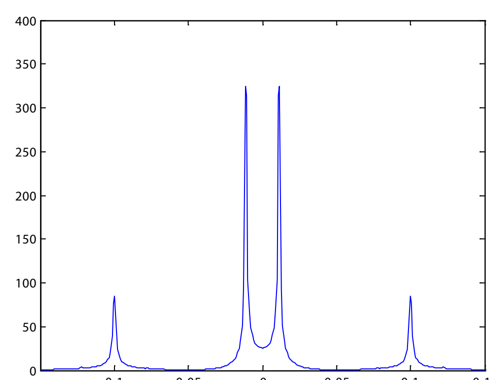

0 50 100 150 200 250 300 350 400
-0.1 -0.05 0 0.05 0.1 0.15
(f)

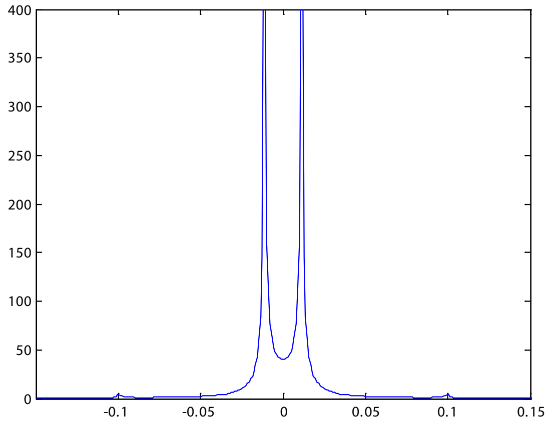

0 50 100 150 200 250 300 350 400
-0.1 -0.05 0 0.05 0.1 0.15
(g)


0 50 100 150 200 250 300 350 400
-0.1 -0.05 0 0.05 0.1 0.15
(h)


0 50 100 150 200 250 300 350 400
-0.1 -0.05 0 0.05 0.1 0.15
(i)


0 50 100 150 200 250 300 350 400
-0.1 -0.05 0 0.05 0.1 0.15
(k)


0 50 100 150 200 250 300 350 400
-0.1 -0.05 0 0.05 0.1 0.15
(j)



(l)

Figure 3. Spectral comparison of recovered signal by different algorithms (**a**) Original s_1 (**b**)Original s_2 (**c**) Recovered s_1 LMS-type algorithm by Douglas[12] (**d**) Recovered s_2 (**e**) Recovered s_1 LMS-type EASI algorithm[13] (**f**) Recovered s_2 (**g**) Recovered s_1 RLS-type algorithm by Pajunen[14] (**h**) Recovered s_2 (**i**) Recovered s_1 RLS type algorithm by X. L. Zhu et. al.[9] (**j**) Recovered s_2 (**k**) Recovered s_1 Proposed TDLMS algorithm (**l**) Recovered s_2

original signal is minimum with DFT-LMS type algorithm while it is maximum with Douglas algorithm. Since the Douglas algorithm and EASI algorithm are unable to remove the correlation among the mixed signals, their performance in recovery of the source signals is poor compared to the RLS type algorithms which can do some decorrelation. The best recovery shown by the DFT-LMS type algorithm is due to more de-correlation of the mixed signals. The harmonic distortion in the recovered signals with different algorithms is calculated and presented in Table 1. From the table, it is clear that harmonic distortion is minimum with DFT-LMS type algorithm which is clearly demonstrated by the spectrums of recovered signal. Hence Figure 3 and Table 1 show the superiority of transform domain LMS type algorithm over other algorithms.

For the defined performance indices the results are shown in Figure 4 (for the average PI of 100 ensemble runs) and Figure 5 (for average deviation from orthogonality of weight matrix).

From these figures, it can be observed that:

(i) In general, RLS type BSS algorithms converge faster than LMS type BSS algorithms.

(ii) However the best convergence is observed in transform domain LMS type BSS algorithm. DFT-LMS type BSS algorithm starts with lower values of PI and have same steady state value as other algorithms.

(iii) The average deviation of weights, **W,** from orthogonality for RLS type algorithm[14] by starts with higher values but it soon converges to near diagonal matrix while the weight matrix **W** for the proposed DFT-LMS type algorithm is having small average deviation from diagonal matrix similar to the rest of the algorithms.

Table 1. Total harmonic distortion (%) in the recovered s_1 and s_2

Recovered Signals	LMS-type algorithms		RLS-type algorithms		TDLMS type algorithm (using DFT)
	EASI	Douglas	Pajunen	Zhu	
s_1	35.9	51.1	30.7	18.0	2.54
s_2	51.4	68.0	34.3	20.0	2.86

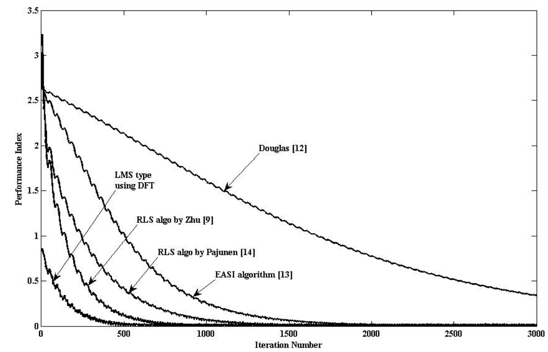


Figure 4. Average cross talk error over 100 ensembles versus iteration number

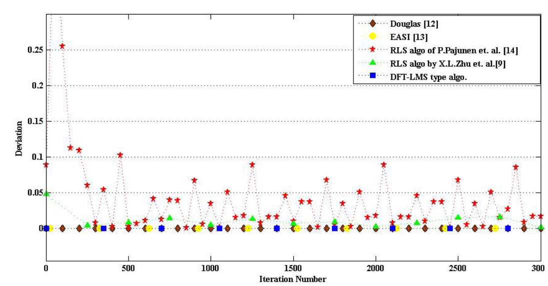


Figure 5. Average deviation of the separating matrix away from orthogonality over 100 ensembles versus iteration number.

Summing up, from the above observations, the performance index for the proposed scheme has better

Table 2. Average simulation time (in seconds) for different BSS algorithms (duration of the source signals is 0.32s)

EASI	Douglas	Zhu (RLS)	TDLMS type BSS (using DFT)
0.0457	*0.0377*	*0.0783*	*0.0499*

convergence performance than the other algorithms while the deviation from orthogonality is similar to the others.

Time taken by different algorithms, during simulation, was also measured and presented in Table 2. The duration of the signals was $0.32s$ i.e. 2560 samples at 8KHz sampling frequency were taken for each source signal. From Table 2, it is observed that the LMS-type algorithms (Douglas[12], EASI[13] and our proposed LMS-DFT algorithm) take less time for computation than the RLS-type algorithm[9]. This is expected because more computation is involved in RLS-type algorithm, primarily due to calculations of the terms $\mathbf{P}_t$ and $\mathbf{Q}_t$. Among the LMS-type ones, DFT-LMS algorithm takes a bit higher computation time than the others which may be attributed to the computation of complex Fourier coefficients.

6. Conclusions

Derivation of an LMS-type and a RLS-type algorithm, using Bussgang criterion and natural gradient on Stiefel manifold, has been presented in this paper. Simulation results for these algorithms have been compared with that for a few related algorithms. It is observed that the LMS-type algorithm which takes pre-whitened input, obtained using an orthogonal transform (DFT), performs the best when evaluated in terms of the standard parameters- performance index (which measures the cross-talk in mixing-separating matrix) and average deviation from orthogonality of the weight matrix. The DFT-LMS type algorithm also recovers the original signals with minimum harmonic distortion compared to the others and its computational complexity is lower than its nearest contender, the RLS algorithm[9], but slightly higher than the LMS-type algorithms[12,13]. Actual time taken during simulation was measured for different algorithms and it is observed that the DFT-LMS type algorithm takes more time than the other LMS-type algorithms but takes less time than RLS-type algorithm, as expected from calculation of computational complexity.

7. Acknowledgements

The authors acknowledge Prof. Vishnu Priye, Dean, R&D, ISM Dhanbad, for the fruitful discussions and help in writing this paper.

8. References

1. Elsabrouty M. Fast converging blind signal separation algorithm using the bussgang cost function and the natural gradient. Int Conf on Signal Proc and Comm, Dubai. 2007. p. 229–32.
2. Panci G, Campisi P, Colonnese S, Scarano G. Multichannel Blind Image Deconvolution using the Bussgang Algorithm: Separation and Multiresolution Approaches. IEEE Trans. on Image Process. 2003; 12(11):1324–37.
3. Karhunen J, Pajunen P, Oja E. The nonlinear PCA criterion in Blind Source Separation: Relationship with other approaches. Neurocomputing. 1998; 22(1–3):5–20.
4. Sivakumar R, Deepa P, Sankaran D. A study of BFO Algorithm based PID controller design for MIMO process using Various Cost Functions. Indian Journal of Science and Technology. 2016; 9(12):1–6.
5. Amari S. Natural Gradient works efficiently in learning. Neural Computing. 1998; 10(2):251–76.
6. Joho M, Mathis H, Moschyte GS. Combined Blind/Nonblind Source Separation based on the Natural Gradient. IEEE Signal Process. Letters. 2001; 8:236–38.
7. Amari S, Cichocki A, Yang HH. A New Learning Algorithm for Blind Source Separation. Advances in Neural Info Process Sys. 1996; 8:757–63.
8. Haykin S. Adaptive Filter Theory. 3rd Ed Prentice Hall, NJ, 2002.
9. Zhu XL, Zhang XD, Ding ZZ, Jia Y. Adaptive Nonlinear PCA algorithms for Blind Source Separation without Prewhitening. IEEE Trans on Circuits and Sys–I. 2006; 53(3):745–53.
10. Edleman A, Arias T, Smith ST. The Geometry of Algorithms with Orthogoinality Constraints. SIAM Journal Matrix Analysis Applications. 1998; 20(2):303–53.
11. Dhar A, Senapati A, Roy JS. Direction of Arrival Estimation for Smart Antenna using a Combined Blind Source Separation and Multiple Signal Classification Algorithm. Indian Journal of Science and Technology. 2016; 9(18):1–8.
12. Douglas SC. Self stabilized Gradient Algorithms for Blind Source Separation with Orthogonality Constraints. IEEE Trans Neural Networks. 2000; 11(6):1490–97.
13. Cardoso JF, Laheld BH. Equivariant Adaptive Source Separation. IEEE Trans. Signal Process. 1996; 44(12):3017–30.

14. Pajunen P, Karhunen J. Least–square methods for Blind Source Separation based on Nonlinear PCA. Int Journal Neural Sys. 1998; 8(6):601–12.
15. Golub H, Loan VCF. Matrix Computations. 3rd Ed. The John Hopkins Univ. Press. Maryland, 1996; 1–723.
16. Cheney W, Kincaid D. Linear Algebra: Theory and Applications. 2nd Ed. Jones and Barlett Learning, Canada. 2012; 1–80.
17. Pannirselvan R, Gopalakrishnan N. Development of Cost Function for Sewage Treatment Plant based on Conventional Activated Sludge Process. Indian Journal of Science and Technology. 2015; 8(30):1–6.
18. Douglas SC. Self stabilized Gradient Algorithms for Blind Source Separation with Orthogonality Constraints. IEEE Trans Neural Networks. 2000; 11(6):1490–97.