

September 2003

## Flexible multichannel blind deconvolution, an investigation

Ah Chung Tsoi  
*University of Wollongong*, [act@uow.edu.au](mailto:act@uow.edu.au)

L. S. Ma  
*University of Wollongong*

Follow this and additional works at: <https://ro.uow.edu.au/infopapers>



Part of the [Physical Sciences and Mathematics Commons](#)

---

### Recommended Citation

Tsoi, Ah Chung and Ma, L. S.: Flexible multichannel blind deconvolution, an investigation 2003.  
<https://ro.uow.edu.au/infopapers/103>

---

## Flexible multichannel blind deconvolution, an investigation

### Abstract

In this paper, we consider the issue of devising a flexible nonlinear function for multichannel blind deconvolution. In particular, we consider the underlying assumption of the source probability density functions. We consider two cases, when the source probability density functions are assumed to be unimodal, and multimodal respectively. In the unimodal case, there are two approaches: Pearson function and generalized exponential function. In the multimodal case, there are three approaches: mixture of Gaussian functions, mixture of Pearson functions, and mixture of generalized exponential functions. It is demonstrated through an illustrating example that the assumption on the source probability density functions gives rise to different performances of source separation algorithms for the multichannel blind deconvolution problem. Further it is observed that these performance differences are not large, indicating that the current formulation of multichannel blind deconvolution problems is robust with respect to the underlying assumption of source probability density functions. It is further speculated that one of the discriminating features among various source separation algorithms appears to be the relative computational efficiencies of various approximation schemes. In other words, the discriminating feature of various source separation algorithms based on assumptions on the source probability density function appears to be an implementation issue rather than one of a theoretical concern.

### Keywords

blind source separation, deconvolution, probability

### Disciplines

Physical Sciences and Mathematics

### Publication Details

This paper originally appeared as: Tsoi, AC & Ma LS, Flexible multichannel blind deconvolution, an investigation, IEEE 13th Workshop on Neural Networks for Signal Processing, 17-19 September 2003, 349-358. Copyright IEEE 2003.

## **FLEXIBLE MULTICHANNEL BLIND DECONVOLUTION, AN INVESTIGATION**

Ah Chung Tsoi, and Liang Suo Ma  
Office of Pro-Vice-Chancellor (IT)

University of Wollongong, Northfields Ave., Wollongong, NSW 2522, Australia

**Abstract.** In this paper, we consider the issue of devising a flexible nonlinear function for multichannel blind deconvolution. In particular, we consider the underlying assumption of the source probability density functions. We will consider two cases, when the source probability density functions are assumed to be unimodal, and multimodal respectively. In the unimodal case, there are two approaches: Pearson function and generalized exponential function. In the multimodal case, there are three approaches: mixture of Gaussian functions, mixture of Pearson functions, and mixture of generalized exponential functions. It is demonstrated through an illustrating example that the assumption on the source probability density functions gives rise to different performances of source separation algorithms for the multichannel blind deconvolution problem. Further it is observed that these performance differences are not large, indicating that the current formulation of multichannel blind deconvolution problems is robust with respect to the underlying assumption of source probability density functions. It is further speculated that one of the discriminating features among various source separation algorithms appears to be the relative computational efficiencies of various approximation schemes. In other words, the discriminating feature of various source separation algorithms based on assumptions on the source probability density function appears to be an implementation issue rather than one of a theoretical concern.

### **1. INTRODUCTION**

Blind Source Separation (BSS) and multichannel blind deconvolution (MBD) have attracted much attention in recent years among signal processing researchers since the publication of the seminal paper by Bell and Sejnowski [2] demonstrating the application of neural network formulation to this problem. Since then, there are various major contributions to this problem, using various approaches, e.g., contrast function [6], infomax [6], natural gradient

method [1], negentropy [6]. Most of these approaches concentrate on the estimation of the parameters of the demixer system to separate the sources. It was shown in [2] that the parameter estimation problem is inherently nonlinear, giving rise to the need of a nonlinear function in the parameter estimation algorithm<sup>1</sup>. There are various approaches to obtain this nonlinearity function, e.g., cumulant, spline functions<sup>2</sup>, hyperbolic tangent function [2]. One of the formulation of the MBD problem is through the minimization of a Kullback-Liebler divergence function [11], when it is required to place an assumption on the source probability density function (pdf). This formulation makes it clear what type of assumptions can be placed on the source pdfs. There are various approaches to resolving this issue, viz., unimodal and multimodal assumptions. In the unimodal pdf assumption, the pdf of the sources are assumed to be unimodal. One may approximate the unimodal pdf by a Pearson function [4] (which can be unimodal or multimodal, dependent on the parameters describing the function), a generalized exponential function [3, 5, 8]. Alternatively, in the multimodal case, the source pdfs are assumed to be multimodal. In this case, there is so far only one approach, viz., approximating the source pdf using a mixture of Gaussian functions [9].

In this paper, we will concentrate on the MBD situation<sup>3</sup>. We will examine the following issues surrounding the underlying assumption of the source pdfs. We first examine the assumption that the source pdfs are unimodal in Section 3. In this case, there are two approaches, viz., a Pearson function, and a generalized exponential function. Then in Section 4, we will extend these cases to the situation when we assume the source pdfs are multimodal to obtain three corresponding cases: mixture of Gaussian functions, mixture of Pearson functions, and mixture of generalized exponential functions. The consideration given in this section, except for the case of mixture of Gaussian functions, is new, as far as we are aware. We will describe the corresponding parameter estimation algorithms in Section 5. Then we will investigate the issue of the mismatch between the assumption on source pdfs and the underlying "real" source pdfs in Section 6 through an illustrating example. The types of questions which we seek answers for include, for example, what happens if the underlying "real" source pdfs are multimodal, and that we assume the source pdfs are unimodal instead. Similarly what happens if the underlying "real" source pdfs are multimodal, and the assumption on source pdfs is multimodal. What are the performance degradations through these mismatch of assumptions. Then we draw a number of observations from our experiments in Section 7. Except in the case of mixture of Gaussians in the multimodal assumption case, the observations from the experiments conducted appear to be new, as far as we are aware. A brief summary of our findings will be given in Section 8.

<sup>1</sup>The parameter estimation algorithm is nonlinear incorporating a nonlinear function  $\varphi(\cdot)$  which will be studied in this paper.

<sup>2</sup>In this paper, we will consider the spline function based nonlinearity, nor the cumulant based studies.

<sup>3</sup>We study the MBD case because it is more complex. Secondly, the corresponding independent component analysis case is a proper subset of the MBD case.

## 2. BRIEF DESCRIPTION OF KULLBACK-LIEBLER DIVERGENCE FUNCTION APPROACH

Given a number of sources,  $s_i$ ,  $i = 1, 2, \dots, n$ . It is assumed that the sources are not available to the sensors, they are independent, and at most one of them is Gaussian distributed [2]. The sources are mixed together by a mixer  $\mathcal{M}$ , which is an unknown linear time invariant dynamical system  $\mathcal{S}$ , which is described by a set of unknown constant parameters  $\Theta$ . In a general manner, the source and the sensors are governed by the following relationship  $\mathbf{u} = \mathcal{S}(\mathbf{s})$ , where  $\mathbf{u}$  and  $\mathbf{s}$  are respectively the vectors denoting the sensor outputs and the sources. The notation  $\mathcal{S}(\cdot)$  is a general description of the dynamical relationship between  $\mathbf{u}$  and  $\mathbf{s}$ . For simplicity we will assume that there are equal number of sources and sensors. The multichannel blind deconvolution problem is to find a demixer  $\mathcal{D}$ , which can recover the sources. We will denote the output of the demixer as  $\mathbf{y}$ . Note that in MBD case, the demixer  $\mathcal{D}$  is a dynamical system. Note further that in this paper, the vectors  $\mathbf{s}$ ,  $\mathbf{u}$ , and  $\mathbf{y}$  all have the same dimension.

One way in which the problem can be resolved is using the following approach. We measure the dependence among the recovered sources  $\mathbf{y}$  using mutual information. Given  $P(\mathbf{y})$ , the probability density function of the recovered signal vector  $\mathbf{y}$ , the mutual information between the recovered signals can be defined as follows:

$$I(\mathbf{y}) = \int P(\mathbf{y}) \frac{P(\mathbf{y})}{\prod_{q=1}^n P(y_q)} d\mathbf{y} = -H(\mathbf{y}) + \sum_{q=1}^n H(y_q), \quad (1)$$

where  $H(\mathbf{y}) = -E[\log(P(\mathbf{y}))]$  is the entropy of  $\mathbf{y}$ ,  $H(y_q) = -E[\log(P(y_q))]$  is the marginal entropy of  $y_q$ ,  $q = 1, 2, \dots, n$ . Observe that  $I(\mathbf{y}) \geq 0$ , and  $I(\mathbf{y}) = 0$  if and only if the components of vector  $\mathbf{y}$  are statistically independent. Therefore  $I(\mathbf{y})$  is an appropriate measurement of the dependence among the recovered signals. Unfortunately, mutual information is difficult to compute explicitly, hence we use a cost function similar to [11]:

$$l(\mathbf{y}, \Omega) = -\log |\det(\mathbf{H}_0)| - \sum_{q=1}^n \log P(y_q), \quad (2)$$

where  $\Omega$  is the set of system parameters of the demixer and source model parameters,  $\det(\cdot)$  is the determinant and  $\mathbf{H}_0$  is the zeroth order Markov parameter [11]. If we assume that the linear time invariant dynamical system is modelled by a state space model:  $\mathbf{x}(t+1) = \mathbf{x}(t) + B\mathbf{s}(t)$ , and  $\mathbf{u}(t) = C\mathbf{x}(t) + D\mathbf{s}(t)$ , where  $\mathbf{x}$  the state is a  $N$  dimensional vector <sup>4</sup>, and  $A$ ,  $B$ ,  $C$ , and  $D$  are respectively constant matrices of appropriate dimensions. Then,

<sup>4</sup>In this paper, we assume that the dimension of the state vector  $N$  is known *a priori*. This can be determined by a number of methods, e.g., using the balanced realization as indicated in [10].

$\mathbf{H}_0 = D$ . We can easily obtain the following parameter updating rules for the matrix  $D$ :

$$\Delta D = \eta(k)(I - \varphi(\mathbf{y})\mathbf{u}^T D^T)D, \quad (3)$$

where  $\varphi(\cdot)$  is a vector nonlinearity related to the source model.  $\varphi(y_q) = -\frac{d \log P(y_q)}{dy_q}$ .

There are a number of possible assumptions for the nonlinearity  $\varphi(\cdot)$ . The simplest assumption is that it is a fixed nonlinearity [2], e.g.,  $\tanh(\cdot)$ . A more complex assumption would be that the sources can be approximated by a symmetrical probability density function (pdf), which will be considered in Section 3. An even more complex assumption would be that the source pdfs are multimodal; this case will be considered in Section 4.

### 3. UNIMODAL ASSUMPTIONS

In this section, we will consider the situation when the source pdfs are assumed to be a unimodal pdf. There are the following approaches:

- Pearson function [4]:  $P(y_q) = \frac{1}{2} [\mathcal{N}(\mu_q, \sigma_q^2) + \mathcal{N}(-\mu_q, \sigma_q^2)]$ , where  $q = 1, 2, \dots, n$ ,  $\mathcal{N}(\mu, \sigma^2)$  denotes a Gaussian function with mean  $\mu$  and variance  $\sigma^2$ . This function could be a unimodal function or a bimodal function [4] dependent on the parameters  $\mu$  and  $\sigma^2$ . In this case, the nonlinear function is given by:

$$\varphi(y_q) = -\frac{d \log P(y_q)}{dy_q} = \frac{y_q}{\sigma_q^2} - \frac{\mu_q}{\sigma_q^2} \tanh\left(\frac{\mu_q}{\sigma_q^2} y_q\right). \quad (4)$$

- Generalized exponential function [5, 3]:  $P(y_q) = \frac{R_q \beta_q^{\frac{1}{R_q}}}{2\Gamma(\frac{1}{R_q})} \exp(-\beta_q |y_q|^{R_q})$ , where  $\Gamma(\cdot)$  is a gamma function. This function has a zero mean, variance determined by  $\frac{1}{\beta_q}$ , and a kurtosis determined by  $R_q$ . In this case, the nonlinear function  $\varphi$  given by:

$$\varphi(y_q) = -\frac{d \log P(y_q)}{dy_q} = \beta_q R_q |y_q|^{R_q-1} \text{sign}(y_q), \quad (5)$$

where  $\text{sign}(y_q) = 1$  for  $y_q \geq 0$ , and  $-1$  for  $y_q < 0$ .

This nonlinear function  $\varphi(\cdot)$  is then used in the parameter update algorithm, e.g., in Eq(3). Note that  $\varphi(\cdot)$  devised in this manner is a time varying nonlinearity, as it depends on the values of  $y_q$ , as well as on other parameters.

Note that because of our assumption on the source pdfs in that at the most only one of the sources is Gaussian distributed [2], it does not make sense to assume the source pdfs to be approximated by a single Gaussian function.

#### 4. MULTIMODAL ASSUMPTIONS

In this case, the source pdfs are assumed to be multimodal in shape. There are three possible approaches:

- Mixture of Gaussians [9]:  $P(y_q) = \sum_{i=1}^P C_{iq} \mathcal{N}(\mu_{iq}, \sigma_{iq}^2)$ ,  $q = 1, 2, \dots, n$ .  $C_{iq}$  are constant but unknown mixing weights.
- Mixture of Pearson functions:

$$P(y_q) = \sum_{i=1}^P \frac{1}{2} C_{iq} [\mathcal{N}(\mu_{iq}, \sigma_{iq}^2) + \mathcal{N}(-\mu_{iq}, \sigma_{iq}^2)],$$

where  $C_{iq}$  are unknown constants.

- Mixture of generalized exponential functions:

$$P(y_q) = \sum_{i=1}^P C_{iq} \frac{R_{iq} \beta_q^{\frac{1}{R_{iq}}}}{2\Gamma(\frac{1}{R_{iq}})} \exp(-\beta_{iq} |y_{iq}|^{R_{iq}}),$$

where  $C_{iq}$  are unknown constants.

The nonlinear function  $\varphi(y_q) = -\frac{d \log P(y_q)}{dy_q}$  in each of the above cases can be obtained relatively easily.

While the extension of Lee *et al.* treatment of Pearson function [7] to the mixture of Pearson functions, and the extension of the generalized exponential function [3, 5, 8] to the mixture of generalized exponential functions are relatively straightforward, these do not appear to have been attempted previously. Yet, viewed from the perspective of this paper, these extensions appear to be logical.

#### 5. DERIVATION OF PARAMETER ESTIMATION ALGORITHM

In this section, we will briefly describe the ways how the parameter update algorithms will be derived. Because of the lack of space we will not give the detailed parameter estimation algorithms for each individual case.

Consider Eq(2), the nonlinearity  $\varphi(\cdot)$  is obtained by differentiating the assumed source probability density function. Thus once the probability density function assumption is defined, the differentiation with respect to the parameters can be carried out. The first term in Eq(2) depends on  $D$ , which can be differentiated quite easily. Once these two terms are differentiated, then they can be combined using some simple algebraic manipulations, bearing in mind the natural gradient trick as discussed in [1].

In summary, the update algorithms for  $\Theta$ , the parameters of the demixer are given as follows:

$$\Delta D = \eta(k)(I - \varphi(\mathbf{y})\mathbf{u}^T D^T)D \quad (6)$$

$$\Delta C = -\eta(k)\varphi(\mathbf{y})\mathbf{x}^T \quad (7)$$

$$\Delta A_{ij} = -\eta(k)\varphi^T(\mathbf{y}) \sum_{\ell=1}^N C_{\bullet\ell} \frac{\partial x_\ell}{\partial A_{ij}} \quad (8)$$

$$\Delta B_{iq} = -\eta(k)\varphi^T(\mathbf{y}) \sum_{\ell=1}^N C_{\bullet\ell} \frac{\partial x_\ell}{\partial B_{iq}}, \quad (9)$$

where  $\ell, i, j = 1, 2, \dots, N$ ,  $q = 1, 2, \dots, n$ ,  $C_{\bullet\ell}$  denotes the  $\ell$ -th column vector of matrix  $C$ .  $\eta(k)$  is a learning parameter.  $\eta(k)$  should decay faster than  $\frac{1}{k}$  to guarantee the convergence of the updating algorithm.  $\frac{\partial x_\ell}{\partial A}$ ,  $\frac{\partial x_\ell}{\partial B}$  can be obtained from following:

$$\frac{\partial x_\ell(k+1)}{\partial A_{ij}} = \sum_{m=1}^N A_{\ell m} \frac{\partial x_m(k)}{\partial A_{ij}} + \delta_{\ell i} x_j \quad (10)$$

$$\frac{\partial x_\ell(k+1)}{\partial B_{iq}} = \sum_{m=1}^N A_{\ell m} \frac{\partial x_m(k)}{\partial B_{iq}} + \delta_{\ell i} u_q, \quad (11)$$

where  $\delta_{\ell i}$  is the Kronecker delta function <sup>5</sup>.

Apart from the update algorithms for  $A, B, C, D$ , for each method, there are specific parameters to be updated. For example, for the generalized exponential function method, there are the following parameters to be updated:  $\beta_q, R_q$ . In the case of mixture of Pearson functions, mixture of Gaussian functions, and mixture of generalized exponential functions, there are also the mixing constants  $C_{iq}$ . The parameter update algorithms for these parameters can be obtained quite easily by differentiating the cost function Eq(2). We will not give the details here due to lack of space.

## 6. MISMATCH BETWEEN THE UNDERLYING PROBABILITY DENSITY FUNCTION AND THE SOURCE PROBABILITY DENSITY FUNCTION ASSUMPTION

In this section we will conduct some experiments to investigate the issue of mismatch between the source pdf assumption and the “real” underlying source probability density function <sup>6</sup>. In the experiment, we synthesize two source signals, both of them are bimodal (see Figure 1(a)). The scatter

<sup>5</sup>For an independent component analysis or blind source separation problem, if the relationship between the source and the mixer output is  $\mathbf{u} = D\mathbf{s}$ ,  $D$  is a constant matrix, then the corresponding parameter update algorithm is the same as the one given here. In this case,  $A, B$  and  $C$  are zero.

<sup>6</sup>We have run a number of examples to test the general validity of the observations of this paper. Here we will only show the results of a simple experiment.



diagram of the mixer outputs is given in Figure 1(b). We use the fixed nonlinearity of  $\tanh(\cdot)$  as the dictum to compare the performance of the algorithms with other approaches (see Figure 1(c)). Then we systematically consider the case when the source pdf is assumed to be unimodal, and the source pdf is assumed to be multimodal respectively. In both cases, we use sources which have a multimodal probability density function (see Figure 1(a)). The statistics of the results for each case are summarized in Table 1.

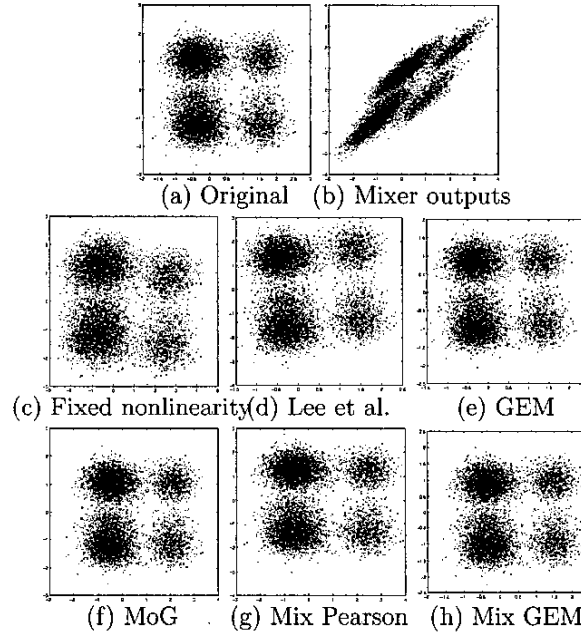


Figure 1: The scatter diagram of (a) the original source signals, (b) the outputs of the mixer system, (c) the one obtained by hyperbolic tangent function, (d) the one obtained by Pearson function, (e) the one obtained by generalized exponential function, (f) the one obtained by mixture of Gaussian functions, (g) the one obtained by mixture of Pearson functions, and (h) the one obtained by using a mixture of penalized exponential functions.

Note that the mean square error (MSE) is the same as the variance because the mean is almost zero.

## 7. OBSERVATIONS

- From Table 1, it is clear that when the “real” underlying source pdf is multimodal, the algorithm based on an assumption of multimodal pdfs performs better than the situation when the algorithm which is based on an assumption that the source pdf is unimodal.
- It is noted that it does not make sense to have a single Gaussian function

Method	source	MSE	Mean	Variance
fixed nonlinearity	source 1	0.0043	2.0830e-005	0.0043
	source 2	0.0089	2.2819e-005	0.0089
Pearson function	source 1	0.0010	-1.1635e-006	0.0010
	source 2	0.0024	7.8791e-006	0.0024
generalized exponential function	source 1	0.0002	-2.9435e-006	0.0002
	source 2	0.0003	-4.6204e-006	0.0003
Mixture of Gaussians	source 1	0.00012	-1.5804e-006	0.00012
	source 2	0.00005	-1.9682e-006	0.00005
Mixture of Pearson functions	source 1	0.00011	-9.9128e-007	0.00011
	source 2	0.00005	-1.3986e-006	0.00005
Mixture of Generalized exponential functions	source 1	0.00004	-8.9346e-007	0.00004
	source 2	0.00009	-1.9902e-007	0.00009

TABLE 1: A TABLE SHOWING THE PERFORMANCE OF VARIOUS METHODS WHEN THERE MIGHT BE A MISMATCH BETWEEN THE ASSUMPTION PLACED ON THE SOURCE PDFS AND THE UNDERLYING SOURCE PDFS.

as the source pdf assumption. This is because such an assumption will contradict directly with one of the assumptions in the MBD formulation [2], viz., at the most only one source signal is Gaussian distributed. Otherwise the source signals cannot be separated.

- It is observed that in the multimodal source pdf assumption situation, the three methods performed equally well, with the variances of errors are all within the same order of magnitude.
- It is surprising to find that while there is noted performance degradation when there is a mismatch between the “real” underlying source pdf and the source pdf assumption, the difference in performance is not significant. Indeed the performance degradation of a fixed nonlinearity is not significant compared with those which gives the best performance, e.g., mixture of generalized exponential functions.
- While it is not shown here, if the source pdf switches from sub-Gaussian to super-Gaussian, all methods as described in this paper continue to work. This is a slight surprise, given the fact that it is the source switching from sub-Gaussian to super-Gaussian which motivated the introduction of Pearson function in Lee *et al.* [7] in the first instance.
- These observations lead us to postulate the following proposition:

**Proposition** The multichannel blind deconvolution problem is robust with respect to the assumption on the nonlinearity, or equivalently the assumption on the source probability density functions.

Unfortunately we are not able to prove this proposition formally.

A corollary of this proposition is that it does not matter whether we use a fixed nonlinear function, e.g., a hyperbolic tangent function, or a mixture of Gaussian functions, the effect on the performance of the parameter update algorithms is secondary, as long as the nonlinear function deployed is one of the following:

- fixed nonlinear function  $\tanh(\cdot)$ .
- a Pearson function
- a generalized exponential function
- a mixture of Gaussians
- a mixture of Pearson functions
- a mixture of generalized exponential functions

This corollary is pleasing in that it does not require a user to know *a priori* the probability density function of the sources. Even if there is a mismatch in the assumption between the source pdfs and the “real” underlying source pdfs, the results will not be affected significantly by the nonlinearity used. This will make the parameter update algorithms attractive to a practitioner, who does not need to paid special attention to the nonlinear function which is being deployed.

- It appears that if the proposition is generally true, then one of the discriminations among various algorithms will be the computational complexity of the algorithm. In other words, which method that a user chooses will be an implementation issue rather than a theoretical issue.
- While it is possible to tune the performance of a particular representation of the nonlinearity in the algorithm with respect to specific examples, we have refrained from doing so, as this is in general dependent on the underlying system.

## 8. CONCLUSIONS

In this paper, we have shown that the assumption of source pdfs contributes in a secondary manner on the formulation of multichannel blind deconvolution problem. It appears that a fixed nonlinearity, e.g., hyperbolic tangent function, or any one of a number of methods, e.g., mixture of Gaussian functions, mixture of Pearson functions, mixture of generalized exponential functions, or their equivalent unimodal counterpart for Pearson function and generalized exponential function performs quite well. Hence it is concluded that this issue of assumption on the source probability density function is an implementation issue rather than one which is a theoretical issue in that the performance of the parameter update algorithms appear to be relatively insensitive to the choice of nonlinear functions, or equivalently to the assumption on source pdfs. In practice which approach one chooses is dependent on

the efficiency of implementation of which a deciding factor may be the computational complexity of the parameter update method. To a practitioner, the results of this paper is pleasing in that it means the practitioner can use any one of the nonlinear functions indicated in this paper, and knows that its effect on the performance of the parameter updating algorithm is secondary.

## 9. ACKNOWLEDGEMENT

The authors acknowledge partial financial support from both the Australian Research Council, and Motorola Australia Research Centre.

## REFERENCES

- [1] Amari, S. (1998). Natural Gradient Working Efficiently in Learning. *Neural Computation*, vol. 10, pp. 251–276.
- [2] Bell, A.J. and Sejnowski, T.J. (1995). An Information Maximization Approach to Blind Separation and Blind Deconvolution. *Neural Computation*, vol. 7, pp. 1129–1159.
- [3] Choi, S., Cichocki, A. and Amari, S. (1998). Flexible Independent Component Analysis. *Proc. IEEE Workshop on Neural Networks for Signal Processing (NNSP'98)*, Cambridge, England. pp. 83–92.
- [4] Douglas, S.C., Cichocki, A. and Amari, S. (1997). Multichannel Blind Separation and Deconvolution of Sources with Arbitrary Distributions. *IEEE Workshop on Neural Networks for Signal Processing*, pp. 436–445.
- [5] Everson, R.M. and Roberts, S.J. (1999). Independent Component Analysis: A Flexible Non-linearity and Decorrelating Manifold Approach. *Neural Computation*, 11(8), pp. 1957–1983.
- [6] Haykin, S. (2000) *Unsupervised Adaptive Filtering*. Vol 1 & 2. J. Wiley & Sons. New York.
- [7] Lee, T-W., Girolami, M. and Sejnowski, T.J. (1999). Independent Component Analysis using an Extended Infomax Algorithm for Mixed Sub-Gaussian and Super-Gaussian Sources. *Neural Computation*, 11(2), pp. 417–441.
- [8] Ma, L.S. and Tsoi, A.C. (2003). Multichannel Blind Deconvolution using a Generalized Exponential Source Model. *4th International Symposium on Independent Component Analysis and Blind Source Separation (ICA2003)*, Nara, Japan, pp.555–560.
- [9] Nakayama, K., Hirano, A. and Sakai, T. (2003). An Adaptive Nonlinear Function Controlled by Estimated Output PDF for Blind Source Separation. *4th International Symposium on Independent Component Analysis and Blind Source Separation (ICA2003)*, Nara, Japan, pp.427–432.
- [10] Tsoi, A.C., and Ma, L.S. (2003). Blind Deconvolution of Dynamical System Using a Balanced Parameterised State Space Approach. *Proc ICASSP 2003*.
- [11] Zhang, L-Q. and Cichocki, A. (2000). Blind Deconvolution of Dynamical Systems: A State Space Approach. *Journal of Signal Processing*, 4(2), pp. 111–130.