# A New Method for Source Separation

Russell H. Lambert
University of Southern California, Electrical Engineering-Systems
3740 McClintock Ave. Room 400, Los Angeles, CA 90089 rlambert@sipi.usc.edu

## ABSTRACT

A new adaptive filtering approach to multichannel source separation is presented. The method is based on an extension of traditional single channel blind equalization. This multichannel extension will be presented theoretically and the results will be demonstrated by simulation using both communications data and speech. The new source separation algorithm is compared to existing adaptive separation algorithms which handle multipath: the Herault-Jutten source separation algorithm [1], and the time-recursive Weinstein et. al. [5] algorithms. The new method offers dramatic performance improvement over existing methods and the ability to handle different source data types in an optimal fashion.

## 1. INTRODUCTION

In recent years the source separation problem has been studied by Cardoso and others [2, 3] who have presented eigenstructure-based methods for source separation. Adaptive filtering approaches such as the Herault-Jutten (HJ) algorithm [1], and the Weinstein et.al. algorithms [5] have also emerged. Recently Cardoso et. al. have introduced adaptive matrix algorithms [4]. While some of the source separation methods are not suited to handle multipath channels, the adaptive filter methods do so in a natural manner.

Blind equalization is an adaptive filter technique which can restore a signal which has been corrupted by a multipath channel back to original condition. A blind cost function $J_b$, is minimized, in much the same way that least mean square (LMS) and recursive least squares (RLS) adaptation (with access to the source) minimize the mean squared error $J = (\hat{x} - x)^2$.

In the blind equalization update equation, an odd nonlinear function $g(\cdot)$ must be found which has been shown to be the *maximum a posteriori* (MAP) estimate of the data plus convolutional noise [6]. The function $g(\cdot)$ and $J_b$ can be found for the a given type of signal [6, 9, 8]. An example of $J_b$ is Gray's Variable

This work was supported in part by the Office of Naval Research under contract No. N00014-91-J-4021.

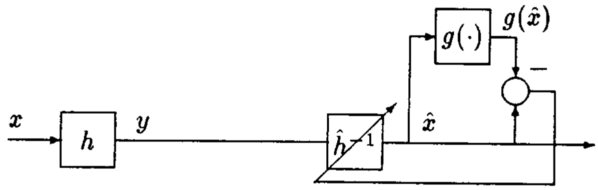


Figure 1: Single Channel Blind Equalization System

Norm $O_\alpha^2 = \frac{E|\hat{x}|^2}{(E|\hat{x}|^\alpha)^{\frac{2}{\alpha}}}$ [8]. $O_4^2$ is equivalent to the Godard or constant modulus criterion $J_b = (\hat{x} - \frac{E|x|^2}{E|x|^4}\hat{x}^3)^2$ [10]. $O_1^2$ is the Sato criterion $J_b = (\hat{x} - \frac{E|x|^2}{E|x|} sign(\hat{x}))^2$ [11]. Using the simple system shown in Fig. 1, we have the following update equations and cost functions: when the reference data $x$ is available as in the LMS algorithm,

$$h^{-1} = h^{-1} + \mu(\hat{x} - x)\underline{y} \qquad J = (\hat{x} - x)^2, \quad (1)$$

and when the reference is not available we can use a blind update,

$$h^{-1} = h^{-1} + \mu(\hat{x} - g(\hat{x}))\underline{y} \qquad J_b = (\hat{x} - g(\hat{x}))^2. \quad (2)$$

Notation convention: the estimate of the deconvolution or inverse filter is $\hat{h}^{-1}$, $\hat{x} = y * \hat{h}^{-1}$, $\underline{y}$ is a vector containing the last $k$ samples of sensor data $y$, where $k$ is the order of the inverse filter.

## 2. MULTICHANNEL EXTENSION OF BLIND EQUALIZATION

Given access to $N$ sensors with an assumed number of sources less than or equal to $N$, all with unknown direct and cross channels, we wish to recover all sources.

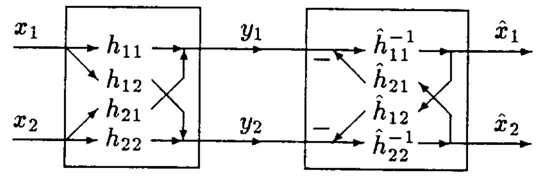


Figure 2: Multichannel Blind Equalization System Using an Inverse System With Feedback

We must decorrelate all the inputs for the source separation problem. If a non-unity direct channel exists we must also invert it thus performing the combined separation/equalization problem.

The multichannel cost function can be made as the simple sum of single channel blind cost functions

$$J_{mb} = J_{b1} + J_{b2} + \ldots + J_{bN}. \tag{3}$$

Referring to Fig. 2, the two-sensor, two-source problem can be formulated:

$$\begin{matrix} y_1 \\ y_2 \end{matrix} = \begin{bmatrix} h_{11} & h_{21} \\ h_{12} & h_{22} \end{bmatrix} * \begin{matrix} x_1 \\ x_2 \end{matrix} \tag{4}$$

The $y$'s are the sensor data and the $x$'s are the unknown sources.

This means that

$$y_1 = x_1 * h_{11} + x_2 * h_{21} \tag{5}$$
$$y_2 = x_1 * h_{12} + x_2 * h_{22} \tag{6}$$

$h_{11}$ and $h_{22}$ are the direct channels and $h_{12}$ and $h_{21}$ are the cross channels.

For source separation only, we study the two-sensor, two-source problem:

$$\begin{matrix} y_1 \\ y_2 \end{matrix} = \begin{bmatrix} 1 & h_{21} \\ h_{12} & 1 \end{bmatrix} * \begin{matrix} x_1 \\ x_2 \end{matrix}. \tag{7}$$

The diagonal (or direct) channel filters are unity. The estimation of the cross channels and the separation of the sources are the two distinct parts of the problem.

The HJ source separation algorithm and the others tested in this paper are time-recursive (as opposed to batch or frequency domain) algorithms. The Weinstein et. al. [5] algorithm has a batch and two time-recursive implementations. We have tested the time-recursive versions of their algorithm. One version has a stochastic gradient update method (like LMS) and the other has a Newton update method (like RLS).

## 3. THE NEW SOURCE SEPARATION METHOD

Considering a two source, two sensor case using the equations (5,6), we make use of feedback by using the current best estimates of $h_{12}$ and $h_{21}$ to obtain estimates of the sources:

$$\hat{x}_1 = y_1 - \hat{x}_2 * \hat{h}_{21} \tag{8}$$
$$\hat{x}_2 = y_2 - \hat{x}_1 * \hat{h}_{12} \tag{9}$$

The above equations are repeated at least twice at each iteration. These equations make up the separation part of the algorithm. The general separation formulations are:

$$\hat{x}_i = y_i - \sum_{j \neq i} \hat{x}_j * \hat{h}_{ji}, \tag{10}$$

the set of equations (10) are repeated at least $N$ times.

Using the multichannel blind cost function $J_{mb}$ in (3), the cross channels are updated with:

$$\hat{h}_{21} = \hat{h}_{21} + \mu \frac{\partial J_{mb}}{\partial \hat{x}_1} \underline{\hat{x}}_2 \tag{11}$$
$$\hat{h}_{12} = \hat{h}_{12} + \mu \frac{\partial J_{mb}}{\partial \hat{x}_2} \underline{\hat{x}}_1. \tag{12}$$

The general $N$-source, $N$-sensor channel update formulations are:

$$\hat{h}_{ij} = \hat{h}_{ij} + \mu \frac{\partial J_{mb}}{\partial \hat{x}_j} \underline{\hat{x}}_i. \tag{13}$$

Comparing this with the single channel blind equalization update in equation (2), we have

$$\hat{h}_{ij} = \hat{h}_{ij} + \mu(\hat{x}_j - g_j(\hat{x}_j)) \underline{\hat{x}}_i. \tag{14}$$

### Example: Two Uniform Sources

From [6, 8], we know that $O_4^2$ is well suited to work with uniformly distributed data. For separation of two uniform sources we can use $J_{mb} = O_4^2(\hat{x}_1) + O_4^2(\hat{x}_2)$, and the updates are:

$$\hat{h}_{21} = \hat{h}_{21} + \mu(\hat{x}_1 - \frac{E|x_1|^2}{E|x_1|^4} \hat{x}_1^3) \underline{\hat{x}}_2 \tag{15}$$
$$\hat{h}_{12} = \hat{h}_{12} + \mu(\hat{x}_2 - \frac{E|x_2|^2}{E|x_2|^4} \hat{x}_2^3) \underline{\hat{x}}_1. \tag{16}$$

Another example uses a Laplace distributed process as source 1 and a uniformly distributed process as source 2. From [6, 8], we know that $O_1^2$ is optimal to work with Laplace (double-sided exponential) distributed data. For separation of these sources we can use $J_{mb} = O_1^2(\hat{x}_1) + O_4^2(\hat{x}_2)$.

## 4. DECORRELATION SEPARATION USING FEEDBACK

A simple decorrelation operation which ignores the *a priori* known statistical information contained in the nonlinearity of (14) $g_i(\hat{x}_i)$, will also work for the source separation problem. This approach has degraded performance with respect to the "multichannel blind equalization separation" described in Section 3 because it does not make use of the known statistics of the sources. It is described by equation (10) which is repeated at least $N$ times, and

$$\hat{h}_{ij} = \hat{h}_{ij} + \mu \, \hat{x}_j \, \underline{\hat{x}}_i. \tag{17}$$

## 5. EXISTING ADAPTIVE FILTER SOURCE SEPARATION METHODS

### The HJ Source Separation Algorithm [1]

The estimated sources are formed as in (8) and (9), and the cross channels are updated with:

$$\hat{h}_{21} = \hat{h}_{21} + \mu f(\hat{x}_1)g(\hat{\underline{x}}_2) \quad (18)$$

$$\hat{h}_{12} = \hat{h}_{12} + \mu f(\hat{x}_2)g(\hat{\underline{x}}_1). \quad (19)$$

The functions $f(\cdot)$ and $g(\cdot)$ are odd. The algorithm uses the Bussgang property of finite variance signals [9]. Although it is a generalization of the decorrelation algorithm of Section 4, the optimal choices for $f(\cdot)$ and $g(\cdot)$ were not known until their relation to blind equalization theory was found. The simplest method which can be used to find the optimum nonlinearities is presented in [6]. In testing this algorithm using uniform data and $f(x) = x^3$ and $g(x) = x$ (i. e. using only the nonlinear term of (14)), the resulting algorithm had nearly identical performance the new algorithm of Section 3.

### Weinstein Separation Algorithms [5]

These algorithms use a different inverse system architecture. An "adjugate system inverse" is used which is a polynomial matrix inverse without dividing by the determinant. Variables $v_1$ and $v_2$, are defined,

$$\hat{v}_1 = y_1 - \hat{y}_2 * \hat{h}_{21} \quad (20)$$

$$\hat{v}_2 = y_2 - \hat{y}_1 * \hat{h}_{12}, \quad (21)$$

and the solution as derived in [5] is:

$$C_{y_2v_2}h_{21} = c_{y_1v_2} \quad (22)$$

$$C_{y_1v_1}h_{12} = c_{y_2v_1}, \quad (23)$$

which is a set of normal equations. Here $C$ is a cross correlation matrix and $c$ is a cross correlation vector.

Just as the solution of the normal equation has batch and various recursive implementations, so does this set. Following the notation used above, and presenting only the version for real data, we have the following updates: for LMS-type

$$\hat{h}_{21} = \hat{h}_{21} + \mu\ \hat{v}_2\ \underline{\hat{v}}_1 \quad (24)$$

$$\hat{h}_{12} = \hat{h}_{12} + \mu\ \hat{v}_1\ \underline{\hat{v}}_2, \quad (25)$$

and for RLS-type

$$\hat{h}_{21} = \hat{h}_{21} + C^{-1}_{y_1v_1}\ \hat{v}_2\ \underline{\hat{v}}_1 \quad (26)$$

$$\hat{h}_{12} = \hat{h}_{12} + C^{-1}_{y_2v_2}\ \hat{v}_1\ \underline{\hat{v}}_2. \quad (27)$$

It was noted in our simulations that the RLS-type update did not offer speed improvement because the filter update must proceed slowly enough for the separation process to be accurate.

## 6. SOURCE SEPARATION SIMULATIONS

The two source signals used in the simulations shown in Figs. 3 and 4 were uniformly distributed processes with unit variance. The cross-channels were $h_{12} = [.4\ .3\ -.5\ .3\ .3\ .1]$ and $h_{21} = [.5\ -.4\ .2\ -.1\ -.1\ .1]$. An interesting discovery was made in testing these source separation techniques. The closer the direct/cross energy ratio is to unity, the longer the algorithm will take to converge. Direct/cross energy ratio is

$$\frac{\sum_k h_{ii}(k)^2}{\sum_{j\neq i}\sum_k h_{ji}(k)^2}. \quad (28)$$

This problem is akin to the eigenvalue spread problem experienced by the LMS algorithm. For these tests the direct/cross energy ratios were: 3.18db for source 1, and 1.61db for source 2.

For all of the following simulations, we repeated the "bootstrapping" separation equations (8,9) four times. After repeating the equations many times a point of diminishing returns is reached.

For Figs. 3 and 4, the mean squared errors (MSE), $(x_1 - \hat{x}_1)^2$ and $(x_2 - \hat{x}_2)^2$, are used so the effectiveness of the separation can be shown. The MSE from 10 Monte Carlo runs were averaged and plotted. Step sizes were adjusted so that a steady state mean squared error of .02 is attained.

Fig. 3 shows the results of comparing the new algorithm of Section 3 to the decorrelation only method of Section 4. The new algorithm which uses the known statistics of the sources converges much faster.

Fig. 4 shows the results of comparing the new algorithm of Section 3 to the Weinstein et. al LMS-type algorithm which uses the "adjugate inverse system". The new algorithm shows faster convergence, suggesting that the use of the *a priori* known statistics of the sources helps to speed the separation process. In comparing Fig. 4 with the decorrelation result of Fig. 3, we see that the use of feedback is also a help to the algorithm.

### Speech Data

In Fig. 5, we show results using the new source separation method with speech signals as inputs. The cost function used was $J_{mb} = O_1^2(\hat{x}_1) + O_1^2(\hat{x}_2)$. The cross channels were of order 25. The results are quite striking as one can see that the waveforms of the original two speech signals are recovered nearly perfectly.

## 7. CONCLUDING REMARKS

A new source separation method allows one to tailor the algorithm to be optimal for given types of source data. It has far-ranging potential uses, and allows one to use the *a priori* known statistics of the sources to obtain a more powerful algorithm than simple decorrelation. Blind equalization cost functions can be extended to handle multichannel systems.

## 8. REFERENCES


[1] C. Jutten and J. Herault "Blind separation of sources, part 1: an adaptive algorithm based on neuromimetic architecture," *Signal Processing*, vol. 24, 1991, pp. 1-10.

[2] J. Cardoso, "Source separation using higher order moments," *ICASSP 1989*, pp. 2109-2119

[3] J. Cardoso, "Eigen-structure of the fourth-order cumulant tensor with application to the blind source separation problem," *ICASSP 1990*, pp. 2655-2658.

[4] J. Cardoso, A. Belouchrani, and B. Laheld, "A new composite criterion for adaptive and iterative blind source separation," *ICASSP 1994*, vol. 4, pp. 273-276.

[5] E. Weinstein, M. Feder and A. V. Oppenheim, "Multi-channel signal separation by decorrelation," *IEEE Trans. Speech and Audio Processing*, Oct. 1993, vol. 1, no. 4, pp. 405-413

[6] R. H. Lambert, and C. L. Nikias, "Blind equalization cost functions and maximum a posteriori", MILCOM 94, Ft. Monmouth, NJ, pp. 291-295

[7] R. H. Lambert, and C. L. Nikias, "Forward/inverse blind equalization", 28th ASILOMAR Conf. on Signals Systems and Computers, 1994, Pacific Grove, CA

[8] W. C. Gray, "Variable Norm Deconvolution," Ph.D. Dissert. Stanford Univ., 1979.

[9] S. Bellini, "Bussgang techniques for blind equalization," Globecom 1986, pp. 1634-1640.

[10] D. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Trans. on Communications*, vol. 28, 1980, pp. 1867-1875.

[11] Y. Sato, "A method of self-recovering equalization for multilevel amplitude-modulation system," *IEEE Trans. on Communications*, vol. 23, 1975, pp. 679-682.


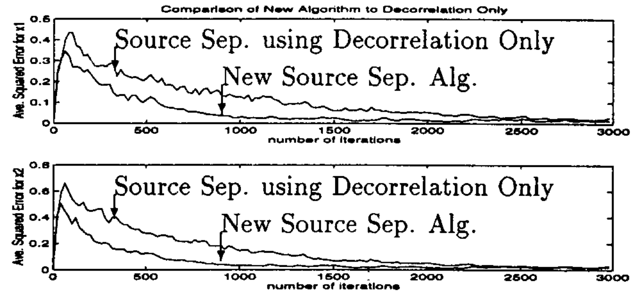


Figure 3: New Algorithm and HJ Algorithm Steady State MSE=.02

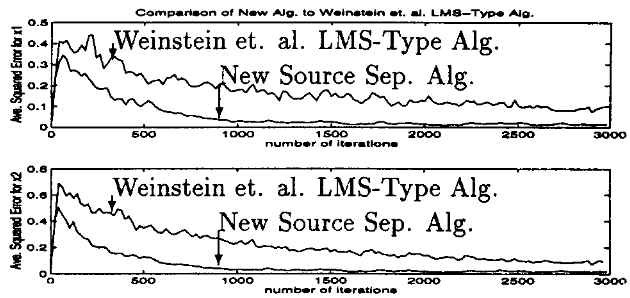


Figure 4: Comparison of new algorithm to the Weinstein et. al. LMS-type Algorithm. The steady state MSE is .02 for both cases.

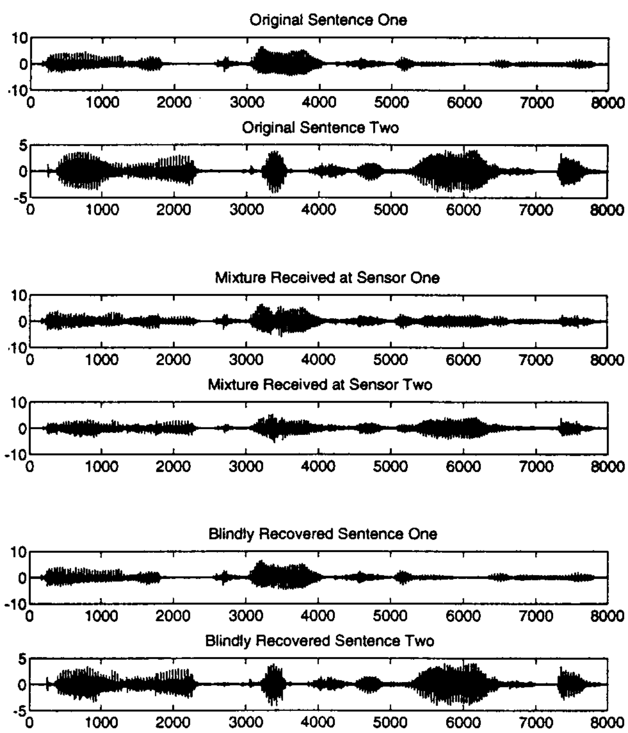


Figure 5: Speech Separation Simulation Using New Algorithm and $O_1^2$