

A BLIND SEPARATION APPROACH FOR MAGNITUDE BOUNDED SOURCES

Alper T. Erdogan

EE Department, Koc University, Istanbul, Turkey
email: alperdogan@ku.edu.tr

ABSTRACT

A novel blind source separation approach for channels with and without memory is introduced. The proposed approach makes use of pre-whitening procedure to convert the original convolutive channel into a lossless and memoryless one. Then a blind subgradient algorithm, which corresponds to an l_∞ norm based criterion, is used for the separation of sources. The proposed separation algorithm exploits the assumed boundedness of the original sources and it has a simple update rule. The typical performance of the algorithm is illustrated through simulation examples where separation is achieved with only small numbers of iterations.

1. INTRODUCTION

Blind Source Separation (BSS) is a subject of interest in signal processing with a variety of applications ranging from channel equalization algorithms for communication links employing multiple antennas to feature vector size reduction in pattern recognition.

As an alternative to the existing BSS approaches (see e.g. [1, 2] and the references therein), we introduce a new algorithm that exploits the assumption that the sources to be recovered are magnitude-bounded. This assumption fits into various practical applications including the equalization of communication channels where the sources are constructed from a bounded finite alphabet.

Our approach assumes initial pre-whitening of the observations, similar to the approach of [2]. Since the original sources are mapped to pre-whitened observations through a memoryless unitary mapping, a Higher Order Statistics (HOS) based approach is needed to recover them [3]. Our approach for separation is based on the following observation, which will be more rigorously stated by Theorem 1 in Section 4 :

Finding the independent components from the whitened observations can be formulated as the geometrical problem of finding an appropriate rotation and/or reflection that minimizes the maximum (real) magnitude component of the transformation output vector over the ensemble.

This work was supported in part by TUBITAK under Career Award Program.

Although the corresponding l_∞ norm based cost function is non-differentiable, we can develop low complexity algorithms based on subgradient optimization, which is the focus of this article.

The organization of the article is as follows: Section 2 provides the BSS setup. In Section 3 the general outline of the approach is provided. Section 4 is the major part where the separation approach that exploits the magnitude boundedness of sources is introduced. The corresponding adaptive algorithm is provided in Section 5. Simulation examples for the proposed algorithm is given in Section 6. Finally, Section 7 is the conclusion.

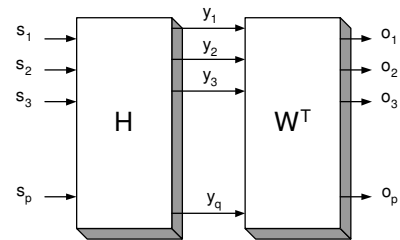


Fig. 1. Blind Source Separation Setup

2. BLIND SOURCE SEPARATION SETUP

The blind source separation setup that we consider throughout the article is shown in Figure 1. In this figure,

- $s_1(k), s_2(k), \dots, s_p(k)$ are source signals. It is assumed that they are all i.i.d with zero mean and unity variance (without loss of generality), and mutually independent of each other. Furthermore, the source signals are considered to be bounded and complex symmetric in the sense that

$$\begin{aligned} \max \Re\{s_l\} &= \max \Im\{s_l\} = -\min \Re\{s_l\} \\ &= -\min \Im\{s_l\} = M. \end{aligned} \quad (1)$$

- The source signals are mixed by a MIMO system with a $q \times p$ transfer matrix $\mathbf{H}(z)$, which has q outputs

denoted by $y_1(k), y_2(k), \dots, y_q(k)$. In z-transform domain, we can write

$$\underbrace{\begin{bmatrix} y_1(z) \\ y_2(z) \\ \vdots \\ y_q(z) \end{bmatrix}}_{\mathbf{y}(z)} = \mathbf{H}(z) \underbrace{\begin{bmatrix} s_1(z) \\ s_2(z) \\ \vdots \\ s_p(z) \end{bmatrix}}_{\mathbf{s}(z)}. \quad (2)$$

We will concentrate on two special cases:

A. Memoryless Channel: In this case $\mathbf{H}(z)$ is assumed to be equivalent to a constant, i.e.,

$$\mathbf{H}(z) = \mathbf{H}_0. \quad (3)$$

The corresponding problem is referred to as the *instantaneous blind source separation*.

B. Convolutional Channel: In this case $\mathbf{H}(z)$ is a polynomial function, i.e., $\mathbf{H}(z) = \sum_{k=0}^L \mathbf{H}_k z^k$, which is not a monomial. We'll assume either of the two conditions below:

1. $\mathbf{H}(z)$ is irreducible [3] i.e.

$$\text{rank}(\mathbf{H}(z)) = p \quad \forall z, \quad (4)$$

and $q > p$. In this case, we can find a proper whitening transformation $\mathbf{W}_{pre}(z)$ to convert $\mathbf{H}(z)$ into a memoryless channel, i.e.,

$$\mathbf{W}_{pre}^T(z) \mathbf{H}(z) = \mathbf{\Phi}, \quad (5)$$

where $\mathbf{\Phi}$ is a $p \times p$ unitary matrix.

2. $\mathbf{H}(z)$ is not irreducible but can be written as

$$\mathbf{H}(z) = \mathbf{H}_I(z) \mathbf{\Phi} \mathbf{\Lambda}(z), \quad (6)$$

where $q > p$, $\mathbf{H}_I(z)$ is a $q \times p$ irreducible transfer matrix, $\mathbf{\Phi}$ is a $p \times p$ unitary matrix and $\mathbf{\Lambda}(z)$ is a $p \times p$ diagonal matrix with monomial entries, i.e.,

$$\mathbf{\Lambda}(z) = \text{diag}([z^{d_1} \quad z^{d_2} \quad \dots \quad z^{d_p}]). \quad (7)$$

In this case we can find [3] a whitening transformation $\mathbf{W}_{pre}(z)$ such that

$$\mathbf{W}_{pre}^T(z) \mathbf{H}(z) = \mathbf{\Phi} \mathbf{\Lambda}(z). \quad (8)$$

Therefore, the resulting output of the whitening transformation would be an instantaneous mixture of unequally delayed versions of the original sources. The sources can be separated in this case, however, relative delays among them would be unknown.

- The purpose of blind source separation is to estimate the source signals (independent components) $s_k(z)$ $k = 1, \dots, p$, from the observations $y_k(z)$, $k = 1, \dots, q$ using a linear system with transfer matrix $\mathbf{W}^T(z)$, i.e.,

$$\mathbf{o}(z) = \mathbf{W}^T(z) \mathbf{y}(z), \quad (9)$$

where $\mathbf{o}(z) = [o_1(z) \quad o_2(z) \quad \dots \quad o_p(z)]^T$ contains the estimates of the original sources. The \mathbf{W} is obtained adaptively from the time samples of \mathbf{y} . No a priori knowledge of \mathbf{H} and no training sequences are assumed.

3. BLIND SOURCE SEPARATION APPROACH

The proposed blind source separation approach consists of two sequential steps:

1. **Whitening Step:** Input vector is whitened based on second order statistics.
2. **Memoryless Source Separation Step:** The whitened vector obtained in the step is transformed to an independent vector using subgradient optimization which implicitly uses HOS.

Therefore, we decompose \mathbf{W} into two operators:

$$\mathbf{W}(z) = \mathbf{W}_{pre}(z) \mathbf{\Theta}, \quad (10)$$

where $\mathbf{W}_{pre}(z)$ is a $q \times p$ whitening transfer matrix such that

$$\underbrace{[x_1(z) \quad x_2(z) \quad \dots \quad x_p(z)]^T}_{\mathbf{x}(z)} = \mathbf{W}_{pre}^T(z) \mathbf{y}(z) \quad (11)$$

is the z-transform of the white vector sequence $\mathbf{x}(k)$ and the components of $\mathbf{x}(k)$ are all unity variance random variables uncorrelated with each other, which we can compactly state as $S_{\mathbf{x}}(z) = \mathbf{I}$.

As a result, we can write

$$\mathbf{o}(z) = \mathbf{W}^T(z) \mathbf{y}(z) \quad (12)$$

$$= \mathbf{\Theta}^T \underbrace{\mathbf{W}_{pre}^T(z) \mathbf{H}(z)}_{\mathbf{C}(z)} \mathbf{s}(z) = \mathbf{\Theta}^T \mathbf{x}(z). \quad (13)$$

Here, we can write $\mathbf{x}(k) = \mathbf{\Phi} \hat{\mathbf{s}}(k)$ where

- Memoryless Channel: $\hat{\mathbf{s}}(k) = \mathbf{s}(k)$
- Convolutional Channel:
 - For Case 1: $\hat{\mathbf{s}}(k) = \mathbf{s}(k)$
 - For Case 2: $\hat{s}_i(k) = s_i(k - d_i)$ for $i = 1, \dots, p$.

The whitening step requires second order statistics and it takes care of the frequency selection and loss effects of the channel and converts the convolutive BSS problem into an instantaneous BSS problem with unitary mapping. Therefore, the next step is to find a unitary matrix Θ which rotates and/or reflects the white vector $\mathbf{x}(k)$ into an independent vector $\mathbf{o}(k)$ and this step requires implicit or explicit use of higher order statistics in general.

4. l_∞ NORM BASED INSTANTANEOUS BLIND SOURCE SEPARATION

In the previous section, the second step of blind source separation problem is formulated as finding a unitary Θ matrix which converts white input vector $\mathbf{x}(k)$ into an independent vector $\mathbf{o}(k)$. The existing methods to solve this problem generally optimize a cost function of the output of Θ under the constraint that Θ is a unitary matrix:

$$\begin{aligned} & \text{maximize } \mathcal{J}(\mathbf{o}(k)) \\ & \text{s.t. } \Theta^H \Theta = I \end{aligned}$$

Here the cost function \mathcal{J} is typically chosen as some measure of nongaussianity such as Multiuser Kurtosis [2] and Negentropy [1] functions.

In our approach, we introduce the following optimization problem (*Problem 1*):

$$\begin{aligned} & \text{minimize } \max_k \|\Re\{\mathbf{o}(k)\}\|_\infty \\ & \text{s.t. } \Theta^H \Theta = I \end{aligned}$$

which is the minimization of the maximum of the infinity norm of the output vector $\mathbf{o}(k)$ over time. The following theorem, whose proof is given in [4] shows that the global minimizers of the above optimization setting are the desired solutions for the source separation problem:

Theorem 1: Given that each i.i.d. $s_l(k)$ satisfies the input condition in (1), and $\mathbf{x} = \Phi \mathbf{s}$ where Φ is a unitary matrix. Let \mathbf{G} be defined as the overall mapping between \mathbf{s} and \mathbf{o} , then Θ_{opt} is a global minima of the *Problem 1* if and only if the corresponding \mathbf{G}_{opt} has the form

$$\mathbf{G}_{opt} = \Theta_{opt}^T \Phi = \mathbf{D}\mathbf{E}, \quad (14)$$

where \mathbf{E} is a permutation matrix and \mathbf{D} is a diagonal matrix of the form $\mathbf{D} = \text{diag}([e^{j\frac{\pi}{2}k_1} \quad e^{j\frac{\pi}{2}k_2} \quad \dots \quad e^{j\frac{\pi}{2}k_p}])$, where k_1, k_2, \dots, k_p are integers.

5. THE ADAPTIVE ALGORITHM

In the previous section, we introduced (*Problem 1*) whose global minimizers are the desired separators. In this section we provide an iterative adaptive algorithm to obtain a solution for this problem.

The gradient search based approach, which is used in various BSS approaches (e.g., [2]) as well as other adaptive signal processing methods is not directly applicable to *Problem 1* as its cost function is not differentiable. Fortunately, it is a convex cost function for which we can use subgradient methods (see [5] and the references therein).

In order to obtain a practical algorithm to minimize the cost function of *Problem 1*, the evaluation of the maximum value of the infinity-norm of $\Re\{\mathbf{o}(k)\}$ should be limited to a finite window of time values $k \in \{0, 1, \dots, \Omega - 1\}$. In this case, we can rewrite the corresponding optimization problem (*Problem 2*) as

$$\begin{aligned} & \text{minimize } f(\Theta) = \max_{k \in \{0, 1, \dots, \Omega - 1\}} \|\Re\{\mathbf{o}(k)\}\|_\infty \\ & \text{s.t. } \Theta^H \Theta = I. \end{aligned}$$

If we define $\mathbf{X} = [\mathbf{x}(0) \quad \mathbf{x}(1) \quad \dots \quad \mathbf{x}(\Omega - 1)]$, as the matrix of input values in the window of interest, then for a given Θ the corresponding outputs can be placed in a matrix \mathbf{O} :

$$\begin{aligned} \mathbf{O} &= [\mathbf{o}(0) \quad \mathbf{o}(1) \quad \dots \quad \mathbf{o}(\Omega - 1)] \quad (15) \\ &= \Theta^T \mathbf{X}. \quad (16) \end{aligned}$$

Based on these definitions, we can write the subdifferential set corresponding to the cost function of *Problem 2* as [4]

$$\partial f(\Theta) = \text{Co}\left\{ \frac{\Re\{\mathbf{O}_{m,n}\}}{|\Re\{\mathbf{O}_{m,n}\}|} \bar{\mathbf{X}}_{:,n} \mathbf{e}_m^T \mid (m,n) \in \mathcal{I} \right\}, \quad (17)$$

where Co is the convex hull operation, $\mathbf{O}_{m,n}$ is the element of the matrix \mathbf{O} located at the m^{th} row and n^{th} column (which is $o_m(n)$), $\bar{\mathbf{X}}_{:,n}$ is the n^{th} column of \mathbf{X} (which is $\mathbf{x}(n)$) with its elements complex conjugated and \mathcal{I} is the set of index pairs for which the maximum real magnitude is achieved in matrix \mathbf{O} , i.e.,

$$\mathcal{I} = \{(m,n) \mid |\Re\{\mathbf{O}_{m,n}\}| = f(\Theta)\}. \quad (18)$$

Based on the subdifferential set in (17), we can write a subgradient based update rule for the blind source separation as

$$\begin{aligned} \underline{\Theta}^{(i+1)} &= \underline{\Theta}^{(i)} - \mu^{(i)} \frac{\Re\{\mathbf{O}_{m^{(i)},n^{(i)}}^{(i)}\}}{|\Re\{\mathbf{O}_{m^{(i)},n^{(i)}}^{(i)}\}|} \bar{\mathbf{X}}_{:,n^{(i)}} \mathbf{e}_{m^{(i)}}^T \quad (19) \\ \underline{\Theta}^{(i+1)} &= \mathcal{P}_U\{\underline{\Theta}^{(i+1)}\}, \quad (20) \end{aligned}$$

where

- $\Theta^{(i)}$ is the value of Θ at the i^{th} iteration,
- $\mathbf{O}^{(i)}$ is the output matrix calculated based on $\Theta^{(i)}$,

- $(m^{(i)}, n^{(i)})$ is the index for a maximum real component magnitude entry of $\mathbf{O}^{(i)}$,
- $\mu^{(i)}$ is the step size at the i^{th} iteration. We use the relaxation rule for fast convergence. (see [5, 4] for the details).
- $\underline{\Theta}^{(i+1)}$ is the unprojected version of the updated Θ ,
- \mathcal{P}_U is the projection operator to the unitary matrix set where we use the minimum-distance projection operator to the set of unitary matrices [4] which is defined as

$$\begin{aligned}\Theta^{(i+1)} &= \mathcal{P}_U\{\underline{\Theta}^{(i+1)}\} \\ &= \mathcal{P}_U\{\mathbf{U}^{(i+1)}\Sigma^{(i+1)}\mathbf{V}^{(i+1)H}\} \\ &= \mathbf{U}^{(i+1)}\mathbf{V}^{(i+1)H},\end{aligned}$$

where $\mathbf{U}^{(i+1)}\Sigma^{(i+1)}\mathbf{V}^{(i+1)H}$ is the SVD of $\underline{\Theta}^{(i+1)}$. These matrices can be conveniently computed using a Gram-Schmidt based algorithm [4] which exploits the special update structure in (19).

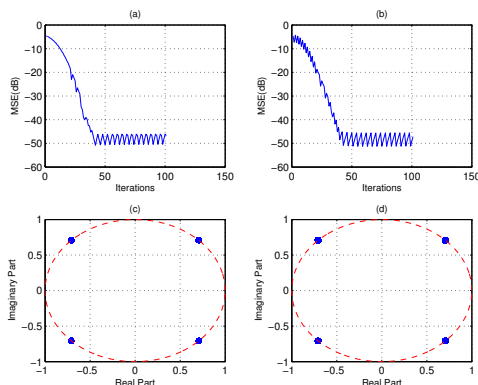


Fig. 2. MSE curves and eye diagrams for the first example.

6. EXAMPLES

As the first example, we consider a 3×2 convolutive channel of order 11. Both sources use 4-QAM constellation. The entries of the impulse response for this channel are randomly generated from a Gaussian distribution with zero mean and unity variance. We chose the whitening filter length as 24 and the window size Ω as 800. In Figure 2, typical convergence diagrams (a)-(b) (after whitening) and the corresponding eye diagrams after convergence (c)-(d) for both sources are shown. It can be seen from these figures that the algorithm achieves almost perfect separation at only 50 iterations, which is a typical behavior on all random 3×2 transfer matrices with independent Gaussian entries we generated. Furthermore, since the phase ambiguities of the algorithm are factors of $\pi/2$, the eye diagrams after convergence appears almost exactly as 4-QAM constellations.

As the second example, we consider a 3^{rd} order 4×2 channel whose output is corrupted by AWGN with a corresponding SNR level of 40dB. Typical convergence curves for a random channel are shown in Figure 3 (a)-(b). We should note that it is possible to have both higher and lower SNR levels for different channel choices depending on the noise enhancement levels of the whitening processes corresponding to those channels [2].

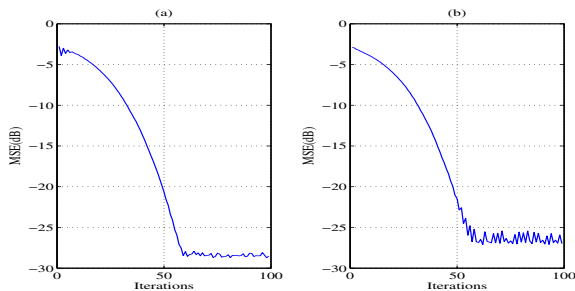


Fig. 3. MSE curves for the second example.

7. CONCLUSION

A novel iterative BSS approach based on subgradient optimization is introduced. The memory of the channel is handled by whitening and the remaining source separation, which requires the explicit or implicit use of higher order statistics, is achieved by optimizing an l_∞ norm based cost function. Since this cost function is non-differentiable, a subgradient based approach is used. The resulting algorithm, which also employs minimum distance mapping to unitary set, has both low complexity and fast convergence.

8. REFERENCES

- [1] Aapo Hyvärinen, Juha Karhunen, and Erkki Oja, *Independent Component Analysis*, John Wiley and Sons Inc., 2001.
- [2] Constantinos Papadias, “Unsupervised receiver processing techniques for linear space-time equalization of wideband multiple input / multiple output channels,” *IEEE Trans. on SP*, vol. 52, pp. 472–482, Feb. 2004.
- [3] Yuijro Inouye and Ruey-Wen Liu, “A system-theoretic foundation for blind equalization of an FIR MIMO channel system,” *IEEE Trans. on Signal Processing*, vol. 49, no. 4, pp. 425–436, April 2002.
- [4] Alper T. Erdogan, “A simple geometric blind source separation method for bounded magnitude sources,” *submitted to IEEE TSP (draft available at http://aspc.ku.edu.tr/Research.htm)*, 2004.
- [5] Alper T. Erdogan and Can Kizilkale, “Fast and low complexity blind equalization via subgradient projections,” *IEEE Trans. on Signal Processing*, to appear.