A BRUTE-FORCE ANALYTICAL FORMULATION OF THE INDEPENDENT COMPONENTS ANALYSIS SOLUTION

Deniz Erdogmus, Anant Hegde, Kenneth E. Hild II, M. Can Ozturk, Jose C. Principe

CNEL, Electrical & Computer Eng. Dept., University of Florida, Gainesville, FL 32611, USA

ABSTRACT

Many algorithms based on information theoretic measures and/or temporal statistics of the signals have been proposed for ICA in the literature. There have also been analytical solutions suggested based on predictive modeling of the signals. In this paper, we show that finding an analytical solution for the ICA problem through solving a system of nonlinear equations is possible. We demonstrate that this solution is robust to decreasing sample size and measurement SNR. Nevertheless, finding *the root* of the nonlinear function proves to be a challenge. Besides the analytical solution approach, we try finding the solution using a least squares approach with the derived analytical equations. Monte Carlo simulations using the least squares approach are performed to investigate the effect of sample size and measurement noise on the performance.

1. INTRODUCTION

Independent components analysis (ICA) has become a useful tool in engineering and basic scientific research. There are many successful algorithms in the literature that find the independent components of a signal set. These algorithms mostly exploit the independence assumption for the sources through information theoretic measures, higher order statistics of the signals, and the temporal structures of the signals through second order statistical measures like cross-correlation at multiple lags. Among the information theoretic approaches we can list Bell and Sejnowski's Infomax [1], Comon's minimum mutual information method [2], Yang and Amari's information theoretic approaches [3], and Mermaid [4] by Hild *et al*. On the higher order statistics front, JADE by Cardoso [5], Pham's decorrelation-of-outputs approach [6], Hyvarinen's celebrated FastICA [7], Karhunen and Oja's nonlinear PCA approaches [8] are among the significant innovations. For nonstationary sources, the time-varying cross-correlation of the signal can be exploited [9,10,11]. An original approach, first proposed by Zibulevsky, is to determine a transformation such that the representations of the sources in the transform space become sparse [12]. Then, the determination of the mixing matrix becomes extremely easy as the observation vectors start lining up along the matrix columns with increasing sparsity factor [13]. None of these approaches, however, provide an analytical expression for the mixing matrix or for its inverse, without resorting to some indirect criterion of optimality. There have even been approaches that utilize the predictive modeling of the signals to determine an analytical solution for ICA [14]. Nevertheless, a generic approach that targets the determination of the mixing matrix without restrictive assumptions has not yet been proposed.

It turns out that the answer to this problem lies in the simplest of all approaches, which is the subject of this paper. We called this approach *brute-force ICA*, because it relies heavily on forcing out a system of equations from the data by which the unknowns can be determined numerically, and perhaps analytically. In fact, while we were working on the formulation of the presented methodology, a paper has appeared that exploited similar ideas to find an analytical solution for the blind equalization problem [15]. In this paper, however, we will not focus on the analytical solution. We will rather concentrate on the change in the performance of the solution obtained through this brute-force approach when the number of training samples and the measurement signal-to-noise-ratio (SNR) are varied.

2. PROBLEM DESCRIPTION AND SOLUTION

Suppose that there exist n mutually independent signals (in the source vector $\mathbf{s}$) that are mixed by an unknown matrix (called the mixing matrix) $\mathbf{H}$ to form the observation vector $\mathbf{x}$ according to $\mathbf{x} = \mathbf{Hs}$. The ij^{th} entry of the mixing matrix is denoted by h_{ij}, and it is the scale factor that multiplies source $\mathbf{s}_j$ in observation $\mathbf{x}_i$. For simplicity, we consider the square mixture case, where the number of observations is also n. In ICA, the task is to determine the independent source signals or the mixing matrix having only the observed vector samples $\mathbf{x}(k)$, where k is the sample (time) index. Since neither $\mathbf{H}$ nor the source signals are available, the solution requires some assumptions regarding the statistics of the source signals.

Here, we will resort to the commonly used independence assumption. In addition, we will assume without loss of generality that the sources are zero-mean and unit-variance. Subtracting the mean of the observation vector from the observed samples satisfies the zero-mean assumption. On the other hand, the unit-variance assumption is acceptable since the source amplitude scale factors cannot be distinguished from the mixing matrix entry scale factors. Formally, these two assumptions are expressed as $E[\mathbf{s}_i]=0$ and $E[\mathbf{s}_i^2]=1$. In addition, due to the source independence assumptions, we obtain the following equalities for integer α.

$$\begin{aligned} &E[\mathbf{s}_i^{\alpha}\mathbf{s}_j]=E[\mathbf{s}_i^{\alpha}]E[\mathbf{s}_j]=0 \quad j\neq i \\ &E[\mathbf{s}_i^{\alpha}\mathbf{s}_j^2]=E[\mathbf{s}_i^{\alpha}]E[\mathbf{s}_j^2]=E[\mathbf{s}_i^{\alpha}] \quad j\neq i \end{aligned} \tag{1}$$

In the last identity of (1), $E[\mathbf{s}_i^{\alpha}]$ is either 0 (for $\alpha=1$), 1 (for $\alpha=2$), or unknown (for $\alpha>2$). These two identities will prove extremely useful in determining the system of equations for which we aim. Now consider the second order joint moments of the observed signals.

$$\begin{aligned} E[\mathbf{x}_i^2] &= E\left[\left(\sum_{l=1}^{n} h_{il}\mathbf{s}_l\right)^2\right] \\ &= \sum_{l=1}^{n} h_{il}^2 E[\mathbf{s}_l^2] + 2\sum_{l=1}^{n}\sum_{m=1}^{n} h_{il}h_{im}E[\mathbf{s}_l\mathbf{s}_m] \\ &= \sum_{l=1}^{n} h_{il}^2 \\ E[\mathbf{x}_i\mathbf{x}_j] &= E\left[\left(\sum_{l=1}^{n} h_{il}\mathbf{s}_l\right)\left(\sum_{m=1}^{n} h_{jm}\mathbf{s}_m\right)\right] \\ &= \sum_{l=1}^{n}\sum_{m=1}^{n} h_{il}h_{jm}E[\mathbf{s}_l\mathbf{s}_m] = \sum_{l=1}^{n} h_{il}h_{jl} \end{aligned} \tag{2}$$

Similarly, we can derive expressions for the fourth order joint moments of the observed signals. Using the simplifications pointed out in (1), these equations become

$$\begin{aligned} E[\mathbf{x}_i^4] &= \sum_{l=1}^{n} h_{il}^4 E[\mathbf{s}_l^4] + 6\sum_{l=1}^{n-1}\sum_{k=l+1}^{n} h_{il}^2 h_{ik}^2 \\ E[\mathbf{x}_i^3\mathbf{x}_j] &= \sum_{l=1}^{n} h_{il}^3 h_{jl} E[\mathbf{s}_l^4] + \sum_{l=1}^{n-1}\sum_{k=l+1}^{n} h_{il}^2 h_{ik} h_{jk} \\ E[\mathbf{x}_i^2\mathbf{x}_j^2] &= \sum_{l=1}^{n} h_{il}^2 h_{jl}^2 E[\mathbf{s}_l^4] + \sum_{l=1}^{n}\sum_{\substack{k=1\\k\neq l}}^{n} h_{il}^2 h_{jk}^2 \\ &\quad + 4\sum_{l=1}^{n}\sum_{\substack{k=1\\k\neq l}}^{n} h_{il}h_{ik}h_{jl}h_{jk} \end{aligned} \tag{3}$$

Notice in (3) that the fourth order moments of the independent sources appear as additional unknowns in the equations. With these n new unknowns and the original n^2 **H** entries, the total number of unknowns when the second and fourth order joint moments of the observed signals are considered becomes n^2+n. In order to determine these unknowns, we have extracted $n+n(n-1)/2$ equations from the second order moments and $n+n(n-1)+n(n-1)/2$ equations from the fourth order moments, which amount to $2n^2$. Clearly, we are interested in the integer values of $n\geq 2$. Therefore, for these cases of interest we always have more equations than unknowns when the second and fourth order joint moments are utilized ($2n^2> n^2+n$).

In the above discussion, we have not considered the third order joint moments, because doing so would introduce equations that contained the third order moments of the independent sources, i.e., $E[\mathbf{s}_i^3]$. For symmetric source distributions this moment will become zero, which reduces the number of independent equations. Therefore, using these moments is not recommended.

Now that we have determined the expressions for the joint moments of the observations in terms of the mixing matrix entries and the fourth order moments of the sources, the problem reduces to finding the root of a system of nonlinear equations of the form **f**(**X**)=**C**, where **X** is the vector of unknowns consisting of the entries h_{ij} and the source moments $E[\mathbf{s}_i^4]$. The constant vector **C**, on the other hand, consists of the sample estimates of the second and fourth order moments of the observations, i.e., $E[\mathbf{x}_i^{\alpha}\mathbf{x}_j^{\beta}]$, $\alpha,\beta=0,1,2,3,4$ and $\alpha+\beta=4$. In practice, it is recommended that all $2n^2$ equations be utilized since the solution found by an overdetermined system of equations is expected to have smaller finite-sample variance and more robustness to noise compared to any solution that will be obtained using a subset (with size n^2+n or more) of these equations. The importance of the equations could also be weighted based on the estimated variance of the entries of **C** due to the finite sample size. We will not however, be concerned with these issues in this paper.

Clearly, the performance of the solution obtained using this approach will be independent of the sign of the kurtosis of the source signals. In fact, since the fourth order moments of the sources are among the unknowns to be determined, the solution of the algorithm can be used to determine the values of the kurtosis of the sources.

It is possible to write out the equations for the higher order joint moments of the observed signals, thus get additional or alternative equations. However, odd moments are not desirable due to the same reason stated earlier for the third order moments. Higher order even moments, on the other hand, are less desirable than the second and fourth order moments, simply because as the moment order increases, sample estimates become more vulnerable to outliers and require more samples for accurate estimation.

3. THE SIMPLEST SPECIAL CASE

In order to demonstrate the level of complexity involved, we will explicitly present the equations for a 2x2 mixture situation in this section. When n=2, the total number of equations given by (2) and (3) is eight. These are given in (4). Since these eight equations in six unknowns do not conflict with each other, we can select an arbitrary subset of six equations.

$$\begin{aligned}
E[\mathbf{x}_1^2] &= h_{11}^2 + h_{12}^2 \\
E[\mathbf{x}_2^2] &= h_{21}^2 + h_{22}^2 \\
E[\mathbf{x}_1\mathbf{x}_2] &= h_{11}h_{21} + h_{12}h_{22} \\
E[\mathbf{x}_1^4] &= h_{11}^4 E[\mathbf{s}_1^4] + h_{12}^4 E[\mathbf{s}_2^4] + 6h_{11}^2 h_{12}^2 \\
E[\mathbf{x}_2^4] &= h_{21}^4 E[\mathbf{s}_1^4] + h_{22}^4 E[\mathbf{s}_2^4] + 6h_{21}^2 h_{22}^2 \\
E[\mathbf{x}_1^3\mathbf{x}_2] &= h_{11}^3 h_{21} E[\mathbf{s}_1^4] + h_{12}^3 h_{22} E[\mathbf{s}_2^4] \\
&\quad + 3h_{11}^2 h_{12} h_{22} \\
E[\mathbf{x}_1\mathbf{x}_2^3] &= h_{21}^3 h_{11} E[\mathbf{s}_1^4] + h_{22}^3 h_{12} E[\mathbf{s}_2^4] \\
&\quad + 3h_{21}^2 h_{22} h_{12} \\
E[\mathbf{x}_1^2\mathbf{x}_2^2] &= h_{11}^2 h_{21}^2 E[\mathbf{s}_1^4] + h_{12}^2 h_{22}^2 E[\mathbf{s}_2^4] \\
&\quad + h_{11}^2 h_{22}^2 + 4h_{11}h_{12}h_{21}h_{22}
\end{aligned} \tag{4}$$

Consider the selection of the first five equations and the last equation. Using the fourth and fifth identities of (4), it is possible to determine $E[\mathbf{s}_1^4]$ and $E[\mathbf{s}_2^4]$ in terms of the mixing matrix entries. In addition, using the first two identities, we can express the diagonal entries of the matrix in terms of the off-diagonal entries. After all the substitutions and taking some combinations of different powers of the third and eighth identities, we finally obtain the following two equations that need to be solved simultaneously for the diagonal entries of the mixing matrix.

$$\begin{aligned}
0 &= h_{11}^4[2h_{22}^4 - 2c_2 h_{22}^2 - c_2] \\
&\quad + h_{11}^2[-2c_1 h_{22}^4 + 2c_1(c_2+1)h_{22}^2] \\
&\quad + [2c_3^2 - c_3^4 - c_1^2 h_{22}^2] \\
0 &= h_{11}^2(c_2 - h_{22}^2)\begin{bmatrix} h_{22}^4(c_4 - 6h_{11}^2(c_1 - h_{11}^2)) \\ (c_1 - h_{11}^2)^2(6h_{22}^2(c_2 - h_{22}^2) - c_5) \end{bmatrix} \\
&\quad + h_{22}^2(c_1 - h_{11}^2)\begin{bmatrix} h_{11}^4(c_5 - 6h_{22}^2(c_2 - h_{22}^2)) \\ (c_2 - h_{22}^2)^2(6h_{11}^2(c_1 - h_{11}^2) - c_4) \end{bmatrix} \\
&\quad + \left[h_{11}^4 h_{22}^4 - (c_1 - h_{11}^2)(c_2 - h_{22}^2)\right]\cdot \\
&\quad \left[h_{11}^2 h_{22}^2 - 2h_{11}^2(c_2 - h_{22}^2) - 2h_{22}^2(c_1 - h_{11}^2)\right] \\
&\quad - (c_8 - 2c_3^2)
\end{aligned} \tag{5}$$

In (5), the constants c_i denote the expectations on the left hand side of (4) and are determined by the data. The first equation in (5) is quadratic in h_{11}^2, if h_{22}^2 is considered to be a constant. From this, we obtain two solutions for h_{11}^2 in terms of h_{22}^2 corresponding to different permutations of the sources. Substituting one of these solutions for h_{11}^2 in the second identity in (5) reveals a complicated equation for h_{22}^2. Although we attempted to solve this equation analytically, we were unsuccessful. However, it is possible to search for this root to determine h_{22}^2. Then, h_{11}^2 is determined by the first identity in (5). Moreover, using the first two identities in (4), it is possible to calculate h_{12}^2 and h_{21}^2. The actual matrix entries can then be determined by taking the square root of all these values. However, care must be taken in selecting the signs of these square roots. These signs must be consistent with all (or only the selected six) equations in (4). To do this, one can arbitrarily choose the signs of the diagonal elements of the solution. The signs of the off-diagonal entries can then be determined using, for example, the third identity in (4), i.e. the c_3 equation.

4. LEAST-SQUARES APPROACH

In the noisy finite-sample case, the solution obtained by simultaneously solving any subset of the equations in (4) (in the general case (2) and (3)) might result in suboptimal mixing matrix estimates in the least square sense. In order to address this issue, it is possible to solve for the matrix entries, as well as the source fourth order moments, by minimizing the following least squares criterion.

$$J(\mathbf{X}) = (\mathbf{f}(\mathbf{X}) - \mathbf{c})^T \mathbf{G}(\mathbf{f}(\mathbf{X}) - \mathbf{c}) \tag{6}$$

where $\mathbf{c}$ is a vector consisting of the observation joint moments appearing on the left side of (4), $\mathbf{X}$ is the vector of unknown parameters consisting of the mixing matrix entries and the source fourth order moments, $\mathbf{f}(\mathbf{X})$ is the nonlinear functions appearing on the right hand side of (4), and finally $\mathbf{G}$ is a positive definite weighting matrix that could be used to weight the importance of each equation in the solution. In order to be strictly consistent with the least squares theory, this weighting matrix could be selected as a diagonal matrix consistent with the finite-sample estimation variances of the entries of $\mathbf{c}$. On the other hand, computing these estimation variances is not an easy task, therefore, one might resort to the simple choice of an identity weighting matrix, i.e. $\mathbf{G}$=$\mathbf{I}$.

The minimization can be carried out using any standard optimization algorithm. For example, if steepest descent is utilized, the update algorithm for $\mathbf{X}$ becomes

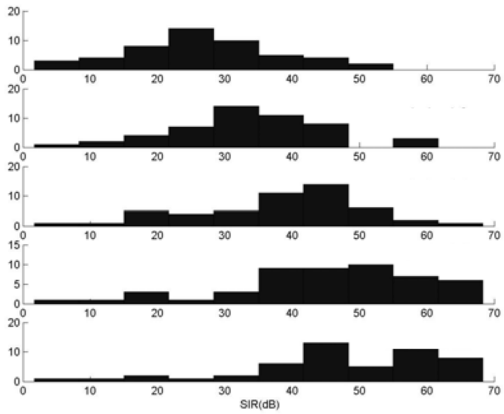

Figure 1. SIR (dB) histograms of 50 Monte Carlo simulations presented from top to bottom for each of the sample sizes 10^2, 10^3, 10^4, 10^5, 10^6.

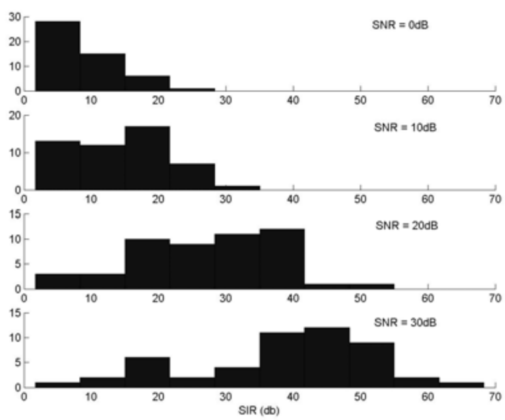

Figure 2. SIR (dB) histograms of 50 Monte Carlo simulations presented from top to bottom for each of the SNR levels 0dB, 10dB, 20dB, 30dB.

$$\mathbf{X}_{k+1} = \mathbf{X}_k - 2\eta \left(\frac{\partial \mathbf{f}(\mathbf{X})}{\partial \mathbf{X}} \right)^T_{\mathbf{X}=\mathbf{X}_k} \mathbf{G} \cdot (\mathbf{f}(\mathbf{X}) - \mathbf{c}) \qquad (7)$$

Since the performance surface given in (6) is highly nonlinear, there will be local minima that might trap the algorithm. In order to reach one of the global optima (there are multiple global optima corresponding to different permutations and signs of the separated sources), the solution offered by (5) could be used as an initial estimate. The least squares procedure then refines the mixing matrix estimate to find the MSE-optimal solution.

5. SIMULATION RESULTS

In order to study the performance of the brute-force ICA solution, we have designed two Monte Carlo experiments. In one of these experiments, we evaluate the performance of the algorithm versus the number of available data samples. In the second experiment, we investigate the robustness of the solution to measurement noise by varying the SNR of the observed signals. For the simulations, we use the 2x2 case. For every Monte Carlo run, each entry of the mixing matrix is selected randomly from a uniform distribution in [-1,1]. As the performance measure, we utilize the signal-to-interference ratio (SIR) defined below. Denoting the actual mixing matrix with **H** and the inverse of the estimated mixing matrix with $\hat{\mathbf{H}}^{-1}$, the overall mixing matrix after separation becomes $\mathbf{Q} = \hat{\mathbf{H}}^{-1}\mathbf{H}$. Letting $\mathbf{q}_i$ denote the i^{th} row of the overall matrix **Q**, the SIR (dB) is defined as

$$SIR = \frac{1}{n} \sum_{i=1}^{n} 10 \log_{10} \frac{\max_j \mathbf{Q}_{ij}^2}{\mathbf{q}_i \mathbf{q}_i^T - \max_j \mathbf{Q}_{ij}^2} \qquad (8)$$

For each of the sample sizes 10^2, 10^3, 10^4, 10^5, and 10^6, we perform 50 Monte Carlo simulations using zero-mean, unit-variance sources and starting from randomly selected matrix estimates. The two source distributions were selected to be uniform and Gaussian. However, notice that the formalism presented above does not impose any restrictions on the source distributions other than independence, nor does its performance critically depend on these distributions.

The results of the first set of Monte Carlo simulations are presented in Fig. 1. For each sample size, the histogram of the final SIR values is shown in a subplot. The subplots are ordered from top to bottom for ascending sample size. We clearly observe the expected increase in performance as the number of samples increase from one hundred to one million.

In the second set of Monte Carlo simulations, we vary the average SNR at the observed signals from 0dB to 30dB in steps of 10dB. Similarly, we perform 50 Monte Carlo simulations using randomly selected matrix estimates. The sample size is kept fixed at 10^4 for all runs. Once again, the source distributions are uniform and Gaussian. The SIR measure is *not* modified to account for the noise in the separated signals, since neither the algorithm nor the demixing structure is tuned to reduce noise. Nevertheless, the current measure gives an idea of how much interference is coming from the *unwanted* source signals in each output channel.

The results of these Monte Carlo simulations are presented in Fig. 2. For each SNR level, the histogram of the final SIR values is shown in a subplot. Again, the subplots are ordered from top to bottom for ascending SNR. As expected, we observe an increase in performance as the noise power drops from being equal to the signal power to values negligible compared to the signal power.

In both sets of simulations, the steepest descent algorithm sometimes resulted in low-quality solutions exhibiting SIR levels less than or around 10dB. These

Mean SIR	$N = 10^2$	$N = 10^3$	$N = 10^4$
Fast ICA	16.7	24.2	30.6
Brute-force	26.9	33.8	38.4

Table 1. Performance of Fast ICA and the presented brute-force ICA approach for the 2x2 case with uniform and Gaussian sources; mean SIR (dB) over 50 Monte Carlo simulations is used as the figure of merit.

separation levels correspond to local minima, therefore they represent suboptimal solutions. The problem of local minima could be avoided by starting adaptation using a standard ICA algorithm. After convergence of the standard algorithm, the solution could be used as the initial condition for the proposed algorithm in order to fine-tune the matrix estimate.

In order to demonstrate this, we present the average performance of Fast ICA [7], a benchmark algorithm, in the same experimental setup (with a 2x2 mixture using one uniform and one Gaussian distributed source). Fast ICA is known to be very successful in the described experimental setting. For each of 100, 1000, and 10000 sample sizes, we have performed 50 Monte Carlo simulations with both of these algorithms. The results of brute-force ICA for the same situation were already presented in Fig. 1 in the form of histograms. The average SIR values obtained by the solutions given by these two algorithms for the three training data sizes are listed in Table 1. Clearly, the presented brute-force ICA approach is able to achieve a much better separation solution with the given data. Thus, it is possible, for example, to use Fast ICA to obtain a sufficiently accurate initial condition for brute-force ICA. The latter approach can then be implemented using the least-squares methodology described above to obtain a more accurate solution.

6. CONCLUSIONS

In the literature, numerous ICA algorithms are proposed, yet the simplest approach (extracting the equations for the solution from the topology using a brute-force approach on the independence assumption) had not been tried. In this paper, we aimed to demonstrate that it is possible to determine a nonlinear system of equations from which the mixing matrix in the ICA problem can be determined. In this formulation, the propagation of second and fourth order moments through the mixing matrix are exploited. As a consequence, the fourth order moments of the source signals appeared in these equations as additional unknowns. This way, the determination of a successful solution has been made independent of the sign of the kurtosis of the source signals.

We have attempted to find the analytical solution for the simplest 2x2 situation. We were able to deduce a single nonlinear equation in only one variable (one of the diagonal entries of the mixing matrix). However, due to the complexity of this final equation, we could not determine the analytical root, which would yield the expression for this matrix entry. Nevertheless, in computer experiments whose results were not presented in this paper, single dimensional numerical zero-finding methods (the standard *fzero* function of Matlab®) were able to determine this root very accurately. Once this value is determined, the other matrix entries and the source fourth order moments could be solved analytically using the presented system of equations.

For general practical purposes, the analytical solution might not be feasible due to increasing complexity with data dimension. In those situations, a least squares approach can be employed. In this paper, we have presented the basics of such a least squares approach and we have presented Monte Carlo simulation results for this approach. In these simulations, we have studied the effect of sample size and measurement noise level on the performance of the algorithm. The results showed that the proposed criterion is able to yield high accuracy separation results for sample sizes as low as 100.

As a future line of research, we will investigate the system of equations that will arise from utilizing the time-delayed correlations of the observed signals. These might lead to simpler equations, thus to an analytical expression for the ICA solution.

Acknowledgments: This work is supported by NSF grant ECS-9900394.

7. REFERENCES

[1] A. Bell, T. Sejnowski, "An Information-Maximization Approach to Blind Separation and Blind Deconvolution," *Neural Computation*, vol. 7, pp. 1129-1159, 1995.

[2] P. Comon, "Independent Component Analysis, a New Concept?" *Signal Processing*, vol. 36, no. 3, pp. 287-314, 1994.

[3] H.H. Yang, S.I. Amari, "Adaptive Online Learning Algorithms for Blind Separation: Maximum Entropy and Minimum Mutual Information," *Neural Computation*, vol. 9, pp. 1457-1482, 1997.

[4] K.E. Hild II, D. Erdogmus, J.C. Principe, "Blind Source Separation Using Renyi's Mutual Information," IEEE Signal Processing Letters, vol. 8, no. 6, pp. 174-176, 2001.

[5] J.F. Cardoso, A. Souloumiac, "Blind Beamforming for Non-Gaussian Signals," *IEE Proc. F Radar and Signal Processing*, vol. 140, no. 6, pp. 362-370, 1993.

[6] D.T. Pham, “Blind Separation of Instantaneous Mixture of Sources via the Gaussian Mutual Information Criterion,” *Signal Processing*, vol. 81, pp. 855-870, 2001.

[7] A. Hyvarinen, “Fast and Robust Fixed-Point Algorithms for Independent Component Analysis,” *IEEE Transactions on Neural Networks*, vol. 10, pp. 626-634, 1999.

[8] A. Hyvarinen, J. Karhunen, E. Oja, *Independent Component Analysis*, Wiley, New York, 2001.

[9] A. Ziehe, K.R. Muller, “TDSEP - An Efficient Algorithm for Blind Separation Using Time Structure,” *Proceedings of ICANN’98*, Skovde, Sweden, pp. 675-680, 1998.

[10] L. Parra, C. Spence, “Convolutive Blind Separation of Non-Stationary Sources,” *IEEE Transactions on Speech and Audio Processing*, vol. 8, no. 3, pp. 320-327, 2000.

[11] H.C. Wu, J.C. Principe, “Generalized Anti-Hebbian Learning for Source Separation,” *Proceedings of ICASSP’99*, Phoenix, Arizona, vol. 2, pp. 1073-1076, 1999.

[12] M. Zibulevsky, B. Pearlmutter, P. Bofill, P. Kisilev, “Blind Source Separation by Sparse Decomposition in a Signal Dictionary,” in *Independent Components Analysis: Principles and Practice*, Cambridge University Press, Cambridge, England, 2000.

[13] L. Vielva, D. Erdogmus, J.C. Principe, “Underdetermined Blind Source Separation Using a Probabilistic Source Sparsity Model,” *Proceedings of ICA'01*, San Diego, California, pp. 675-679, 2001.

[14] L. Molgedey, H. Schuster, “Separation of Independent Sources Using Time-Delayed Correlations,” *Physical Review Letters*, vol. 72, pp. 3634-3637, 1994.

[15] O. Grellier, P. Comon, B. Mourrain, P. Tebuchet, “Analytical Blind Channel Identification,” *IEEE Transactions on Signal Processing*, vol. 50, no. 9, pp. 2196-2207, 2002.