SOURCE SEPARATION OF TEMPORALLY CORRELATED SOURCE USING BANK OF BAND PASS FILTERS

Andrzej Cichocki ♢ *and Adel Belouchrani* ‡

♢ Laboratory for Open Information Systems
Brain Science Institute, Riken, Japan.
E-mail: cia@brain.riken.go.jp
‡ Electrical Engineering Department
Ecole Nationale Polytechnique, Algiers, Algeria.
E-mail: belouchrani@hotmail.com

ABSTRACT

This paper introduces a new source separation algorithms exploiting the difference in the spectra shapes of the source signals. The proposed approach relies only on second-order statistics and estimates the mixing matrix by using eigenvalue decomposition of covariance matrix of prewithened sensor signals or alternatively an input output identification procedure using as inputs linear band pass filtered versions of the estimated colored sources. An adaptive implementation of the proposed technique is presented. The new algorithm shows to be computationally very simple and efficient. In addition and in contrast to other existing techniques, the covariance of the noise do not need to be modeled. The effectiveness of the proposed method is illustrated by some numerical simulations.

1. INTRODUCTION

In practical situations, one has to process multidimensional observation of the form

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) \tag{1}$$

where $\mathbf{x}(t)$ is a noisy instantaneous linear mixture of source signals, sampled at time t, $\mathbf{A}$ is $m \times n$ mixing matrix $\mathbf{n}(t)$ is the additive noise and $\mathbf{s}(t) = [s_1(t), \cdots, s_n(t)]^T$ consists of n source signals. Generally the waveform of source signals as well as their number are unknown and should be also estimated. Model (1) has show to be a good approximation in various areas as biomedical signal processing [1], factor analysis [2] or financial time series [3].

Blind source separation consists of retrieving the source signals without resorting to any *a priori* information about the mixing matrix $\mathbf{A}$; it exploits only the information carried by the received signals themselves, hence, the term blind.

Usually, signal separation algorithms are based on the main assumption of mutual independence of the source signals or their different autocorrelation functions what is equivalent to different spectra. Various techniques have been proposed. A first class of them are based on batch processing of higher order cumulants [4]. Another class is based on non-linear spatial adaptive filters [5, 6]. Both these classes assume non Gaussian source signals but do not require the exact knowledge of their distributions. When the source distributions are known, exact Maximum Likelihood approaches to solve the signal separation problem become possible [7, 8, 9]. In the case of time coherent source signals and even Gaussian source signals, solutions based on second order statistics are possible [10]. For non stationary source signals, blind source separation based on time frequency distributions have been considered in [11].

In this paper, we propose a new signal separation algorithm for sources with different spectra shapes. The proposed approach relies only on second-order statistics and estimates the mixing matrix by an input output identification procedure using as inputs band pass filtered versions of the estimated sources. The transfer functions of these filters are chosen to do not overlap in the frequency domain. The proposed method is implemented in an adaptive fashion. The new algorithm is computationally very simple and efficient. In contrast to other existing techniques, the proposed adaptive algorithm is robust to additive correlated noise.

2. PROBLEM STATEMENT

Assumptions: The source signal vector $\mathbf{s}(t)$ is assumed to be either *H1)* a deterministic sequence or *H2)* a stationary multivariate process with

$$H1) \qquad \lim_{T\to\infty} \frac{1}{T} \sum_{t=1,T} \mathbf{s}(t+\tau)\mathbf{s}(t)^* \overset{\text{def}}{=} E[\mathbf{s}(t+\tau)\mathbf{s}(t)^*]$$
$$= \mathrm{diag}[\rho_1(\tau), \ldots, \rho_n(\tau)]$$

$H2) \qquad E[\mathbf{s}(t+\tau)\mathbf{s}(t)^*] = \mathrm{diag}[\rho_1(\tau), \ldots, \rho_n(\tau)]$

where the superscript * denotes the conjugate transpose of a vector and $\mathrm{diag}[\cdot]$ is the diagonal matrix formed with the elements of its vector valued argument. For simplicity, the same notation E is used for the deterministic averaging operation under hypothesis (*H1*) and for ensemble averaging under (*H2*). This convention holds throughout. Assumptions (*H1*) or (*H2*) mean that the component processes $\mathbf{s}_i(t)$, $1 \leq i \leq n$ are mutually uncorrelated and $\rho_i(\tau) = E[\mathbf{s}_i(t+\tau)\mathbf{s}_i^*(t)]$ denotes the auto-correlation of $\mathbf{s}_i(t)$. The additive noise $\mathbf{n}(t)$ is modeled as a stationary, zero-mean random process and assumed to be decorreladed from the source signals. Contrary to classical assumptions, no assumption is made on either its distribution or its temporal or spatial correlation properties. The $m \times n$ matrix $\mathbf{A}$ is assumed to have full rank but is otherwise unknown.

3. THE NEW SECOND ORDER SEPARATION PRINCIPLE

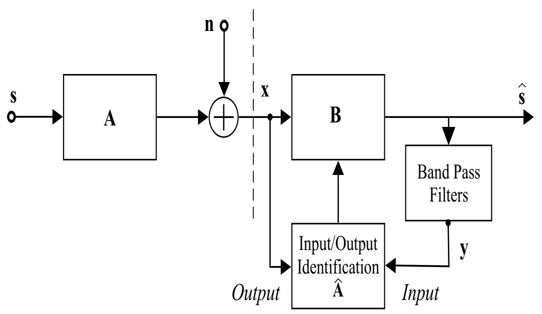

Fig. 1. The new second order separation principle.

The principle of the proposed algorithm is depicted in Figure 1. The concept is to estimate the mixing matrix by an input output identification procedure using as inputs a filtered version of the estimated sources. The transfer functions of the filters associated to each source are chosen to do not overlap in the frequency domain.
Let us start with data-assisted case when we know the source signals. In this case, it would be a simple matter to estimate the mixing matrix $\mathbf{A}$ using an input output identification. For this purpose, multiply equation (1) by the transpose conjugate of $\mathbf{s}(t)$ and compute the expectation of the obtained equation. This leads to

$$E[\mathbf{x}(t)\mathbf{s}(t)^*] = \mathbf{A}E[\mathbf{s}(t)\mathbf{s}(t)^*] + E[\mathbf{n}(t)\mathbf{s}(t)^*] \qquad (2)$$

Since, the additive noise is decorrelated from the source signals, we have

$$E[\mathbf{n}(t)\mathbf{s}(t)^*] = 0 \qquad (3)$$

Finally, we obtain

$$\mathbf{A} = E[\mathbf{x}(t)\mathbf{s}(t)^*]E[\mathbf{s}(t)\mathbf{s}(t)^*]^{-1} \qquad (4)$$

Equations (3) and (4) show that the proposed algorithm will have the important feature of being robust to colored additive noise. In practice, only an estimate over T samples of the expectations are available. Hence, the solution consists of evaluating the following statistics:

$$\mathbf{R}_{xs} = \frac{1}{T}\sum_{t=1}^{T} \mathbf{x}(t)\mathbf{s}(t)^H, \qquad (5)$$

$$\mathbf{R}_{ss} = \frac{1}{T}\sum_{t=1}^{T} \mathbf{s}(t)\mathbf{s}(t)^H \qquad (6)$$

and commuting the estimated mixing matrix by

$$\hat{\mathbf{A}} = \mathbf{R}_{xs}\mathbf{R}_{ss}^{-1} \qquad (7)$$

where ^{H} denotes the transpose conjugate operator.
In our problem, the source signals $\mathbf{s}(t)$ are of course unknown, otherwise we will have no need to perform the source separation. In this case, we propose to use instead of the original source signal a filtered version of the recurrently estimated sources. The transfer functions of the filters associated to each source are chosen to do not overlap in the frequency domain. Hence, the estimated matrix $\hat{\mathbf{A}}$ can be computed iteratively by employing a recurrent input output identification.

Comments: If the source signals do not overlap in the frequency domain, a simple spectral filtering of only one observed signal is sufficient. In the case of overlapping spectra sources, the spatial diversity provided by the multi-sensor array is necessary to achieve the separation by actually a spatial filtering, commonly known as beamforming. The proposed new approach suggests the combination in a recursive fashion of both the spectral and spatial filtering. Note that the spatial filtering is achieved by the band pass filters and the spatial filtering is achieved by the unmixing matrix. The latter is estimated by an Input-Output like identification, which links together the spectral and the spatial filtering.

4. IMPLEMENTATION

4.1. Bandpass filter implementation

The estimated source signal obtained at each iteration are filtered by IIR bandpass filters and in special case by a first order AR (autoregressive) filter described by

$$\mathbf{y}(t) = \hat{\mathbf{s}}(t) + \mathbf{a} \odot \mathbf{y}(t-1) \qquad (8)$$

where $\odot$ designates the Hadamard product and
$\mathbf{a} = [a_1, a_2, \ldots, a_n]^T$ a vector that contains the different complex-valued coefficients $a_i = |a_i|e^{j\phi_i}$. The argument of a_i controls the center frequency of the filter (that we would

like to be fit with the one of the source) and its modulus controls the bandwidth of the filter.

In the special case of real valued data, the following second order bandpass filtering can be used,

$$\begin{aligned}\mathbf{y}_1(t) &= \hat{\mathbf{s}}(t) + \mathbf{a} \odot \mathbf{y}_1(t-1) \\ \mathbf{y}_2(t) &= \hat{\mathbf{s}}(t) + \mathbf{a}^* \odot \mathbf{y}_2(t-1) \\ \mathbf{y}(t) &= \mathbf{y}_1(t) + \mathbf{y}_2(t)\end{aligned}$$

Note that $\mathbf{a}$ should be complex if we would like to obtain a band pass filter.

Another alternative realization of real bandpass filter with extremely easy adjustable central frequency and bandwidth is the 4-th order Buttherworth filter with the transfer function:

$$H(z) = \frac{b_0 + b_2 z^{-2} + b_4 z^{-4}}{1 + a_1 \omega z^{-1} (a_2 \omega^2 + a_2') z^{-2} + a_3 \omega z^{-3} + a_4 z^{-4}} \tag{9}$$

where

$$\begin{aligned}b_0 &= b_4 = 1/(d^2 + 2^{0.5} d + 1), \\ b_2 &= -2b_0, \\ b_1 &= b_3 = -4b_0, \\ a_1 &= -2d(2d + 2^{0.5})b_0, \\ a_2 &= 4d^2 b_0, \\ a_2' &= 2(d^2 - 1)b_0, \\ a_3 &= 2d(-2d + 2^{0.5})b_0, \\ a_4 &= (d^2 - 2^{0.5} d + 1)b_0, \\ d &= \mathrm{cotan}(\pi B),\end{aligned}$$

and

$$\omega = \frac{\cos(\pi f_1 + f_2)}{\cos(\pi B)} < 1, \tag{10}$$

where f_1 and f_2 are normalized lower and higher cutoff frequencies and B is normalized bandwidth. It should be noted that for fixed (constant) bandwidth B, ω is an only center frequency dependent parameter. It is worthwhile mentioning that the stability constraints on $H(z)$ are provided if

$$|\omega| < 1 \ \ \text{and} d > 0. \tag{11}$$

To show the advantage of the fourth-order Buttherworth bandpass filter in comparison with other simpler realization for wide band source signals lets us consider the standard second order band pass filter with transfer function

$$H_2(z) = (1-r)\frac{\omega z^{-1}/(r + r^2) - 1}{1 - \omega z^{-1} + r^2 z^{-2}} \tag{12}$$

where $\omega(n) = 2r\cos(2\pi f_0 n)$ is the only center frequency term and the parameter r is a fixed design one related to the frequency bandwidth B as follows

$$B = (1-r)/2$$

Such filter provides unity gain only at center frequency alone and make rather large distortion for source signal which are usually not pure sinusoids with some frequency variability. Therefore, by choosing a very narrow bandwidth, this filter can be applied for tracking and enhancement of single sinusoid in white noise. In contrast the forth-order Buttherworth filter has flat characteristic around central frequency and enable enhance arbitrary narrow band source signal with low distortion.

By maximizing the output power of the band pass filter we can adjust automatically the center frequency to the input bandpass signal.

The filter output is given by

$$\begin{aligned}y(n) &= b_0 s(n) + b_2 s(n - 2_+ b_4 s(n-4) - a_1 \\ &\quad \omega(n) y(n-1) - (a_2 \omega^2(n) + a_2') y(n-2) \\ &\quad a_3 \omega(n) y(n-3) - a_4 y(n-4)\end{aligned} \tag{13}$$

In order to find an optimal value of the parameter $\omega(n)$ for a specific bandwidth, we can maximize cost function $E\{y^2(n)\}$ using gradient ascent procedure and obtain simple learning rule

$$\omega(n+1) = \omega(n) + \mu(n) y(n) \alpha(n) \tag{14}$$

where $\mu(n) = \mu/r(n)$, $r(n) = \lambda r(n-1) + \alpha^2(n)$ and

$$\begin{aligned}\alpha(n) &= \frac{\partial y(n)}{\partial \omega(n)} \\ &= -a_1 y(n-1) - 2a_2 \omega(n) y(n-2) \\ &\quad -a_3 y(n-3) - a_1 \omega(n) \alpha(n-1) \\ &\quad -(a_2' + a_2 \omega(n)^2) \alpha(n-2) \\ &\quad -a_3 \omega(n) \alpha(n-3) - a_4 \alpha(n-4)\end{aligned} \tag{15}$$

4.2. Outline of the proposed adaptive algorithm

AT TIME-INSTANT $t+1$:

1. Get a filtered version $\mathbf{y}(t)$ of the previous estimate as described above.
2. Compute the statistics:

$$\begin{aligned}\mathbf{R}_{xy}^{(t)} &= (1-\mu)\mathbf{R}_{xy}^{(t-1)} + \mu \mathbf{x}(t)\mathbf{y}(t)^T \\ \mathbf{R}_{yy}^{(t)} &= (1-\mu)\mathbf{R}_{yy}^{(t-1)} + \mu \mathbf{y}(t)\mathbf{y}(t)^T\end{aligned}$$

   where μ is a decreasing and positive sequence.
3. Update the value of the mixing matrix:

$$\hat{\mathbf{A}}^{(t)} = \mathbf{R}_{xy}^{(t)} {\mathbf{R}_{yy}^{(t)}}^{-1}$$

4. Estimate the source signals at time $t+1$ by

$$\hat{\mathbf{s}}(t+1) = \mathbf{R}_{yy}^{(t)} {\mathbf{R}_{xy}^{(t)}}^{\#} \mathbf{x}(t+1)$$

To avoid the matrix inversion in step 3 and 4, one can update directly $\mathbf{R}_{yy}^{-1}$ or $\mathbf{R}_{xy}^{\#}$ using the Woodbury's identity. This leads to the following adaptation,

$$\mathbf{R}_{yy}^{-1\,(t)} = \frac{1}{1-\mu}[\mathbf{R}_{yy}^{-1\,(t-1)} - \frac{\mu \mathbf{R}_{yy}^{-1\,(t-1)}\mathbf{y}(t)\mathbf{y}(t)^H\mathbf{R}_{yy}^{-1\,(t-1)}}{(1-\mu)-\mu \mathbf{y}(t)^H \mathbf{R}_{yy}^{-1\,(t-1)}\mathbf{y}(t)}]$$

$$\mathbf{R}_{xy}^{\#\,(t)} = \frac{1}{1-\mu}[\mathbf{R}_{xy}^{\#\,(t-1)} - \frac{\mu \mathbf{R}_{xy}^{\#\,(t-1)}\mathbf{x}(t)\mathbf{y}(t)^H\mathbf{R}_{xy}^{\#\,(t-1)}}{(1-\mu)-\mu \mathbf{x}(t)^H \mathbf{R}_{xy}^{\#\,(t-1)}\mathbf{y}(t)}]$$

5. SEPARATION PROCESS ANALYSIS

In this section, we analyze the separation process of the proposed algorithm. This analysis is done in the noiseless case. Substituting $\mathbf{x}(t)$ by $\mathbf{As}(t)$ in the equation at step 4 of the above adaptive algorithm, we get

$$\hat{\mathbf{s}}(t+1) = \mathbf{R}_{yy}^{(t)}\mathbf{R}_{sy}^{(t)\#}\mathbf{s}(t+1) \tag{16}$$

This equation shows that our algorithm satisfies the equivariant property stated in [12]. The performance of such algorithm do not depend on the mixing matrix $\mathbf{A}$ in the noiseless case.

Note now that we can also write

$$\mathbf{y}(t) = \mathbf{R}_{yy}^{(t)}\mathbf{R}_{sy}^{(t)\#}\mathbf{s}(t) \tag{17}$$

According to equations (16) and (17), we have

$$[\mathbf{R}_{\hat{s}\hat{s}}^{(t+1)}]_{ij} = [\mathbf{R}_{yy}^{(t)}]_{ij}, \text{ for } i \neq j \tag{18}$$

and

$$[\mathbf{R}_{s\hat{s}}^{(t+1)}]_{ij} = [\mathbf{R}_{sy}^{(t)}]_{ij}, \text{ for } i \neq j \tag{19}$$

Note that $[\mathbf{R}_{\hat{s}\hat{s}}^{(t+1)}]_{ij}$ ($[\mathbf{R}_{s\hat{s}}^{(t+1)}]_{ij}$) and $[\mathbf{R}_{yy}^{(t)}]_{ij}$ ($[\mathbf{R}_{sy}^{(t)}]_{ij}$) are nothing than the correlation between $\hat{s}_i$ and $\hat{s}_j$ (s_i and $\hat{s}_j$), and between y_i and y_j (s_i and y_j), respectively.

Under the assumption of no overlapping band pass filters and source signals with different spectra shape, we can see that

$$[\mathbf{R}_{yy}^{(t+1)}]_{ij} < [\mathbf{R}_{\hat{s}\hat{s}}^{(t+1)}]_{ij}, \text{ for } i \neq j \tag{20}$$

and

$$[\mathbf{R}_{sy}^{(t+1)}]_{ij} < [\mathbf{R}_{s\hat{s}}^{(t+1)}]_{ij}, \text{ for } i \neq j \tag{21}$$

From equations (18), (19), (20) and (21), we obtain

$$[\mathbf{R}_{yy}^{(t+1)}]_{ij} < [\mathbf{R}_{yy}^{(t)}]_{ij}, \text{ for } i \neq j \tag{22}$$

and

$$[\mathbf{R}_{sy}^{(t+1)}]_{ij} < [\mathbf{R}_{sy}^{(t)}]_{ij}, \text{ for } i \neq j \tag{23}$$

Equations (22) and (23) suggest that at convergence ($t \to \infty$), matrices $\mathbf{R}_{yy}^{(\infty)}$ and $\mathbf{R}_{sy}^{(\infty)}$ are diagonal matrices[1]. Hence, $\hat{\mathbf{s}}$ converges to the correct source signals up to a permutation and a scalar factor.

[1] Note that the diagonal elements of $\mathbf{R}_{yy}^{(\infty)}$ and $\mathbf{R}_{sy}^{(\infty)}$ are different from zero as long as source signal energy exists in the frequency band of each filter.

6. NUMERICAL SIMULATION

In this section, the performance of the proposed blind source separation method, as investigated via computer simulations, is reported.

In the simulated environment, a 2-element uniform linear array having half wavelength sensor spacing receives two signals arriving from different directions $\phi_1 = 0$ and $\phi_2 = 20$ degrees (the particular structure of the array manifold is of course not exploited by the proposed algorithm). The additive noise is generated from a zero mean and temporally white Gaussian process with the following covariance matrix,

$$\mathbf{R}_n = \sigma^2 \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \tag{24}$$

where σ^2 is the noise power and ρ is the coefficient of noise correlation. The step size parameter μ is fixed to 0.5.

Sample run: In this example, the source signal are a QAM-16[2] and a complex sinusoid at normalized frequency of 0.3. A Signal to Noise Ratio (SNR) of 20 dB is chosen for this experiment with $\rho = 0$. Figures 2 and 3 present a sample run corresponding to this example using the signal space representation and the spectral representation, respectively. The observed results show clearly that the proposed algorithm has succeeded in separating the source signals.

Figure 4 displays the rejection level defined as

$$\frac{1}{n(n-1)} \sum_{p \neq q} |(\hat{\mathbf{A}}^{\#}\mathbf{A})_{pq}|^2 \tag{25}$$

where $\hat{\mathbf{A}}$ is the estimated mixture matrix and $^{\#}$ denotes the pseudo-inverse. It is plotted in dB versus time iterations. For this figure the plotting convention is: solid line for the proposed algorithm; dashed line for the EASI algorithm [12]. The plots show fast convergence and small steady state error with respect to the EASI technique.

Performance evaluation: This section deals with the performance evaluation of the proposed algorithm through series of experiments. For this purpose, the source signals are generated by filtering a complex circular white Gaussian processes by an AR model of order one with coefficient $a_1 = 0.85\exp(j0.3)$ and $a_2 = 0.85\rho_2\exp(j(0.3+\delta_s))$, respectively. The parameter δ_s accounts for the spectral shift between the spectrums of the two sources. The overall rejection level is evaluated over 100 independent runs.

In Figure 5, the SNR is kept constant first at 5 dB (for $\rho = 0$ and $\rho = 0.5$) and than at 20 dB (for $\rho = 0$ and $\rho = 0.5$). The curves show the mean rejection level in dB plotted as against the 'spectral shift' δ_s. These plots show that the proposed approach is robust with respect to spatially colored noise.

[2] Quadrature Amplitude Modulation with 16 constellations.

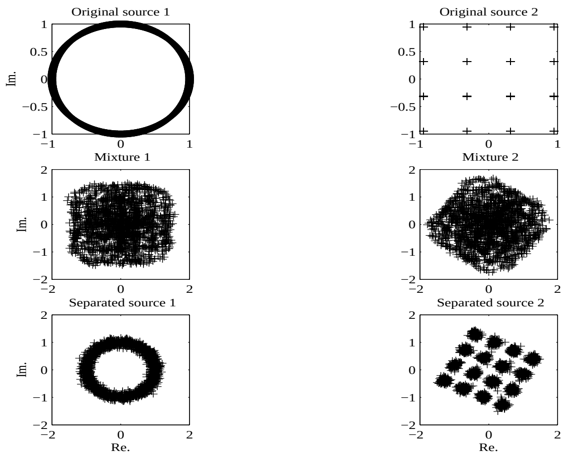

Fig. 2. Sample run.

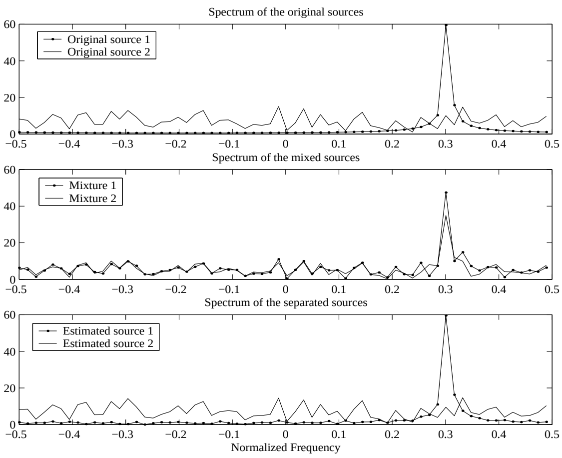

Fig. 3. Spectral representation of sample run.

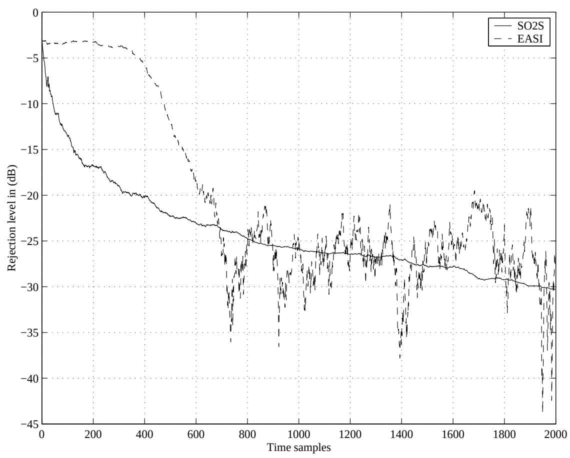

Fig. 4. Sample run: Rejection level in dB versus time iterations.

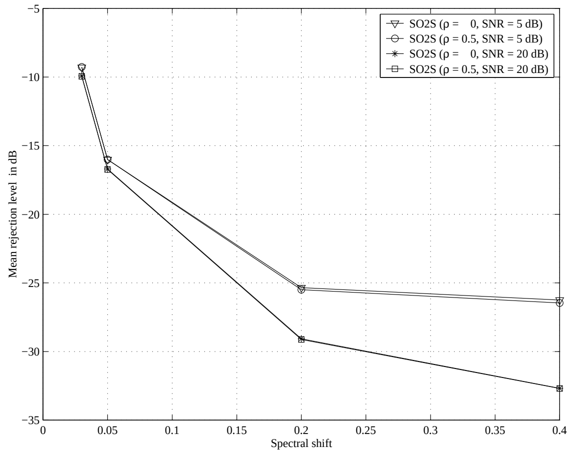

Fig. 5. Performance evaluation: Mean rejection level in dB versus spectral shift.

7. CONCLUSION

In this contribution, we have introduced a new blind source separation algorithm for sources with different spectra shapes. The technique is based on a recurrent input output identification using as inputs band pass filtered versions of the estimated sources. This method shows a number of attractive features: i) it relies only on second order statistics of the received signals, ii) allows -in contrast to higher order statistic techniques- the separation of Gaussian sources, iii) improvement of the quality of the separation, iv) ability to deal with additive noise of unknown covariance, v) computationally very simple and efficient. An adaptive implementation of this approach was proposed.

8. REFERENCES

[1] T.-P. Jung, C. Humphries, T.-W. Lee, S. Makeig, M. J. McKeown, V. Iragui and T.J. Sejnowski, “Extended ICA removes artifacts from electroencephalographic recordings,” *Advances in Neural Information Processing Systems* (NIPS), vol. 10, Cambrige, MA: MIT Press, 1998.

[2] E. E. Cureton and R. B. D’Agostino, *FACTOR ANALYSIS An Applied Approach*. Lawrence Erlbaum Associates, 1983.

[3] A. D. Back and A. S. Weigend,“A first application of Independent Component Analysis to extracting structure from stock returns,” *International Journal of Neural Systems*, Special Issue on Data Mining in Finance, vol. 8, No. 5, October 1997.

[4] J.-F. Cardoso and A. Souloumiac, “An efficient technique for blind separation of complex sources,” in

Proc. IEEE SP Workshop on Higher-Order Stat., Lake Tahoe, USA, 1993.

[5] A. Cichocki, W. Kasprzak and S. Amari, “Multi-layer neural networks with a local adaptive learning rule for blind separation of source signals,” in *Proc. Nolta, Las Vegas, USA*, Dec. 1995.

[6] A. Belouchrani and K. Abed Meraim, “Constant modulus blind source separation technique: A new approach,” in *Proc. ISSPA, Gold Coast, Australia*, August 1996.

[7] A. Belouchrani and J.-F. Cardoso, “Maximum likelihood source separation for discrete sources,” in *Proc. EUSIPCO*, pp. 768–771, 1994.

[8] A. Belouchrani and J.-F. Cardoso, “Maximum likelihood source separation by the expectation-maximization technique: Deterministic and stochastic implementation,” in *Proc. International Symposium on Nonlinear Theory and its Applications, Nolta’95, Las Vegas, Nevada, U.S.A.*, Dec. 1995. Invited paper.

[9] A. Belouchrani and K. Abed Meraim, “An adaptive optimization for digital signal separation via the em algorithm,” in *In Proc. of 4th International Conference on Optimization: Techniques and Applications*, July 1998.

[10] A. Belouchrani and K. Abed Meraim and J.-F Cardoso and E. Moulines, “A blind source separation technique using second order statistics,” *IEEE Trans. on SP*, vol. 45, pp. 434–444, Feb. 1997.

[11] A. Belouchrani and M. G. Amin, “Blind source separation based on time-frequency signal representation,” *IEEE Trans. on SP*, vol. 46, pp. 2888–2898, Nov. 1998.

[12] J.-F. Cardoso and B. Laheld, “Equivariant adaptive source separation,” *IEEE Trans. on SP*, vol. 43, pp. 3017–3029, Dec. 1996.