

# BAYESIAN BLIND SOURCE SEPARATION FOR BRAIN IMAGING.

*Hichem Snoussi<sup>a,\*</sup>, Vince D. Calhoun<sup>b</sup>*

<sup>a</sup> IRCCyN, Institut de Recherche en Communications et Cybernétiques de Nantes,  
Ecole Centrale de Nantes, 44321, Nantes, France,  
email: Hichem.Snoussi@irccyn.ec-nantes.fr

<sup>b</sup> Olin Neuropsychiatry Research Center, Institute of Living, Hartford, CT 06106,  
email: vince.calhoun@yale.edu

## ABSTRACT

This paper deals with the problem of blind source separation in fMRI data analysis. Our main contribution is to present a maximum likelihood based method to blindly separate the brain activations in an fMRI experiment. Choosing the time frequency domain as the signal representation space, our method relies on the second order statistics and exploits the inter-source diversity. It is efficiently implemented by the EM (Expectation-Maximization) algorithm where the time courses of the brain activations are considered as the hidden variables. The estimation variance of the STFT (Short Time Fourier Transform) is reduced by averaging across time frequency sub-domains. The successful separation of the right and left visual cortex activations during a visual fMRI experiment, in a block design, and the extraction of only the relevant tasks corroborate the effectiveness of our proposed separating algorithm.

## 1. INTRODUCTION

Independent Component Analysis (ICA) consists in linearly demixing a set of observed data. The separation principle in this method is based on the statistical independence of the reconstructed sources [1]. However, ICA is designed to efficiently work in the noiseless case. In addition, with the i.i.d assumption in source modeling, the separation necessarily relies on high order statistics and treating the noisy case with the maximum likelihood approach leads to complicated algorithms [2].

Discarding the i.i.d assumption, blind source separation can be achieved with second order statistics. For instance, second order correlation diversity in the time domain, frequency domain or time frequency domain [3] are successfully used to blindly separate the sources. Non stationary second order based methods are also proposed in [4]. Stationarity and non stationarity can approximately be seen as dual under Fourier transformation. We have recently proposed a maximum likelihood method to separate noisy mixture of Gaussian stationary sources exploiting this temporal / spectral duality [5]. The Gaussian model of sources allows an efficient implementation of the EM algorithm [6]. In this contribution, we extend this approach to deal with non stationary sources. Relying on the maximum likelihood principle and the Short Time Fourier Transform (STFT), our approach can be interpreted as a regularized blind identification of the sources spectra. Our method is based on the estimation of the mixing matrix, the sources spectra and the noise covariance matrix. Thus, the same algorithm is

applied to over-determinate and underdeterminate cases without a prewhitening step. The method implicitly incorporates a denoising procedure and it is consequently robust to high level noise. The equations involved in the EM algorithm are very simple to implement. An interesting property of the proposed solution is the exploitation of second order spectral non stationarity. The frequency marginal spectrum (integrating over time the spectrograms) corresponds to an improved version of the separation of noisy stationary sources [5] as the smoothed periodograms, obtained by marginalization, are used instead of the empirical periodograms (corresponding to the Wigner-Ville distribution marginals).

The paper is organized as follows. Section 2 is devoted to the main contribution of this paper. We develop the EM algorithm implementing the maximum likelihood solution in the time frequency domain. The likelihood criteria is interpreted as a regularized matching between the spectral covariances. In Section 3, results on real fMRI signals illustrate the effectiveness of our proposed method comparing to the ICA solution.

## 2. REGULARIZED SPECTRAL MATCHING

The temporal fMRI separation relies on the following mixing model:

$$\mathbf{X} = \mathbf{A}\mathbf{S} + \mathbf{N},$$

where  $\mathbf{X}$  is the  $(M \times T)$  matrix of observations, the column  $\mathbf{X}(:, t)$  contains the scanned image acquired at the time  $t$  and  $M$  is the number of voxels. The  $(N \times T)$ -matrix  $\mathbf{S}$  (the sources) contains the  $N$  time courses rows. The  $(M \times N)$  mixing matrix  $\mathbf{A}$  contains the  $N$  brain region activations. Each column  $(M \times 1)$  of  $\mathbf{A}$  represents a source image and it is time invariant. Thus, the column  $\mathbf{S}(:, t)$  represents the linear combination at the instant  $t$  to produce the image  $\mathbf{X}(:, t)$ . The matrix  $\mathbf{N}$  models the noise corrupting the observations. The advantage of taking into account the noise in the model is to allow the separating algorithm to only extract the relevant sources, that is the task related brain activations. This is possible when the spectral profile of the noise is flat comparing to the more concentrated source spectra. In fact, the time courses of task related activations present the same frequency content as the stimulus.

Some of the real signals collected in fMRI imaging are obviously non stationary. The difficulties thus arising when separating the different **temporal** brain activations are the fact that observations are mixture of two types of sources: stationary sources (task

related activations, thermal noise,...) and non stationary sources (artifacts). In addition, the noise  $N$  is not white and especially when the observation model is modified to segregate task related signals from non task related signals.

In the following, we outline the proposed EM algorithm called hereafter the "Regularized Spectral EM" as it is the extension of the Spectral EM [5] and we show some separation results on noisy real fMRI data.

The Short Time Fourier Transform (STFT) of a signal  $\{x(t)\}$  is a windowed Fourier transform defined as:

$$S_x(t, \omega) = \int x(\tau)h(\tau - t)e^{-j\omega\tau} d\tau,$$

where  $h$  is the moving window capturing the signal non stationarity. It is shown that the squared modulus of  $S_x$  (called the spectrogram) belongs to the Cohen's Class with the kernel  $\phi$  equal to  $W_h$ , the Wigner-Ville distribution of the window  $h$ . Thus, the spectrogram enjoys the positivity property but does not conserve the marginal properties of the Wigner-Ville distribution.

Exploiting the linearity of the STFT transform, the noisy linear mixture model conserves its algebraic form under this transformation:

$$\mathbf{x}(t, \omega) = \mathbf{A}\mathbf{s}(t, \omega) + \mathbf{n}(t, \omega), t = 1..T, \omega = 1..F,$$

where, for the sake of clarity,  $\mathbf{x}$ ,  $\mathbf{s}$  and  $\mathbf{n}$  also denote the STFT transforms of the observations, the sources and the noise respectively. Assuming that the noise is stationary white (with unknown covariance  $\mathbf{R}_n$ ) and that the sources are decorrelated in the time frequency domain<sup>1</sup> (with unknown diagonal covariances  $\{\mathbf{P}(t, \omega) = \mathbb{E}[\mathbf{s}(t, \omega)\mathbf{s}(t, \omega)^*]\}_{t=1..T}^{\omega=1..F}$ ), the likelihood is as follows:

$$\begin{aligned} p(\mathbf{X} | \boldsymbol{\theta}) &= \int p(\mathbf{X} | \mathbf{S}, \mathbf{A}, \mathbf{R}_n) p(\mathbf{S} | \{\mathbf{P}(t, \omega)\}) d\mathbf{S} \\ &= \prod_{t, \omega} |2\pi \mathbf{R}_{t, \omega}|^{-1} \exp[-\text{Tr}(\mathbf{R}_{t, \omega}^{-1} \mathbf{x}(t, \omega)\mathbf{x}(t, \omega)^*)], \end{aligned} \quad (1)$$

where  $\mathbf{R}_{t, \omega} = \mathbf{A}\mathbf{P}_{t, \omega}\mathbf{A}^* + \mathbf{R}_n$  and  $\boldsymbol{\theta}$  is the whole parameter to be estimated ( $\mathbf{A}, \mathbf{R}_n, \{\mathbf{P}(t, \omega)\}$ ).

The likelihood (1) can be interpreted as the matching between STFT covariances matrices  $\mathbf{R}_{t, \omega} = \mathbf{A}\mathbf{P}_{t, \omega}\mathbf{A}^* + \mathbf{R}_n$  and empirical covariances  $\hat{\mathbf{R}}_{t, \omega} = \mathbf{x}(t, \omega)\mathbf{x}(t, \omega)^*$ , in the Kullback-Leibler metric:

$$\log p(\mathbf{X} | \boldsymbol{\theta}) = - \sum_{t, \omega} D_{KL}(\mathbf{R}_{t, \omega} \parallel \hat{\mathbf{R}}_{t, \omega}) \quad (2)$$

## 2.1. Time frequency EM algorithm

The first step of the EM algorithm is the computation of the functional  $\mathcal{Q}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(m-1)})$ :

$$\mathcal{Q}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(m-1)}) = \mathbb{E}[\log p(\mathbf{X}, \mathbf{S} | \boldsymbol{\theta}) | \mathbf{X}, \boldsymbol{\theta}^{(m-1)}]$$

<sup>1</sup>The decorrelation assumption of the time frequency source points is only statistically valid for underspread signals, i.e. the ambiguity function is concentrated in a small neighborhood of the origin [7]. However, our main objective is the estimation of the unknown parameters and not the filtering of sources.

Defining the following statistics which will be computed later:

$$\begin{cases} \mathbf{R}_{xx}(t, \omega) &= \mathbf{x}_{t, \omega}\mathbf{x}_{t, \omega}^* \\ \mathbf{R}_{xs}(t, \omega) &= \mathbf{x}_{t, \omega}\mathbb{E}[\mathbf{s}_{t, \omega} | \mathbf{x}_{t, \omega}, \boldsymbol{\theta}^{(m-1)}]^* \\ \mathbf{R}_{ss}(t, \omega) &= \mathbb{E}[\mathbf{s}_{t, \omega}\mathbf{s}_{t, \omega}^* | \mathbf{x}_{t, \omega}, \boldsymbol{\theta}^{(m-1)}] \end{cases} \quad (3)$$

the functional  $\mathcal{Q}$  can be rewritten in the following form:

$$\begin{aligned} \mathcal{Q}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(m-1)}) &= \sum_{t, \omega} -\log |\mathbf{R}_n| - \text{Tr}(\mathbf{R}_n^{-1}[\mathbf{R}_{xx}(t, \omega) \\ &+ \mathbf{A}\mathbf{R}_{ss}(t, \omega)\mathbf{A}^* - \mathbf{A}\mathbf{R}_{sx}(t, \omega) - \mathbf{R}_{sx}^*(t, \omega)\mathbf{A}^*]) \\ &+ \sum_{t, \omega} -\log |\mathbf{P}_s(t, \omega)| - \text{Tr}(\mathbf{P}_s^{-1}(t, \omega)\mathbf{R}_{ss}(t, \omega)) \end{aligned} \quad (4)$$

The second step is the update of the parameter  $\boldsymbol{\theta}$  by maximizing the functional  $\mathcal{Q}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(m-1)})$ :

$$\boldsymbol{\theta}^{(m)} = \arg \max_{\boldsymbol{\theta}} \mathcal{Q}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(m-1)})$$

This can be achieved by differentiating the functional  $\mathcal{Q}$  (4) with respect to the parameter  $\boldsymbol{\theta}$  and then equating to zero the partial derivatives. We obtain the following simple updating equations:

$$\begin{cases} \mathbf{A}^{(m)} &= \mathbf{R}_{xs}\mathbf{R}_{ss}^{-1} \\ \mathbf{R}_n^{(m)} &= \mathbf{R}_{xx} - \mathbf{R}_{xs}\mathbf{R}_{ss}^{-1}\mathbf{R}_{sx} \\ \mathbf{P}_s(t, \omega) &= \text{diag}(\mathbf{R}_{ss}(t, \omega)) \end{cases} \quad (5)$$

where the matrices  $\mathbf{R}_{xx}$ ,  $\mathbf{R}_{xs}$  and  $\mathbf{R}_{ss}$  are the average of the statistic matrices  $\mathbf{R}_{xx}(t, \omega)$ ,  $\mathbf{R}_{xs}(t, \omega)$  and  $\mathbf{R}_{ss}(t, \omega)$  defined in (3), over the time frequency domain.

The computation of the statistic matrices (3) is essentially based on the computation of the *a posteriori* first and second moments of the source vector  $\mathbf{s}_{t, \omega}$ . Thanks to the *a priori* Gaussianity of sources and noise, the *a posteriori* distribution of the sources is also Gaussian with the following moments:

$$\begin{cases} \mathbb{E}[\mathbf{s}_{t, \omega}] &= \mathbf{W}_{t, \omega}\mathbf{x}_{t, \omega} \\ \mathbb{E}[\mathbf{s}_{t, \omega}\mathbf{s}_{t, \omega}^*] &= \mathbf{V}_{t, \omega} + \mathbb{E}[\mathbf{s}_{t, \omega}]\mathbb{E}[\mathbf{s}_{t, \omega}^*]^* \end{cases}$$

where the matrices  $\mathbf{W}_{t, \omega}$  (Wiener matrix) and  $\mathbf{V}_{t, \omega}$  (*a posteriori* covariance) have the following expressions:

$$\begin{cases} \mathbf{W}_{t, \omega} &= [\mathbf{A}^*\mathbf{R}_n^{-1}\mathbf{A} + \mathbf{P}_s^{-1}(t, \omega)]^{-1}\mathbf{A}^*\mathbf{R}_n^{-1} \\ \mathbf{V}_{t, \omega} &= [\mathbf{A}^*\mathbf{R}_n^{-1}\mathbf{A} + \mathbf{P}_s^{-1}(t, \omega)]^{-1} \end{cases}$$

We note that the equations are very similar to a time frequency Wiener filtering. Consequently, the EM algorithm involves an implicit denoising procedure when computing the first *a posteriori* moment of the sources. In other words, we have an optimal source reconstruction at each step of the algorithm. It is worth noting that the achievement of the separation solution is strongly linked to the diversity of the sources spectrograms (the diagonal time frequency distributions of the matrices  $\mathbf{P}(t, \omega)$  are different). This is the fundamental reason to perform the separation in the frequency domain when the only temporal statistics are not able to provide such diversity.

## 2.2. Spectrum Regularization

The estimation of the parameter  $\theta$  involves the estimation of the whole spectrograms ( $\{P_s(t, \omega)\}_{t=1..T}^{\omega=1..F}$ ) which are smoothed versions of the Wigner-Ville spectra. In order to accelerate the EM algorithm, we can partition the time frequency domain into  $L$  horizontal sub-domains  $\{\mathcal{D}_l\}_{l=1}^L$  and then estimate the averaged spectrograms inside these domains. This is algorithmically equivalent to assume that the spectrograms are constant in the sub-domains in the partitioned time frequency 2-D field. Figure 1 illustrate the horizontal segmentation of the time frequency domain.

We assume then that  $P_s(t, \omega) = P_l$  for all  $(t, \omega) \in \mathcal{D}_l$ . The statistics  $R_{xx}(l)$ ,  $R_{xs}(l)$  and  $R_{ss}(l)$  (3) are also constant in the domain  $\mathcal{D}_l$ . Partitioning the time frequency domain into horizontal bands sub-domains  $\{\mathcal{D}_l\}$ , the matching of STFT spectra leads to the same algorithm as in [5] exploiting the temporal stationarity but with regularized spectra. In fact, the projection of the STFT spectrum yields the windowed power spectrum. Maximizing the likelihood is then equivalent to matching the windowed periodograms according to equation (2). Thus, the method will essentially consist in maximizing the likelihood of the parameters based on the Gaussian modeling of the sources.

As the computation of the observations spectra is performed off-line, the structure of the algorithm is independent of the partition choice. In fact, the algorithm is only based on matching the computed matrices to structured matrices according to the mixture model. Hereafter the pseudo code of the Regularized Spectral EM algorithm:

**Regularized Spectral EM**

```

1: Initializing:
2:   Off line computation of the smoothed
   covariances  $R_{xx}(l)$ 
3:   Initial values for  $A$ ,  $R_n$  and  $P_l$ 
4: repeat until convergence,
5:   //----- E-step -----//
6:   for l=1 to L, compute statistics:
7:      $V_l = (AR_n^{-1}A^* + P_l^{-1})^{-1}$ 
8:      $R_{xs}(l) = R_{xx}(l)R_n^{-1}AV_l$ 
9:      $R_{ss}(l) = V_lA^*R_n^{-1}R_{xx}(l)R_n^{-1}AV_l + V_l$ 
10:  end of loop on l,
11:   $R_{xs} = \frac{1}{L} \sum w_l R_{xs}(l)$ 
12:   $R_{ss} = \frac{1}{L} \sum w_l R_{ss}(l)$ 
13:  //-----M-step-----//
14:   $A = R_{xs}R_{ss}^{-1}$ 
15:   $R_n = \text{diag}(R_{xx} - R_{xs}R_{ss}^{-1}R_{xs}^*)$ 
16:   $P_l = \text{diag}(R_{ss}(l))$ , for l=1 to L
17:  Renormalize  $A$  and  $P_l$ 
18: end of repeat

```

## 3. ILLUSTRATION ON FMRI DATA

The Regularized Spectral EM algorithm was applied to separate the time courses of fMRI data acquired at the FM Kirby Center for Functional Brain Imaging. The experiment consisted of presenting two periodic visual stimulus, shifted by 20 s from one another, to the subject. The stimuli consisted of an 8-Hz reversing checkerboard pattern presented for 15 s in the right visual hemifield, followed by 5 s of an asterisk fixation, followed by 15 s of checkerboard presented to the left visual hemifield, followed by 20 s of an

asterisk fixation. The 55 s set of events was repeated four times for a total of 220 s. Scans were acquired on a Philips NT 1.5-Tesla scanner. A sagittal localizer scan was performed first, followed by a  $T_1$ -weighted anatomic scan [repeat time ( $TR$ ) = 500 ms, echo time (TE) = 30 ms, field of view = 24 cm, matrix = 256 × 256, slice thickness = 5 mm, gap = 0.5 mm] consisting of 18 slices through the entire brain including most of the cerebellum. Next, we acquired functional scans over the same 18 slices consisting of a single-shot, echo-planar scan (TR=1 s, TE= 39 ms, field of view = 24 cm, matrix= 64 × 64, slice thickness = 5 mm, gap = 0.5 mm, flip angle = 90 degrees) obtained consistently over a 3-min, 40-s period for a total of 220 scans.

Our method is tested on two different slices where we expect, in each, two different task related components corresponding to the alternating activation of the right and left visual cortex as a response to an alternating visual stimulus presented to the subject. However, we show results for only one slice (slice 10). In Figure 2, we have plotted three recovered image sources (the three columns of the estimated mixing matrix  $\hat{A}$ ) within their corresponding estimated time courses. We note the ability of the algorithm to extract the sources which have a time course correlated with the stimulus. The first and third sources correspond to the alternative activations of the right and left visual cortex as expected from the conditions of the stimulus presented to the subject. However, the algorithm extracts also another source (the second one) which has a spectral density similar to the first two sources. This shows that fixing the number of sources by intuitive expectation based on the experiment paradigm leads to wrong results. The results shown in Figure 2 were obtained by varying the number of sources and then studying *a posteriori* the results after convergence of the EM algorithm. An automatic selection of the number of components is thus needed for a complete blind analysis of the fMRI data. Figure 3 illustrates the times courses of the right and left visual cortex regions. We note the periodicity of the time courses and their relative inter-delay (around 20 s) corresponding to the inter-delay of the stimulus (the checkerboard pattern was presented alternatively to the right and left visual hemifields). For comparison purposes, we reported the separation results of a temporal ICA InfoMax [8] algorithm on the same data set, in Figure 4. We have fixed the number of sources to 3. The ICA algorithm fails to extract the third source from the two alternative task related sources identified with the EM algorithm, mixing them with a higher frequency component.

## 4. CONCLUSION

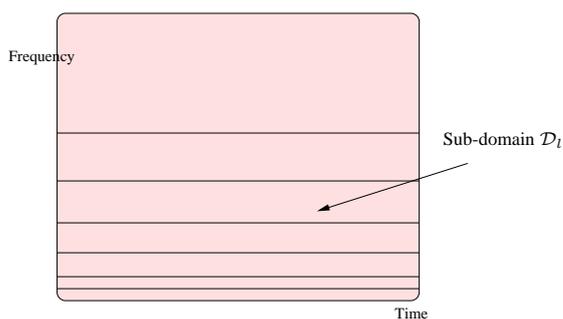
We have presented an EM algorithm to deal with real data suffering from non stationarity and a lack of enough points for spectral analysis. The separation method is essentially based on the diversity of the source smoothed periodograms. The non stationarity of second order statistics allows the mixing matrix identification without resorting to higher order statistics. The use of second order statistics (in other words, the Gaussian modeling) leads to an efficient and fast implementation of the EM algorithm. In fMRI data analysis, we have exploited this diversity between the time course spectra. The spatial pixel distributions represent the columns of the mixing matrix. Therefore, the Regularized Spectral EM algorithm allows a blind joint estimation of the brain source images within the spectra of their time courses. The task related sources are easily distinguished by the signature of the stimulus in their time course spectra.

## Acknowledgment

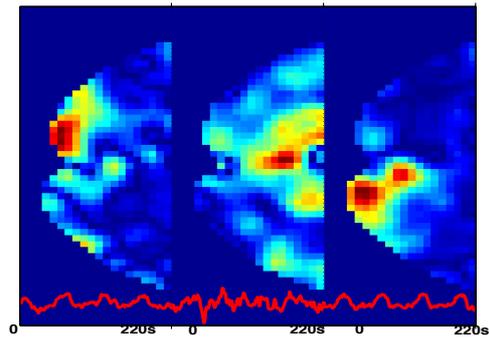
This work was supported in part by the National Institutes of Health under grant 1 R 01 EB 000840-01 (to VC).

### 5. REFERENCES

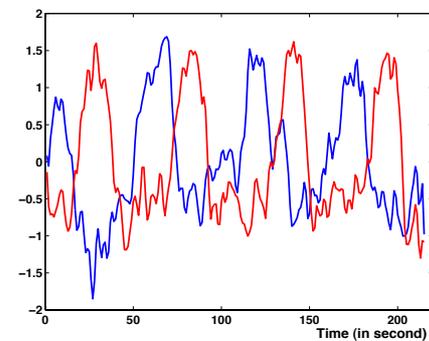
- [1] P. Comon, "Independent Component Analysis, a new concept?", *Signal processing, Special issue on Higher-Order Statistics, Elsevier*, vol. 36 (3), pp. 287–314, April 1994.
- [2] H. Snoussi and A. Mohammad-Djafari, "Bayesian unsupervised learning for source separation with mixture of gaussians prior", *Int. Journal of VLSI Signal Processing Systems for Signal, Image and Video Technology*, vol. 37, no. 2–3, pp. 263–279, June–July 2004.
- [3] A. Belouchrani and M. Amin, "Blind source separation using time-frequency distributions: algorithm and asymptotic performance", in *Proc. ICASSP, Munchen, 1997*, pp. 3469 – 3472.
- [4] K. Matsuoka, M. Ohya, and M. Kawamoto, "A neural net for blind separation of nonstationary sources", *Neural Networks*, vol. 8(3), pp. 411–419, 1995.
- [5] H. Snoussi, G. Patanchon, J. Macías-Pérez, A. Mohammad-Djafari, and J. Delabrouille, "Bayesian blind component separation for cosmic microwave background observations.", in *Bayesian Inference and Maximum Entropy Methods*, R. L. Fry, Ed. MaxEnt Workshops, August 2001, pp. 125–140, Amer. Inst. Physics.
- [6] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm", *J. R. Statist. Soc. B*, vol. 39, pp. 1–38, 1977.
- [7] F. Hlawatsch, G. Matz, H. Kirchauer, and W. Kozek, "Time-frequency formulation, design, and implementation of time varying optimal filters for signal estimation", *IEEE Trans. Signal Processing*, vol. 48, no. 5, pp. 1417–1432, May 2000.
- [8] A. J. Bell and T. J. Sejnowski, "An information maximization approach to blind separation and blind deconvolution", *Neural Computation*, vol. 7, no. 6, pp. 1129–1159, 1995.



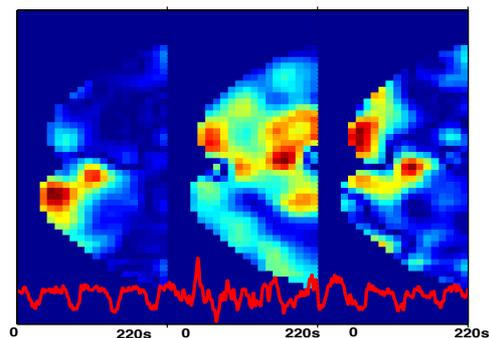
**Fig. 1.** Marginal partitioning of the Time Frequency domain: exploitation of the spectral non stationarity.



**Fig. 2.** The recovered sources with the Regularized Spectral EM algorithm for slice 10. The first and third sources correspond to the left and right visual cortex activations.



**Fig. 3.** The recovered time courses with the Regularized Spectral EM algorithm. Their temporal inter-delay is about 20 s corresponding to inter-delay between the alternative temporal stimulus presented to the subject.



**Fig. 4.** The recovered sources with the ICA InfoMax algorithm for the Slice 10. The alternative task related sources are not separated from the transient signal (the image in the middle of the Figure 2).