# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

**5,300**
Open access books available

**130,000**
International authors and editors

**155M**
Downloads

Our authors are among the

**154**
Countries delivered to

**TOP 1%**
most cited scientists

**12.2%**
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

**WEB OF SCIENCE**™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

## Interested in publishing with us?
Contact book.department@intechopen.com

# A Comparison of Non Negative Blind Source Separation Methods for Identifying Astrophysical Ice Compounds

Jorge Igual and Raul llinares
*Universidad Politecnica de Valencia*
*Spain*

## 1. Introduction

The analysis of astrophysical ices and the determination of the compounds that are present in the molecular clouds play a fundamental role in order to predict the future evolution of the cloud, e.g., its transformation to protostellar bodies or the appearance of new radicals and molecules. Because of the difficulties of obtaining satellite data, the process is simulated first in the laboratory generating ice analogs under controlled conditions. In this case, the ice mixture is carried on allocating the different components in the appropriate concentrations in a deposit chamber with a substrate and recording the spectrum of the aggregated ice when the chamber is filled through a gas inlet with the corresponding compounds (see Figure 1). This process tries to simulate the real process of forming ice mantles under the environmental conditions of the Interstellar Medium. The spectrum is obtained analyzing the transfer function of the ice when it is excited by a source beam with a known spectrum and measuring with a detector the output spectrum after crossing the ice.

The spectrum of each ice can be modelled as the linear instantaneous superposition of the spectrum of the different compounds, so a Source Separation approach is proper. We review and compare in this chapter a set of Source Separation algorithms that approach to the problem in different ways.

Initially, the problem can be addressed as a classical Blind Source Separation problem. In this case, nothing is assumed about the statistical distribution of the compounds present in the ices or how the mixture of these compounds is produced in the ice. The goal is the identification of the elements present in it after recovering the spectrum of the components that there exist in the ice analog. In addition, if we are able to obtain the demixing matrix, we will get a whole description of the mixing process, obtaining the abundances of every component in every ice. But the Blind Source Separation can be slightly modified if we introduce in the modelling of the problem some constraints based on physical properties of the spectrum and mixing process. The priors involve two characteristics: on the one hand, the non-negativeness of the spectrum and the abundances; on the other hand, the sparseness of the spectrum and sometimes the sparseness of the mixing matrix, depending on the kind of astrophysical ice. We will review these approaches and will present the family of algorithms that can be obtained when this information is exploited.
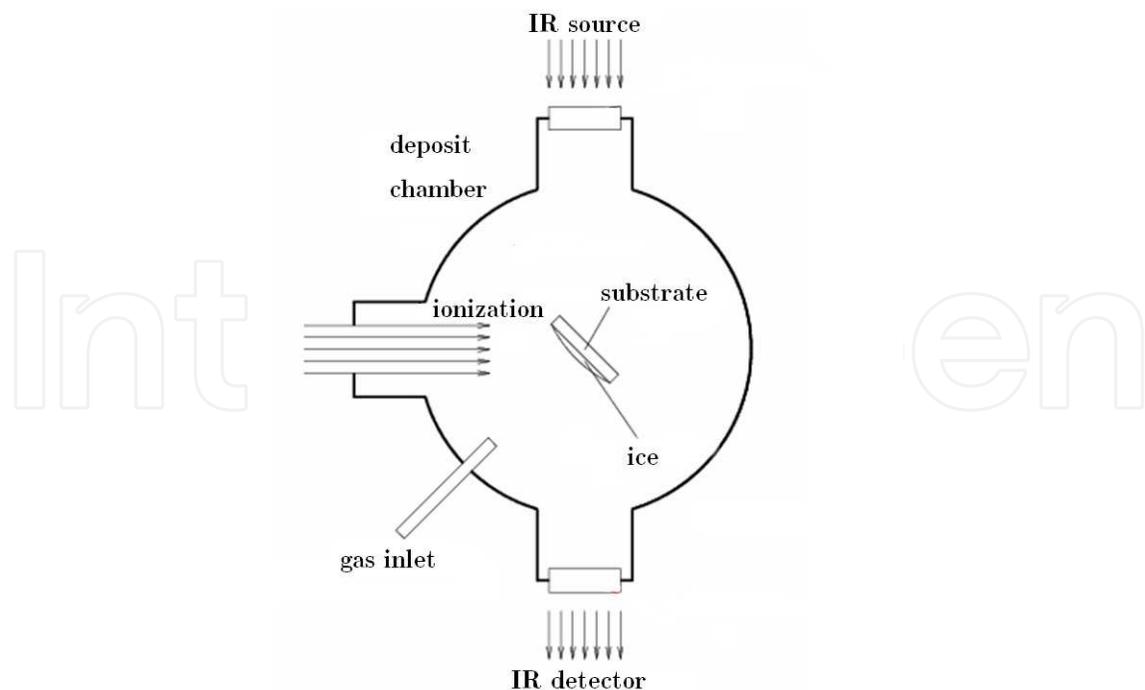
Fig. 1. Obtention of ices in the laboratory.

Blind Source Separation consists of the recovering of some independent sources up to a permutation, scaling and sign factor starting from instantaneous linear mixtures of them with the only assumption of the independence of the sources. The formulation is:

$$\mathbf{x} = \mathbf{As} \tag{1}$$

where $\mathbf{x}$ is the observed vector, $\mathbf{s}$ the source vector and $\mathbf{A}$ the mixing matrix.

To solve the Blind Source Separation problem, the statistical technique applied is the Independence Component Analysis (Oja et al., 2001; Cardoso, 1998 ; Comon, 1994), i.e., the generative model that try to decompose the data as a linear combination of statistically independent random variables. Traditional solutions are based on the minimization-maximization of a function that measures the statistical independence of the recovered signals. As we do not know anything about the distributions of the sources, a complete statistical analysis of the problem is difficult and it is not possible to guarantee exactly the independence of the recovered signals. In order to approximate the statistical independence of the sources, many approaches have been proposed. All of them use higher order statistics or nonlinear functions in order to approximate the independence criterion. In all these methods, no prior knowledge is available. However, in many applications, more information can be included in the model.

There are also different ways to modify the Blind Source Separation formulation to introduce this additional information. The best theoretical way to do it is to state the Bayesian formulation of the problem, where the prior information is combined with the likelihood function to obtain the posterior distribution. This approach has the advantage that it is statistically well grounded, since we do not obtain a point estimate but a distribution. But in practice it has some drawbacks. The most important is that it is difficult to implement even in the case when the distributions are approximated with some known

expressions. We will implement it through the variational Bayesian approach. If we are interested in algorithms that do not require so high computational resources, another statement is necessary. The mathematical formulation is still based on the fact that the spectrum of the sources are decoupled, but some priors are imposed in the Independent Component Analysis solutions. The non-negativeness is enforced modifying classical Blind Source Separation solutions obtaining algorithms where some nonlinear function such as a threshold function (e.g., the unit step function) are used to assure that there are not negative values in the estimated spectra. The idea is that the sources must be projected into the space of feasible solution, i.e., the sources can not have negative values. But there is still another option to address the problem including the positive restriction. In this case, the independence assumption is relaxed and substituted by a matrix factorization where the observations are the product of two matrices; the restriction in this case is related to the sign of the values of these two matrices: they can not have negative entries. This factorization is referred to as Non Negative Matrix Factorization (Lee & Seung, 1999). Its main advantages are that it is relatively easy to obtain fast algorithms and to modify them in order to include some additional restrictions, such as the non-negativeness of the mixing matrix (the abundances or concentrations of the compounds that there exist in an ice). The problem is that the numerical analysis of these algorithms reveals problems in convergence issues; nevertheless, as we will see in the Results section, they work pretty well in the case of our application.

## 2. Laboratory simulations of astrophysical ice mixtures

The signature of each ice is its infrared absorption spectrum, where the absorption bands correspond to the specific vibrational mode of the molecules, each one with different atoms and bonds. The spectrum is relevant because we know that the frequencies corresponding to the middle infrared spectrum (4000-400 cm-1; 2.5-25 μm) span the same range as the vibrational frequencies of the adjacent atoms in molecules associated with the most common species. This is why the infrared spectroscopy is an adequate tool to detect the compounds combined in different concentrations in an ice. Specially, this technique is appropriate to detect substances formed by molecules that have dipolar moment, e.g., $CO$, $CO_2$, $H_2O$, $NH_3$ and $CH_4$. This dipolar moment can be permanent or induced, due to the presence of other molecules. The detection of the existence of methane in water or in beaches after fuel waste or other environmental disasters is a typical example of the importance of the detection of these kinds of molecules. Each molecule (called endmember) has its own signature, i.e., an unique spectrum. The infrared spectrum has indeed another interesting property. It not only allows to know which molecules are present; in addition, it provides other kind of information, as state of the sample (different spectra for the gas and the ice phase of the molecule), and specifically the interatomic bonds of the molecules.

The molecules are combined in different quantities, called concentrations or abundances depending on if we talk about percentages or absolute quantities. They are named "ice mixtures" or to abbreviate "ices" when they are found in the frozen state. The data are the different spectra recorded by a satellite, but due to the high cost of the equipment and the problems inherent to satellite recordings such as the time slot, focusing in the proper direction, sensors or noise, a lot of preprocessing must be carried out before the data are obtained and after that before they are ready to be used.

A solution is to simulate the real ISM processes in the laboratory. In the laboratory, we are interested in the analysis of the composition of different ices and their behaviour in specific conditions that reproduce situations in which these substances are found outside of the laboratory. For example: to raise the temperature gradually or to radiate the samples with ions that are similar to the ones originated outside of the planet Earth. This is the reason because in the deposit chamber used to obtain the ice (see Fig. 1) in the laboratory there is an entrance for the ion beam. In this way, we can obtain in the laboratory molecules of interest as carbonic acid ($H_2CO_3$) from simple molecules as CO by protons implantation in the same way that is supposed it happens in the cloud. The measured infrared spectrum is formed by absorption bands around some specific wavelengths determined by the atomic composition and bond structure of the ice; their peak position and width depend on the presence or absence of some molecules that can affect to their dipolar moment, in addition to temperature and particle shape. Besides, these bands usually have an area and a width related to the compound abundance in the ice.

## 3. Signal processing of laboratory ices

The analysis of laboratory simulations of astrophysical ice mixtures consists of the study of the spectra of the ices in order to establish the compounds present in them. The measured spectrum corresponds to the superposition of the different spectra of the molecules present in the ice; i.e., it describes the absorption features of the ice mixture as a linear combination of the features of the different compounds, e.g., the 2140 $cm^{-1}$ C≡O stretching band. The basic model corresponds to the linear instantaneous mixture model of the Blind Source Separation problem (1).

The infrared absorption spectrum $x_{ij}$ measured for $i = 1,...,M$ ice mixtures in the spectral band $j = 1,...,N$, typically corresponding to 4000 up to 400 $cm^{-1}$ with resolutions 1 or 2 $cm^{-1}$, is the linear combination of the independent absorption spectra $s_{kj}$ of the molecules (sources) $k = 1,...,K$ present in those ices. The concentration of molecule $k$ in ice $i$ is the mixing matrix entry $a_{ik}$. The concentrations $a_{ik}$ and absorption spectra $s_{kj}$ are non-negative, although some preprocessing tasks such as baseline removal, noise and complex physics in the measurement process can produce a negative $x_{ij}$. All these processes will be resumed in a noise term $n_{ij}$ for ice $i$ in wavelength $j$. In this case, we obtain the noisy instantaneous mixture Blind Source Separation model:

$$\mathbf{x} = \mathbf{As} + \mathbf{n} \tag{2}$$

The *Mx1* data vector $\mathbf{x}$ (measured ices) is modeled as the linear instantaneous combination of a *Kx1* source vector $\mathbf{s}$ (spectrum of the compounds), where $\mathbf{n}$ is an additive noise representing the error in the measurements and the goal is to recover the independent components of the source vector and/or estimate the *MxK* mixing matrix $\mathbf{A}$ (abundances of each molecule in every ice). In the case of *M=K*, we have the square problem and the mixing matrix can be inverted obtaining the demixing matrix; remember that the product of the mixing and demixing matrices is a permuted diagonal matrix, i.e., a matrix with one and only one non zero value in every row and column that corresponds to the sign, order and amplitude indetermination of the problem. The noise term is usually considered a centered, white and Gaussian random vector with a given diagonal covariance matrix $\mathbf{R_n}$.

### 3.1 Bayesian source separation

Given the set of observations $\mathbf{x}_j, j = 1 \ldots N$, the posterior distribution of the unknown interesting variables $\mathbf{A}, \mathbf{s}_j, j = 1 \ldots N$ of our generative model is, applying Bayes' Theorem:

$$p(\mathbf{A}, \mathbf{s}_j \mid \mathbf{x}_j) = p(\mathbf{A}, \mathbf{s}_j) \cdot \frac{p(\mathbf{x}_j \mid \mathbf{A}, \mathbf{s}_j)}{p(\mathbf{x}_j)} \tag{3}$$

where $p(\mathbf{A}, \mathbf{s}_j)$ is called the prior probability, $p(\mathbf{x}_j \mid \mathbf{A}, \mathbf{s}_j)$ the likelihood and $p(\mathbf{x}_j)$ the marginal likelihood or evidence.

The prior describes our knowledge before we obtain the observations, i.e., it encodes in advance the possible values of the model parameters, e.g., if we know before any observation that the sources follow a known distribution; the likelihood describes the goodness of the model, i.e., the probability that the observations follow the model; the evidence, in our case, because we assume the Blind Source Separation model, is just a normalization factor that can be dropped in the optimization step. Shortly, Bayes rule updates the prior after a new data is observed, obtaining the posterior probability. Because the spectrum of the molecules and the concentrations in the ices are independent, we can factorize the prior in two terms, obtaining the posterior:

$$p(\mathbf{A}, \mathbf{s}_{1 \ldots N} \mid \mathbf{x}_{1 \ldots N}) \propto p(\mathbf{A}) \cdot p(\mathbf{s}_{1 \ldots N}) \cdot p(\mathbf{x}_{1 \ldots N} \mid \mathbf{A}, \mathbf{s}_{1 \ldots N}) \tag{4}$$

Bayesian inference with the posterior density is intractable in a general framework. The first step consists of approximating the distributions using some parametric distributions. In our case, the prior distribution for the sources and the mixing matrix corresponds to a mixture of Gaussians with hyperparameters the mixing proportions, the mean and the variance value of the Gaussians (Igual & Llinares, 2007). This mixture model can be optimized attending to the non-negativeness of the spectrum. In this case, the distribution used is the truncated or positive Gaussian, i.e. a distribution with zero probability of taking negative values and two times the probability of the Gaussian for positive values.

The variational approach is a good solution to obtain an algorithm that can implement the parameterized posterior distribution. The objective function to be maximised is the negative free energy $E$ (Choudrey & Roberts, 2001):

$$E = \left\langle \log p(\mathbf{x}_{1 \ldots N}, \boldsymbol{\Theta}) \right\rangle_{\hat{p}(\boldsymbol{\Theta})} + H(\hat{p}(\boldsymbol{\Theta})) \tag{5}$$

where $\boldsymbol{\Theta}$ is the set of all parameters: the mixing matrix, the noise covariance matrix, the variance of the mixing matrix entries, the sources, the mixing proportions, mean and variances of the mixture of Gaussian model used for the molecular ice distribution (sources); $\hat{p}(\boldsymbol{\Theta})$ is the approximating posterior. The first term in (5) is the expectation of the joint density with respect $\hat{p}(\boldsymbol{\Theta})$; the second one is the entropy. Maximizing (5) is equivalent to minimizing the Kullback-Leibler distance between the true and the approximating posteriors. The approximated pdf $\hat{p}(\boldsymbol{\Theta})$ is chosen such that it can be factorized over the set $\boldsymbol{\Theta}$. Therefore, the maximization can be done individually with coupled terms. The factorization we use is:

$$\hat{p}(\boldsymbol{\Theta}) = \hat{p}(\mathbf{A})\hat{p}(\mathbf{R})\hat{p}(\mathbf{s}_{1\ldots N})\hat{p}(\boldsymbol{\pi})\hat{p}(\sigma^2) \qquad (6)$$

where $\left\{\boldsymbol{\pi}, \sigma^2\right\}$ are the vectors defined by the mixing proportions and variances of the different mixture of truncated Gaussians used to model the source distribution. Their distribution is a product of symmetric Dirichlets for the mixing proportions and a product of inverted Gammas for the variance. The prior used for the noise variance is also a product of inverted Gammas. Maximizing (5) yields updating rules for the parameters of the posteriors. The parameters of the posteriors are updated versions of the priors. Unfortunately, the update equations are coupled and the implementation of the algorithm requires iterations; these are carried out starting with some initial value for the variables and iterating until convergence.

### 3.2 Non negative matrix factorization

Since the Bayesian approach involves the use of statistical distributions, it allows a physically well grounded exposition of the solution, at least satisfying the physical restrictions of the problem, in particular, the non-negativeness of the absorption spectrum and abundances. However, it has the drawback that in some point some approximations must be done in order to obtain an implementation of the algorithm. The good thing is that we work with distributions, so we can infer not only a point estimate, but a posterior distribution. This is the classical advantage of Bayesian approach. But there are more ways for obtaining an algorithm that enforce the physical restrictions.

One option is extending the Blind Source Separation model adding constraints about the non-negativeness. Another related approach is to model the observations as the product of two non-negative matrix, the so called Non Negative Matrix Factorization (Lee & Seung, 1999). It is closely related to Independent Component Analysis (Cichocki et al., 2006). Although it was not motivated originally by a Bayesian framework, it has been shown that both of them are the same under mild assumptions about the distributions of the sources (Igual & Llinares, 2008).

The NMF statement of the problem is (in matricial form): the infrared absorption spectrum $\mathbf{X}$ of the ices is factorized as the product of the matrix of concentrations $\mathbf{A}$ and the matrix of absorption spectra $\mathbf{S}$ of the molecules present in those ices:

$$\mathbf{X} = \mathbf{AS}, \mathbf{A} \geq 0, \mathbf{S} \geq 0 \qquad (7)$$

Note that in this case we include in the formulation of the problem the whole matrices, i.e., we do not assume any generative model nor time series approach (in our case "time" corresponds to "wavelength"). In real measurements, as it was explained before, an additive noise term must be considered. Although in the laboratory the conditions are well controlled (a high signal to noise ratio), a perfect reconstruction such as (7) is not possible. For example, we experience difficulties to obtain pure mixtures of the different compounds and to control the temperature and pressure conditions in the vacuum chamber where the ice is aggregated. Therefore, in order to use a more complete model, we will assume the noisy model as we did in the Blind Source Separation in order to take into account all these sources of noise. It reads:

$$\mathbf{X} = \mathbf{AS} + \mathbf{V}, \quad \mathbf{A} \geq 0, \mathbf{S} \geq 0 \tag{8}$$

where $\mathbf{V} \in \Re^{M \times N}$ is the noise matrix. Thanks to this term, we will be able to explain measured spectra with negative absorptions, i.e., we will suppose that the negative values of our data are due to the noise vector. In addition, adding the noise term will also allow to understand (8) as a probabilistic generative model connected to (2) although this was not the original motivation of the Non Negative Matrix Factorization.

Assuming in the model the term $\mathbf{V}$ as an i.i.d. Gaussian noise, the maximum likelihood estimates of the matrices $\mathbf{A}$ and $\mathbf{S}$ are:

$$(\mathbf{A}^{\mathrm{ML}}, \mathbf{S}^{\mathrm{ML}}) = \underset{\mathbf{A}, \mathbf{S}}{\arg\min}(-\log \mathrm{p}(\mathbf{X} / \mathbf{A}, \mathbf{S})) = \underset{\mathbf{A}, \mathbf{S}}{\arg\min} \mid\mid \mathbf{X} - \mathbf{AS} \mid\mid^2, \mathbf{A} \geq 0, \mathbf{S} \geq 0 \tag{9}$$

Therefore, for a Gaussian assumption, the cost function to be minimized results rather intuitive: the ML estimate tries to minimize the squared Frobenius norm between the spectra of the ice mixtures and the factorization.

The original multiplicative version of the algorithm is:

$$\mathbf{A}_{i,k} \leftarrow \mathbf{A}_{i,k} \frac{(\mathbf{XS}^{\mathrm{T}})_{i,k}}{(\mathbf{ASS}^{\mathrm{T}})_{i,k}}$$

$$\mathbf{S}_{k,j} \leftarrow \mathbf{S}_{k,j} \frac{(\mathbf{A}^{\mathrm{T}}\mathbf{X})_{k,j}}{(\mathbf{A}^{\mathrm{T}}\mathbf{AS})_{k,j}} \tag{10}$$

If we know that the spectra of the ices do not include negative values, the multiplicative version guarantees that the matrix of abundances and the spectra of the compounds are positive just initializing the algorithm with non negative matrices. However, if there exist negative values due to the noise term, the non-negativeness can not be guaranteed since all the spectral values of the ices are used in the updating step. If we do not take care of it, we can obtain for example compounds with negative absorptions, something physically impossible. When it happens, the rule is modified constraining the algorithm to enforce $\mathbf{A} \geq 0, \mathbf{S} \geq 0$ simply converting the negative or zero values of $\mathbf{A}$ and $\mathbf{S}$ to small positive numbers.

### 3.3 Sparseness of the spectrum

The algorithms explained in the previous section are based on the superposition of non negative signals. In the case of our application, and considering the typical molecules involved in the formation of ices, e.g., CO and $CO_2$, there exist another hypothesis that can be introduced in the statement of the problem: the sparseness of the signals, i.e., their value is zero excepting in some few wavelengths determined by the atomic composition and bond structure of the corresponding compound.

The consequence from a statistical signal processing point of view is that the spectra are characterized by deep narrow absorption peaks around some specific wavelengths, so they can be modeled from a flexible distribution with several parameters, such as a mixture of

Gaussians, to a simple supergaussian distribution such as the Laplacian one, with only one parameter. In the case of a mixture of Gaussians, it reads:

$$p(\mathbf{s}_j) = \prod_{k=1}^{K} p(\mathbf{S}_{k,j}) = \prod_{k=1}^{K} \sum_{q_k=1}^{Q_k} \pi_{k,q_k} \frac{1}{\sqrt{2\pi\sigma_{k,q_k}^2}} \exp\left[-\frac{(\mathbf{S}_{k,j} - \mu_{k,q_k})^2}{2\sigma_{k,q_k}^2}\right] \tag{11}$$

where the pdf of the *k*-source at the wavelength $j$ corresponds to a mixture of $Q_k$ Gaussians with mixing proportions $\pi_{k,q_k}$, $\sum_{q_k=1}^{Q_k} \pi_{k,q_k} = 1, k = 1,\ldots,K$, and hyperparameters the mean $\mu_{k,q_k}$ and the variance $\sigma_{k,q_k}^2$ for each Gaussian.

Note that this information is easily included in the Bayesian approach, since we can control the sparseness simply adjusting the corresponding parameters. In the case of the Non Negative Matrix Factorization, since we are not using the distributions, the sparseness must be incorporated in a different way. We have to modify the cost function, i.e., the squared Frobenius norm between the spectra of the ice mixtures and the factorization, to add a constraint in the optimization procedure that enforces the sparseness hypothesis.

The new cost function to be minimized is:

$$D(\mathbf{X} \mid\mid \mathbf{AS}) = \frac{1}{2}\sum_{i,j}(\mathbf{X}_{i,j} - (\mathbf{AS})_{i,j})^2 + \gamma J(\mathbf{S}) \quad \gamma, \mathbf{A} \geq 0, \mathbf{S} \geq 0 \tag{12}$$

where $\gamma$ is a regularization parameter. With respect to the regularization function $J(\mathbf{S})$, to enforce the sparseness, it is defined such as $J(\mathbf{S}) = \sum_{k,j} \mathbf{S}_{k,j}$, that is equivalent to model the source with a Laplacian distribution. We can calculate the derivatives and use a gradient descent algorithm that minimizes (12). We call this algorithm the regularized Non Negative Matrix Factorization estimate. The updating rules are:

$$\mathbf{A}_{i,k} \leftarrow \mathbf{A}_{i,k} \frac{(\mathbf{XS}^T)_{i,k}}{(\mathbf{ASS}^T)_{i,k}},$$

$$\mathbf{A}_{i,k} \leftarrow \frac{\mathbf{A}_{i,k}}{\sum_{k=1}^{K} \mathbf{A}_{i,k}} \tag{13}$$

$$\mathbf{S}_{k,j} \leftarrow \mathbf{S}_{k,j} \frac{(\mathbf{A}^T\mathbf{X})_{k,j} - \gamma}{(\mathbf{A}^T\mathbf{AS})_{k,j}}$$

In order to prevent that $\mathbf{A}$ and $\mathbf{S}$ can take negative values, in every iteration we apply the nonlinear function $\mathbf{A}_{i,k} = \max(\mathbf{A}_{i,k}, \varepsilon), \mathbf{S}_{k,j} = \max(\mathbf{S}_{k,j}, \varepsilon)$, where $\varepsilon$ is a very small value, i.e., the abundance and spectra are projected to the subspace of possible values as we did in previous sectionand it is typical to all implementations of NMF algorithms to prevent numerical problems.

## 4. Results

The data corresponds to the public ice analogs database of the University of Leiden. This database contains the infrared spectra of laboratory analogs of interstellar ices. Different mixtures of molecules (from one up to three components selected from $H_2$, $H_2O$, $NH_3$, $CH_4$, CO, $H_2CO$, $CH_3OH$, HCOOH, $O_2$, $N_2$ and $CO_2$) at different temperatures and UV radiation exposures were produced, the final spectrum being calculated rationing the measured and the background spectrum. The units of the data are absorbance and $cm^{-1}$. Figure 2 shows the spectra of the pure ices (molecules) used in the experiments.

The baseline was removed with Origin software and the useful wavelengths intervals were selected. Among all the preprocessing tasks, the most important one is the baseline removal, because in much attenuated absorption bands, a bad approximation of the baseline can mask some compounds or produce negative values of the optical depth.

The measure of performance used in this section is the Signal-to-Interference Ratio (SIR) defined as:

$$SIR_i = 10\log\frac{E\left[S_i^2\right]}{E\left[\left(S_i - \hat{S}_i\right)^2\right]} \tag{14}$$

where $S_i$ represents the original source $i$ and $\hat{S}_i$ the corresponding recovered source, normalized to the same power. In this index, a high value means high quality results.

The spectra of the ices were mixed with a 10x5 uniform random positive mixing matrix in the range 0-1 in order to simulate nine real laboratory ices. Figure 3 shows an example of such a mixture. The first experiment involved the recovery of the five ices using the algorithms explained in previous sections: NMF (Non Negative Matrix Factorization - Eq. 10), RNMF (Regularized Non Negative Matrix Factorization - Eq. 13) and VBICA (Bayesian Source Separation). Figures 4 to 6 represent the ices recovered by the three algorithms.

Remember that the sources can be recovered in any order. The separation of CO and $H_2O$ obtained by NMF algorithm is almost perfect. For the other sources, there is some remaining of the other molecules. The algorithm was not able to cancel in the rest of compounds the peak around 2300 $cm^{-1}$ due to CO. The same problem occurs for the water, which is contaminating the spectrum of the estimated $CO_2$. In addition, the algorithm fails absolutely in the extraction of the $CH_3OH$, obtaining two HCOOH (third and fourth signals in Figure 4).

In the case of RNMF, the results improve. Only for the water (the fourth recovered source in the Figure 5) it is still visible the same noisy peak than in the case of the NMF algorithm: the very narrow absorption band around 2300 $cm^{-1}$ due to CO. For the rest of compounds, the separation is excellent.

VBICA was able to recover all the sources but the $CH_3OH$ (fourth source in Figure 6). It was not able to estimate correctly the right side of the spectrum of the $CH_3OH$. However, for the other sources, the algorithm worked very well, as we can see in the same Figure comparing with Figure 2.
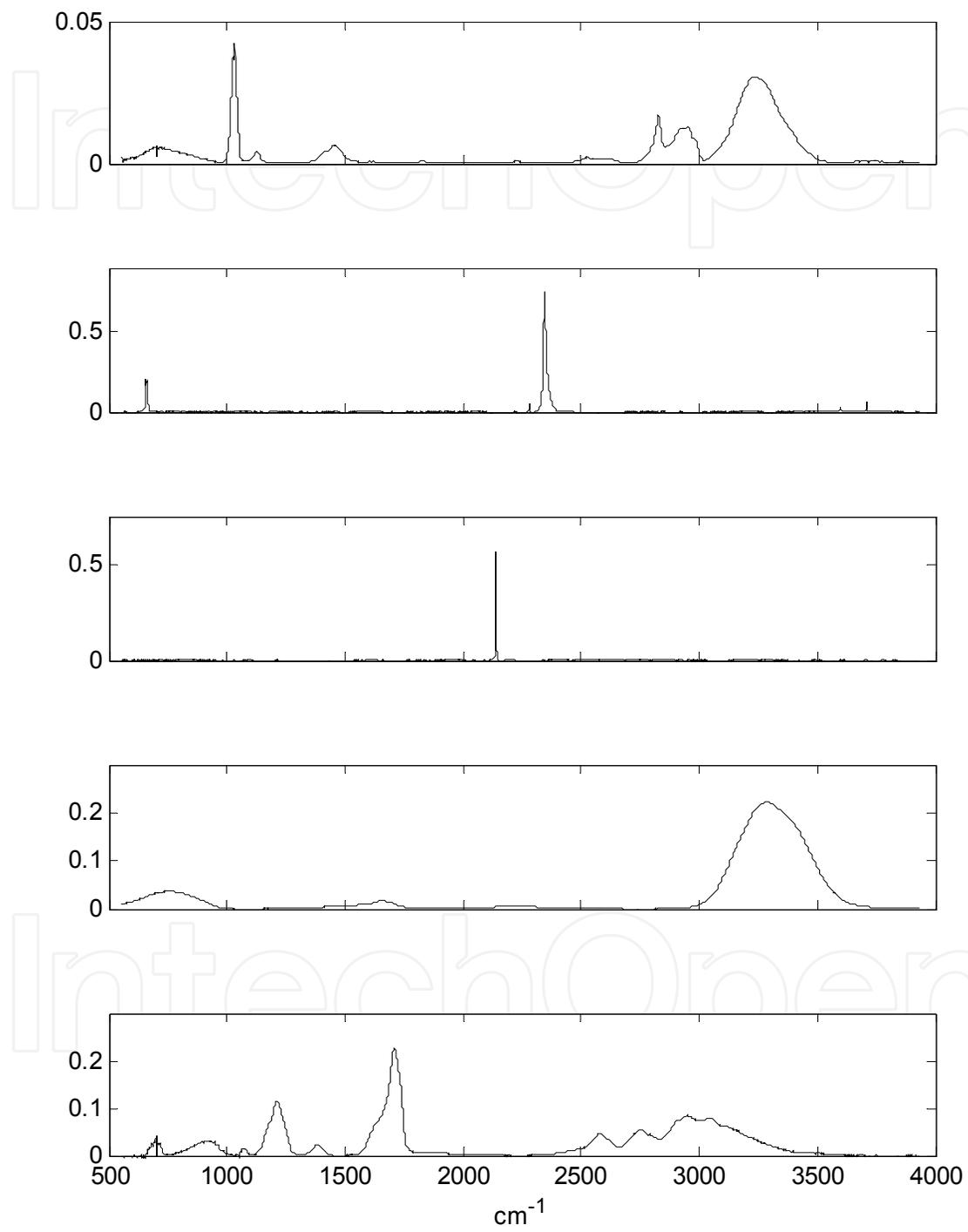
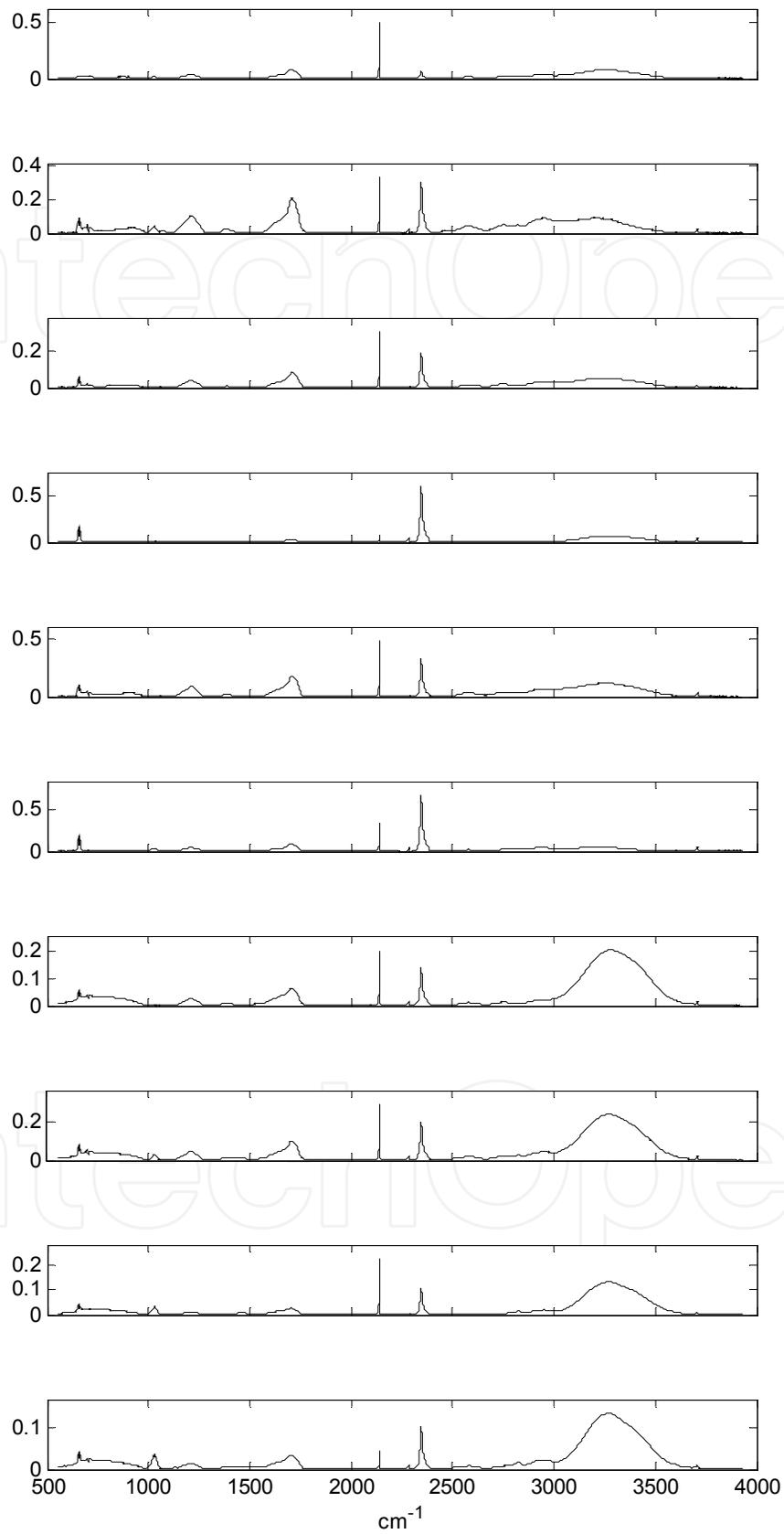Fig. 2. Spectrum of the pure ices: CH₃OH, CO, CO₂, H₂O and HCOOH.

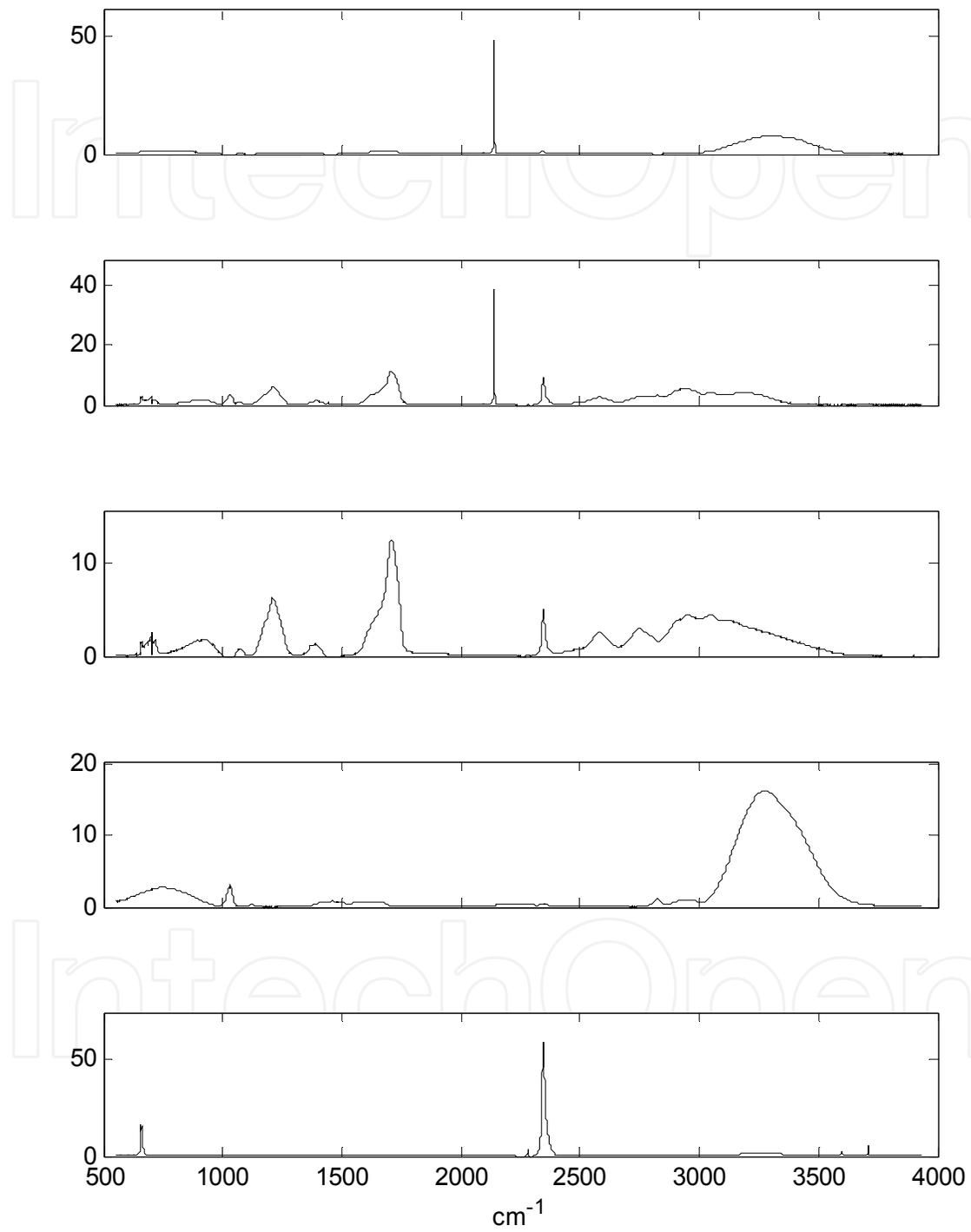Fig. 3. A mixture of ices: CH$_3$OH, CO, CO$_2$, H$_2$O and HCOOH.

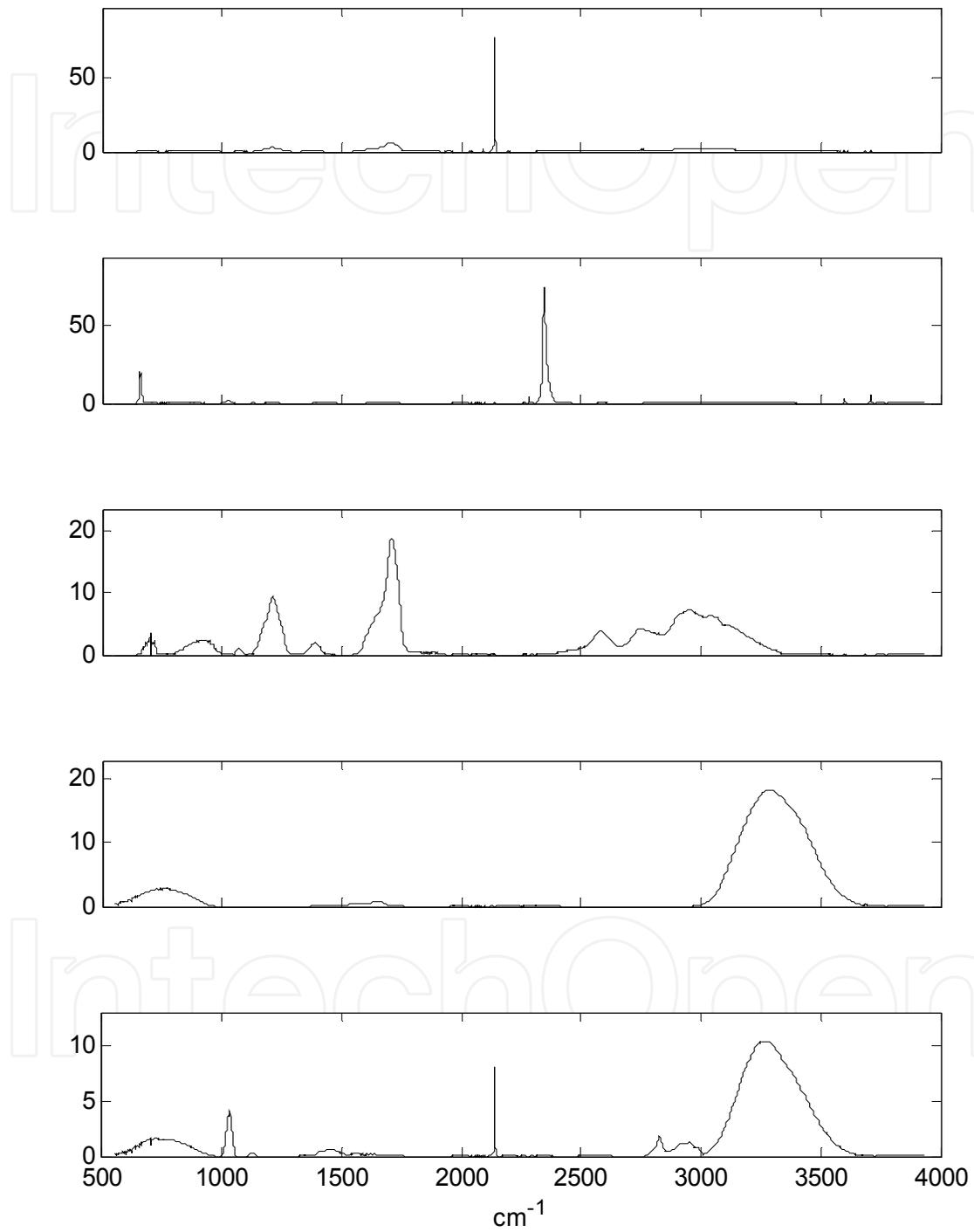Fig. 4. Ices recovered by NMF. $CO_2$, HCOOH, HCOOH, $H_2O$, CO.

Fig. 5. Ices recovered by RNMF: $CO_2$, $CO$, $HCOOH$, $H_2O$, $CH_3OH$.
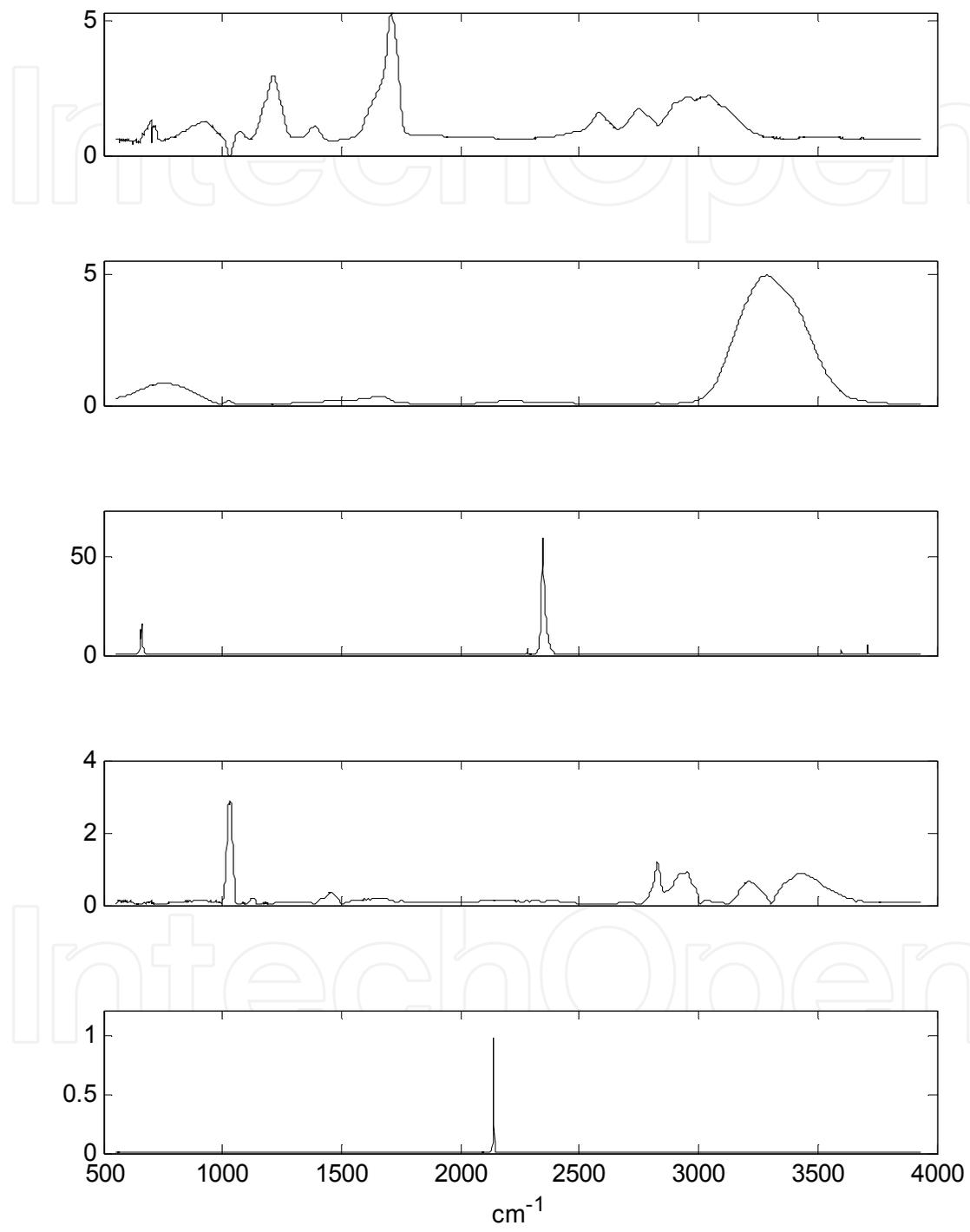
Fig. 6. Ices recovered by VBICA: HCOOH, H₂O, CO, CH₃OH and CO₂.

## 5. Conclusion

We have reviewed an application of Signal Processing techniques in Astrophysics: the study and analysis of ices obtained in the laboratories that try to simulate the astrophysical ices found in the outer space. We have exposed the problem from a Signal Processing point of view, showing the goodness of the Blind Source Separation approach to model the recorded absorption spectra as the combination of the spectra of the molecules that there exist in the obtained ice by aggregation of the molecules.

The problem can be addressed in a Bayesian framework, where the prior distributions about the spectrum of the compounds and the abundances of them in every ice can be treated in a proper way when they are approximated by models such as mixture of Gaussians. But these algorithms are very slow when compared to classical optimization methods based on a cost function that is maximized according to some gradient based algorithm. We have shown that the statistical approach can be relaxed when the Non Negative Matrix Factorization is introduced. The obtained algorithms are faster although they can suffer from convergence problems due to the initialization and adjustment of some parameters that control the performance of the algorithms.

Both procedures work, although attending to all the factors, we can conclude that the algorithms that enforce the non-negativeness condition in the optimization step obtain the best results, since they are fast and are able to cancel the remaining spectrum of the compounds demixing the spectrum of the ices. Nevertheless, the extraction is not always perfect, especially when some large narrow peaks are involved. On the other hand, when the sparseness restriction is included, the results improve, as we have seen with the RNMF algorithm, which obtained similar results to VBICA with a lower computational cost.

## 6. Acknowledgment

## 7. References

Cardoso, J.F.; Blind Signal Separation: statistical principles. *Proc.of the IEEE*, Vol. 86, No. 10, (1998), pp. 2009-2025

Choudrey, R.; Roberts, S.J.; Flexible Bayesian independent component analysis for blind source separation. *Proceedings of 3rd International Conference on Independent Component Analysis and Blind Signal Separation*, pp. 90-95, San Diego, USA, 2001

Cichocki, A.; Zdunek, R.; Amari, S.; New algorithms for non-negative matrix factorization in applications to blind source separation. *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing ICASSP*, 2006

Comon, P.; Independent Component Analysis, a new concept? *Signal Processing*, Vol. 36, (1994), pp. 287-314.

Igual, J.; Llinares, R.; An informed source separation of astrophysical ice analogs. Digital Signal Procesing, Vol. 17 (2007), pp. 947–964, ISSN 1051-2004.

Igual, J. ; Llinares, R.; Nonnegative Matrix Factorization of Laboratory Astrophysical Ice Mixtures. IEEE Journal of Selected Topics in Signal Processing, Vol. 2, pp. 697 – 706, ISSN 1932-4553.

Lee, D.D.; Seung, S.; Learning the parts of objects by non-negative matrix factorization. Nature, Vol. 401, No. 6755 (1999), pp. 788-791.

Oja, E.; Hyvarinen, A.; Karhunen, J. (2001). Independent Component Analysis. John Wiley & Sons.
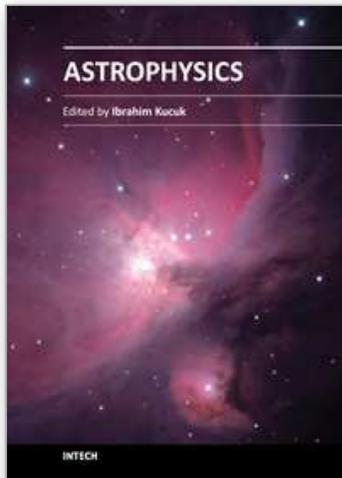
This book provides readers with a clear progress to theoretical and observational astrophysics. It is not surprising that astrophysics is continually growing because very sophisticated telescopes are being developed and they bring the universe closer and make it accessible. Astrophysics Book presents a unique opportunity for readers to demonstrate processes do occur in Nature. The unique feature of this book is to cover different aspects in astrophysics covering the topics: • Astronomy • Theoretical Astrophysics • Observational Astrophysics • Cosmology • The Solar System • Stars • Planets • Galaxies • Observation • Spectroscopy • Dark Matter • Neutron Stars • High Energy Astrophysics

**How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Jorge Igual and Raul llinares (2012). A Comparison of Non Negative Blind Source Separation Methods for Identifying Astrophysical Ice Compounds, Astrophysics, Prof. Ibrahim Kucuk (Ed.), ISBN: 978-953-51-0473-5, InTech, Available from: http://www.intechopen.com/books/astrophysics/a-comparison-of-non-negative-blind-source-separation-methods-for-identifying-astrophysical-ice-compo

# INTECH
open science | open minds