



Blind unmixing of linear mixtures using a hierarchical Bayesian model. Application to spectroscopic signal analysis

Nicolas Dobigeon, Jean-Yves Tournet, Saïd Moussaoui

► To cite this version:

Nicolas Dobigeon, Jean-Yves Tournet, Saïd Moussaoui. Blind unmixing of linear mixtures using a hierarchical Bayesian model. Application to spectroscopic signal analysis. IEEE Workshop on Statistical Signal Processing, Aug 2007, Madison, United States. pp.CDRom, 2007, <10.1109/SSP.2007.4301222>. <hal-00455587>

HAL Id: hal-00455587

<https://hal.archives-ouvertes.fr/hal-00455587>

Submitted on 16 Feb 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

BLIND UNMIXING OF LINEAR MIXTURES USING A HIERARCHICAL BAYESIAN MODEL. APPLICATION TO SPECTROSCOPIC SIGNAL ANALYSIS

Nicolas Dobigeon, Jean-Yves Tournéret

IRIT/ENSEEIH/TéSA
2 rue Charles Camichel, BP 7122
31071 Toulouse cedex 7, France
{dobigeon, tournéret}@enseeiht.fr

Said Moussaoui*

IRCCyN - CNRS UMR 6597
ECN, 1 rue de la Noë, BP 92101
44321 Nantes cedex 3, France
said.moussaoui@irccyn.ec-nantes.fr

ABSTRACT

This paper addresses the problem of spectral unmixing when positivity and additivity constraints are imposed on the mixing coefficients. A hierarchical Bayesian model is introduced to satisfy these two constraints. A Gibbs sampler is then proposed to generate samples distributed according to the posterior distribution of the unknown parameters associated to this Bayesian model. Simulation results conducted with synthetic data illustrate the performance of the proposed algorithm. The accuracy of this approach is also illustrated by unmixing spectra resulting from a multicomponent chemical mixture analysis by infrared spectroscopy.

Index Terms— Spectral unmixing, non-negativity, additivity, Bayesian inference, Monte Carlo methods.

1. INTRODUCTION

Linear regression models are widely used in signal processing applications. The popularity of linear regression may be justified by the variety of existing algorithms allowing one to estimate the unknown model parameters. However, the estimation problem is more complicated when the mixing coefficients or/and the regressors have to satisfy specific constraints. Constraints which have recently received much attention in the literature include monotony [1], positivity [2] or sparsity [3]. This paper studies linear regression models where the mixing coefficients satisfy positivity and additivity constraints. These constraints are for example required for the analysis of spectral mixture data, i.e. in multicomponent material analysis by spectroscopy [2] or hyperspectral imaging [4].

The estimation of constrained linear regression parameters can be achieved by solving constrained optimization problems. One typically defines an appropriate cost function such as the least squares or weighted least squares criteria and one minimizes this cost function by ensuring that the constraints

are satisfied. This strategy has been for instance used successfully for the unmixing of hyperspectral images yielding the well known fully constrained least squares algorithm [5]. However, Bayesian estimators also offer nice alternatives for fitting constrained linear regression models. Prior distributions can be defined in order to satisfy the linear regression constraints. The model parameters are then classically estimated from their posterior distribution. The posterior distribution can also be used to provide information about the uncertainties of the parameter estimates such as confidence intervals. Bayesian estimators for linear regression problems with monotony, positivity or sparsity constraints have been recently studied in [1–3]. This paper studies Bayesian estimators for linear regression parameters with positivity and additivity constraints. The regression coefficients are assigned prior distributions defined on a simplex. A Metropolis-within-Gibbs algorithm is then studied to generate samples according to the full posterior of regressors and mixture coefficients. The proposed methodology is similar to the algorithm developed in [6]. However, it allows one to estimate the mixture spectra, contrary to the algorithm in [6] which assumed that the spectra were identified beforehand.

The paper is organized as follows. Section 3 defines a hierarchical Bayesian model (HBM) for the proposed blind unmixing problem. In particular, prior distributions concentrated on a simplex and satisfying the positivity and additivity constraints are introduced. Section 4 describes a Gibbs sampling strategy that allows one to overcome the computational complexity inherent to this HBM. Simulation results and conclusion are reported in Section 5 and 6.

2. PROBLEM STATEMENT

The linear mixture model is defined as:

$$y_{i,j} = \sum_{m=1}^M c_{i,m} s_{m,j} + e_{i,j}, \quad (1)$$

where $y_{i,j}$ is the observed spectrum at time index i ($i = 1, \dots, N$) in the j^{th} spectral band ($j = 1, \dots, L$), N is the

*This work is part of a young researcher project supported by the CNRS and GdR-ISIS.

number of samples, M is the number of mixture components and L is the number of spectral bands. The coefficient $c_{i,m}$ is the contribution of the m^{th} component in the i^{th} mixture and $e_{i,j}$ is an additive noise modeling measurement errors and model uncertainties. The noise sequences $\mathbf{e}_i = [e_{i,1}, \dots, e_{i,L}]^T$ are assumed to be independent and identically distributed (i.i.d.) according to zero-mean Gaussian distributions with variances $\sigma_{e,i}^2$. Standard matrix notations used for the N observations in the L spectral bands yield:

$$\mathbf{Y} = \mathbf{C}\mathbf{S} + \mathbf{E}, \quad (2)$$

where $\mathbf{Y} = [y_{i,j}]_{i,j} \in \mathbb{R}^{N \times L}$, $\mathbf{C} = [c_{i,m}]_{i,m} \in \mathbb{R}^{N \times M}$, $\mathbf{S} = [s_{m,j}]_{m,j} \in \mathbb{R}^{M \times L}$ and $\mathbf{E} = [e_{i,j}]_{i,j} \in \mathbb{R}^{N \times L}$. Due to obvious physical considerations, the concentrations $\mathbf{c}_i = [c_{i,1}, \dots, c_{i,M}]^T$ and the spectra $\mathbf{s}_m = [s_{m,1}, \dots, s_{m,L}]^T$, for $m = 1, \dots, M$, have to satisfy positivity constraints. Moreover, some applications such as kinetic reactions [7] require that the mixing coefficients are subjected to additivity constraints. The blind unmixing problem addressed in this paper consists of estimating jointly the concentration matrix \mathbf{C} and the spectral sources \mathbf{S} under the following positivity and additivity constraints:

$$\begin{cases} s_{m,j} \geq 0 \text{ and } c_{i,m} \geq 0, & \forall (i, m, j), \\ \sum_{m=1}^M c_{i,m} = 1 & \forall i. \end{cases} \quad (3)$$

A Bayesian model was proposed in [2] to perform the blind unmixing under positivity constraints on the source spectra and the mixing coefficients. This paper goes a step further by including the additivity constraint in the Bayesian model. Note that this constraint allows one to resolve the scale indeterminacy inherent to blind source separation problems.

3. HIERARCHICAL BAYESIAN MODEL

3.1. Likelihood

The statistical properties of the noise vector \mathbf{e}_i and the linear mixing model described in (1) allow one to write:

$$\mathbf{y}_i \sim \mathcal{N}(\mathbf{S}^T \mathbf{c}_i, \sigma_{e,i}^2 \mathbf{I}_L),$$

where $\mathbf{y}_i = [y_{i,1}, \dots, y_{i,L}]^T$, $\mathcal{N}(\cdot, \cdot)$ denotes the Gaussian distribution and \mathbf{I}_L is the $L \times L$ identity matrix. By assuming the independence between the vectors $\mathbf{e}_1, \dots, \mathbf{e}_i$, the likelihood of \mathbf{Y} is:

$$f(\mathbf{Y}|\mathbf{C}, \mathbf{S}, \sigma_e^2) \propto \frac{1}{\prod_{i=1}^N \sigma_{e,i}^L} \exp \left[\sum_{i=1}^N -\frac{\|\mathbf{y}_i - \mathbf{S}\mathbf{c}_i\|^2}{2\sigma_{e,i}^2} \right], \quad (4)$$

where $\sigma_e^2 = [\sigma_{e,1}^2, \dots, \sigma_{e,N}^2]^T$, $\|\mathbf{x}\|^2 = \mathbf{x}^T \mathbf{x}$ stands for the standard ℓ_2 norm and \propto means ‘‘proportional to.’’

3.2. Parameter priors

3.2.1. Concentration coefficients

Using the additivity constraints introduced in (3), the regression vectors can be written $\mathbf{c}_i = [\mathbf{a}_i^T, c_{i,M}]^T$ with $\mathbf{a}_i^T = [c_{i,1}, \dots, c_{i,M-1}]$ and $c_{i,M} = 1 - \sum_{m=1}^{M-1} c_{i,m}$. The natural priors for \mathbf{a}_i , $i = 1, \dots, N$, are uniform distributions on the following simplex¹:

$$\mathbb{S} = \left\{ \mathbf{a}_i \left| a_{i,m} \geq 0, \forall m = 1, \dots, M-1, \sum_{m=1}^{M-1} a_{i,m} \leq 1 \right. \right\}. \quad (5)$$

By assuming the *a priori* independence between the vectors \mathbf{a}_i , the prior distribution for the mixing matrix $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_N]^T$ reduces to:

$$f(\mathbf{A}) = \prod_{i=1}^N \mathbf{1}_{\mathbb{S}}(\mathbf{a}_i), \quad (6)$$

where $\mathbf{1}_{\mathbb{S}}(\cdot)$ denotes the indicator function defined on \mathbb{S} .

3.2.2. Spectral sources

Several distributions with positive support can be chosen as prior distribution for a spectral source, provided that the constraints in (3) are satisfied. However, for computation reasons, choosing a conjugate prior is judicious when no additional information is available. Therefore, our proposed model is initially based on a truncated Gaussian prior. In section 5, simulation results obtained with this Gaussian prior will be compared to results obtained with an exponential prior, which seems to be more appropriate to incorporate information like sparsity.

A truncated Gaussian prior distribution is assigned to each spectral source \mathbf{s}_m denoted as:

$$\mathbf{s}_m | \sigma_{s,m}^2 \sim \mathcal{N}^+(\mathbf{0}_L, \sigma_{s,m}^2 \mathbf{I}_L), \quad (7)$$

where $\mathbf{0}_L$ is the vector made of L zeros and $\mathcal{N}^+(\mathbf{u}, \mathbf{V})$ denotes the positive truncated Gaussian distribution with mean vector \mathbf{u} and covariance matrix \mathbf{V} . The probability density function (pdf) of this multivariate truncated distribution denoted $\phi^+(\cdot | \boldsymbol{\theta}, \boldsymbol{\Sigma})$ satisfies the relation:

$$\phi^+(\mathbf{x} | \boldsymbol{\theta}, \boldsymbol{\Sigma}) \propto \phi(\mathbf{x} | \boldsymbol{\theta}, \boldsymbol{\Sigma}) \mathbf{1}_{\mathbb{R}_+^L}(\mathbf{x}), \quad (8)$$

where $\phi(\cdot | \boldsymbol{\theta}, \boldsymbol{\Sigma})$ is the pdf of the usual multivariate Gaussian distribution defined on \mathbb{R}^L with mean vector $\boldsymbol{\theta}$ and covariance matrix $\boldsymbol{\Sigma}$. By assuming the independence between the spectral sources \mathbf{s}_m ($m = 1, \dots, M$), the prior distribution for \mathbf{S} can be written:

$$f(\mathbf{S} | \sigma_s^2) = \prod_{m=1}^M \phi^+(\mathbf{s}_m | \mathbf{0}_L, \sigma_{s,m}^2 \mathbf{I}_L), \quad (9)$$

with $\sigma_s^2 = [\sigma_{s,1}^2, \dots, \sigma_{s,M}^2]^T$.

¹Note that this choice is equivalent to choose Dirichlet distributions $\mathcal{D}_M(1, \dots, 1)$ as prior distributions for \mathbf{c}_i ($i = 1, \dots, N$).

3.2.3. Noise variances

Conjugate priors which are here inverse Gamma (IG) distributions are chosen for $\sigma_{e,i}^2$:

$$\sigma_{e,i}^2 | \rho_e, \psi_e \sim \mathcal{IG} \left(\frac{\rho_e}{2}, \frac{\psi_e}{2} \right), \quad (10)$$

where $\mathcal{IG} \left(\frac{\rho_e}{2}, \frac{\psi_e}{2} \right)$ denotes the inverse Gamma distribution with parameters $\frac{\rho_e}{2}$ and $\frac{\psi_e}{2}$. By assuming the independence between the noise variances $\sigma_{e,i}^2$, $i = 1, \dots, N$, the prior distribution of σ_e^2 is:

$$f(\sigma_e^2 | \rho_e, \psi_e) = \prod_{i=1}^N f(\sigma_{e,i}^2 | \rho_e, \psi_e). \quad (11)$$

The hyperparameter ρ_e will be fixed to $\rho_e = 2$ whereas ψ_e is an adjustable hyperparameter as in [8].

3.3. Hyperparameter priors

A conjugate IG distribution is chosen as prior distribution for each variance $\sigma_{s,m}^2$:

$$\sigma_{s,m}^2 | \rho_s, \psi_s \sim \mathcal{IG} \left(\frac{\rho_s}{2}, \frac{\psi_s}{2} \right), \quad (12)$$

By assuming a priori independence between the different variances, the hyperprior of σ_s^2 expresses as:

$$f(\sigma_s^2 | \rho_s, \psi_s) = \prod_{m=1}^M f(\sigma_{s,m}^2 | \rho_s, \psi_s). \quad (13)$$

The parameters ρ_s and ψ_s will be adjusted in order to obtain a vague hyperprior.

The prior for ψ_e is a non-informative Jeffrey's prior which reflects the lack of knowledge regarding this hyperparameter:

$$f(\psi_e) = \frac{1}{\psi_e} \mathbf{1}_{\mathbb{R}^+}(\psi_e). \quad (14)$$

Assuming the independence between the different hyperparameters, the hyperparameter prior distribution $\Phi = \{\sigma_s^2, \psi_e\}$ can be written:

$$f(\Phi) \propto \prod_{m=1}^M \left[\frac{1}{\sigma_{s,m}^{\rho_s+2}} \exp \left(-\frac{\psi_s}{2\sigma_{s,m}^2} \right) \right] \frac{1}{\psi_e} \mathbf{1}_{\mathbb{R}^+}(\psi_e). \quad (15)$$

3.4. Posterior distribution of Θ

The posterior distribution of the unknown parameter vector $\Theta = \{\mathbf{A}, \mathbf{S}, \sigma_e^2\}$ can be computed from the following hierarchical structure:

$$f(\Theta | \mathbf{Y}) \propto \int f(\mathbf{Y} | \Theta) f(\Theta | \Phi) f(\Phi) d\Phi, \quad (16)$$

where $f(\mathbf{Y} | \Theta)$ and $f(\Phi)$ have been defined in (4) and (15). Moreover, by assuming the independence between \mathbf{A} , \mathbf{S} and σ_e^2 , the following result can be obtained:

$$f(\Theta | \Phi) = f(\mathbf{A}) f(\mathbf{S} | \sigma_s^2) f(\sigma_e^2 | \rho_e, \psi_e), \quad (17)$$

where $f(\mathbf{A})$, $f(\mathbf{S} | \sigma_s^2)$ and $f(\sigma_e^2 | \rho_e, \psi_e)$ have been defined previously. This hierarchical structure allows one to integrate out the hyperparameter vector $\Phi = \{\psi_e, \sigma_s^2\}$ from the joint distribution $f(\Theta, \Phi | \mathbf{Y})$, yielding:

$$f(\mathbf{A}, \mathbf{S}, \sigma_e^2 | \mathbf{Y}) \propto \prod_{m=1}^M \left[\left[\psi_s + \|\mathbf{s}_m\|^2 \right]^{-\frac{L+\rho_s}{2}} \mathbf{1}_{\mathbb{R}^L_+}(\mathbf{s}_m) \right] \\ \times \prod_{i=1}^N \mathbf{1}_{\mathbb{S}}(\mathbf{a}_i) \prod_{i=1}^N \left[\left(\frac{1}{\sigma_{e,i}^2} \right)^{\frac{L}{2}+1} \exp \left[-\frac{\|\mathbf{y}_i - \mathbf{S} \mathbf{c}_i^T\|^2}{2\sigma_{e,i}^2} \right] \right].$$

4. GIBBS SAMPLER

Generating random samples according to $f(\mathbf{A}, \mathbf{S}, \sigma_e^2 | \mathbf{Y})$ is achieved by a Gibbs sampler whose main steps at each iteration are detailed below.

4.1. Generation according to $f(\mathbf{A} | \mathbf{S}, \sigma_e^2, \mathbf{Y})$

Straightforward computations yield for each observation:

$$f(\mathbf{a}_i | \mathbf{S}, \sigma_{e,i}^2, \mathbf{y}_i) \propto \exp \left[-\frac{(\mathbf{a}_i - \boldsymbol{\mu}_i)^T \boldsymbol{\Lambda}_i^{-1} (\mathbf{a}_i - \boldsymbol{\mu}_i)}{2} \right] \mathbf{1}_{\mathbb{T}}(\mathbf{a}_i),$$

where:

$$\begin{cases} \boldsymbol{\Lambda}_i = \left[\frac{1}{\sigma_{e,i}^2} (\mathbf{S} - \mathbf{s}_M \mathbf{u}^T)^T (\mathbf{S} - \mathbf{s}_M \mathbf{u}^T) \right]^{-1}, \\ \boldsymbol{\mu}_i = \boldsymbol{\Lambda}_i \left[\frac{1}{\sigma_{e,i}^2} (\mathbf{S} - \mathbf{s}_M \mathbf{u}^T)^T (\mathbf{y}_i - \mathbf{s}_M) \right], \end{cases} \quad (18)$$

with $\mathbf{u} = [1, \dots, 1]^T \in \mathbb{R}^{M-1}$. As a consequence, $\mathbf{a}_i | \mathbf{S}, \sigma_{e,i}^2, \mathbf{y}_i$ is distributed according to a truncated Gaussian distribution on the simplex \mathbb{S} :

$$\mathbf{a}_i | \mathbf{S}, \sigma_{e,i}^2, \mathbf{y}_i \sim \mathcal{N}_{\mathbb{S}}(\boldsymbol{\mu}_i, \boldsymbol{\Lambda}_i). \quad (19)$$

When the number of chemical components is relatively small, the generation of $\mathbf{a}_i | \mathbf{S}, \sigma_{e,i}^2, \mathbf{y}_i$ can be achieved using a standard Metropolis Hastings (MH) step. If the proposal distribution of this MH step is the Gaussian distribution $\mathcal{N}(\boldsymbol{\mu}_i, \boldsymbol{\Lambda}_i)$, the acceptance ratio of the MH algorithm reduces to 1 if the candidate is inside the simplex \mathbb{S} and 0 otherwise. For higher dimension problems, the acceptance ratio of the MH algorithm can be small, leading to poor mixing properties. In such cases, an alternative strategy based on a Gibbs sampler can be used (see [9] and [10]).

4.2. Generation according to $f(\sigma_e^2 | \mathbf{A}, \mathbf{S}, \mathbf{Y})$

This generation can be achieved thanks to the two following steps:

4.2.1. Generation according to $f(\psi_e | \sigma_e^2, \mathbf{A}, \mathbf{S}, \mathbf{Y})$

The conditional distribution is expressed as the following IG distribution:

$$\psi_e | \sigma_e^2, \rho_e \sim \text{IG} \left(\frac{N\rho_e}{2}, \frac{1}{2} \sum_{i=1}^N \frac{1}{\sigma_{e,i}^2} \right) \quad (20)$$

4.2.2. Generation according to $f(\sigma_e^2 | \psi_e, \mathbf{A}, \mathbf{S}, \mathbf{Y})$

From $f(\sigma_e^2, \mathbf{A}, \psi_e | \mathbf{S}, \mathbf{Y})$, it can be shown that the conditional distribution of the noise variance in each observation spectrum is the following IG distribution:

$$\sigma_{e,i}^2 | \psi_e, \mathbf{a}_i, \mathbf{S}, \mathbf{y}_i \sim \text{IG} \left(\frac{\rho_e + L}{2}, \frac{\psi_e + \|\mathbf{y}_i - \mathbf{S}\mathbf{c}_i^\top\|^2}{2} \right).$$

4.3. Generation according to $f(\mathbf{S} | \mathbf{A}, \sigma_e^2, \mathbf{Y})$

To sample according to $f(\mathbf{S} | \mathbf{A}, \sigma_e^2, \mathbf{Y})$, it is very convenient to generate samples from $f(\sigma_s^2, \mathbf{S} | \mathbf{A}, \sigma_e^2, \mathbf{Y})$ by using the following steps:

4.3.1. Generation according to $f(\sigma_s^2 | \mathbf{S}, \mathbf{A}, \sigma_e^2, \mathbf{Y})$

Looking at the joint distribution $f(\mathbf{A}, \mathbf{S}, \sigma_s^2, \sigma_e^2 | \mathbf{Y})$, the following result can be obtained, for each source spectrum, $m = 1, \dots, M$:

$$\sigma_{s,m}^2 | \mathbf{s}_m \sim \text{IG} \left(\frac{L + \rho_s}{2}, \frac{\psi_s + \|\mathbf{s}_m\|^2}{2} \right). \quad (21)$$

4.3.2. Generation according to $f(\mathbf{S} | \sigma_s^2, \mathbf{A}, \sigma_e^2, \mathbf{Y})$

Straightforward computations show that the posterior distribution of each source spectrum $f(\mathbf{s}_m | \sigma_s^2, \mathbf{A}, \sigma_e^2, \mathbf{Y})$, $m = 1, \dots, M$, is the following truncated Gaussian distribution:

$$\mathbf{s}_m | \sigma_s^2, \mathbf{A}, \sigma_e^2, \mathbf{Y} \sim \mathcal{N}^+ (\boldsymbol{\lambda}_m, \delta_m^2 \mathbf{I}_L), \quad (22)$$

with

$$\begin{cases} \boldsymbol{\lambda}_m = \delta_m^2 \left[\sum_{i=0}^N \frac{\boldsymbol{\epsilon}_{i,m}}{\eta_{i,m}^2} \right]^\top, \\ \delta_m^2 = \left[\sum_{i=0}^N \frac{1}{\eta_{i,m}^2} \right]^{-1}, \end{cases} \quad \text{and} \quad \begin{cases} \boldsymbol{\epsilon}_{0,m} = \mathbf{0}_L, \\ \boldsymbol{\epsilon}_{i,m} = \frac{\mathbf{y}_i - \mathbf{S}_{-m} \mathbf{c}_{i,-m}^\top}{c_{i,m}}, \\ \eta_{0,m}^2 = \sigma_{s,m}^2, \\ \eta_{i,m}^2 = \frac{\sigma_{e,i}^2}{c_{i,m}^2}, \end{cases} \quad (23)$$

and where \mathbf{S}_{-m} (resp. $\mathbf{c}_{i,-m}$) denotes the matrix \mathbf{S} (resp. the vector \mathbf{c}_i) where the m^{th} column (resp. the m^{th} coefficient) has been removed.

5. ILLUSTRATIONS

5.1. Synthetic data

This section presents some simulation results illustrating the performance of the proposed Bayesian blind unmixing procedure. The spectral sources used in the mixtures are simulated to get observed signals similar to real spectrometric data. Each spectrum is obtained as a superposition of Gaussian and Lorentzian shapes with different and randomly chosen parameters (location, amplitude and width). The mixing coefficients are also chosen to get evolution profiles similar to kinetic reactions. Figure 1 shows an example of $M = 3$ source signals of $L = 1000$ spectral bands. The abundance evolution profiles are simulated for $N = 15$ observation times. An i.i.d. Gaussian sequence is added to each observation to obtain a signal to noise ratio (SNR) equal to 20dB in each mixture.

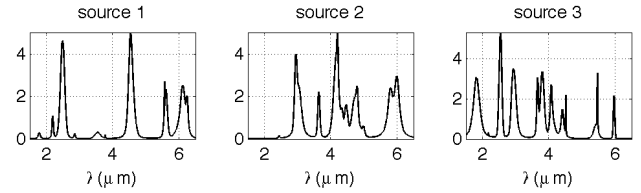


Fig. 1. Example of simulated spectral sources.

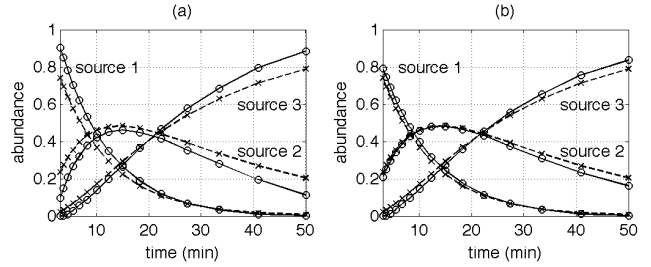


Fig. 2. Simulated (cross) and estimated (circles) mixing coefficients with positive Gaussian (a) and exponential (b) priors on the sources.

Figure 2 summarizes the result of 100 Monte Carlo runs for which the mixing matrix is kept unchanged while new sources and noise sequences are generated at each simulation. Figure 2-a shows a comparison between the true concentrations and their minimum mean square error (MMSE) estimates obtained for a Markov chain of $N_{\text{MC}} = 1000$ iterations including $N_{\text{b-i}} = 500$ burn-in iterations. The proposed Bayesian model has also been implemented with exponential priors for the spectral sources. The results obtained with such priors are shown in Figure 2-b. It can be firstly noticed that the estimated mixing coefficients satisfy positivity and additivity constraints, whatever the source prior. Moreover, an improvement in the abundance estimation accuracy is noted when using an exponential source prior. One may conclude

that the exponential prior seems more adequate for absorption spectra since this prior allows one to encode sparsity more efficiently than the Gaussian prior.

5.2. Application to infrared absorption spectra

To validate the algorithm with real data, an experiment has been performed on mixture data obtained from near infrared (NIR) spectroscopy measurements. Three known chemical species (cyclopentane, cyclohexane and n-pentane) are mixed experimentally with monitored concentrations. These species have been chosen for two main reasons: their available spectra in the NIR frequency bands are highly overlapping. In addition, as these species are inert solvents, they do not interact when they are mixed, which ensures that no new component appears. Figure 3 shows the pure spectra of the chemical species.

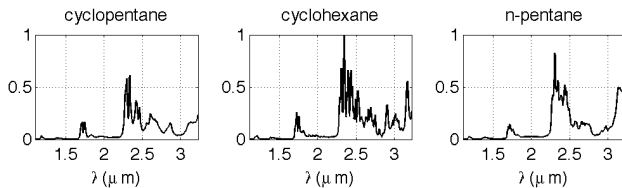


Fig. 3. Measured pure spectra of the mixed alkanes.

The results of concentration matrix estimation are reported in figure 4. It can be noticed that the abundance fraction profiles are estimated correctly, while a small difference between the actual abundances and the estimated ones can be noted. This is explained by the baseline that can be observed in the actual spectra in the spectral domain ranging from $2.4\mu m$ to $3.5\mu m$. As a consequence, the component spectra in the NIR region are less parsimonious than those used in subsection 5.1.

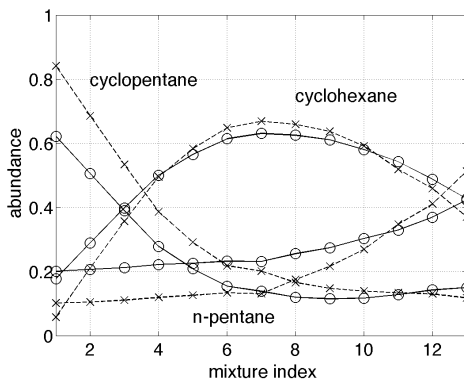


Fig. 4. True (cross) and estimated (circles) mixing coefficients using exponential priors on the sources.

6. CONCLUSION

This paper addresses the problem of unmixing linear mixtures whose regression coefficients satisfy additivity and positivity constraints. A hierarchical Bayesian model is defined with appropriate priors ensuring the model constraints. Estimation of the sources as well as the mixing coefficients is then performed by using samples distributed according to their joint posterior, generated by using a Gibbs sampling strategy. The application to spectroscopic signal analysis seems very promising. Indeed, the results conducted on synthetic and real data show that the proposed unmixing algorithm allows one to perform the separation under the additivity and positivity constraint efficiently. However the estimation can be improved by using more elaborated prior on the sources. This will be one of our future investigation.

7. REFERENCES

- [1] M.-H. Chen and J. J. Deely, "Bayesian analysis for a constrained linear multiple regression problem for predicting the new crop of apples," *J. of Agricultural, Biological, and Environmental Stat.*, vol. 1, no. 4, pp. 467–489, Dec. 1996.
- [2] S. Moussaoui, D. Brie, A. Mohammad-Djafari, and C. Carteret, "Separation of non-negative mixture of non-negative sources using a Bayesian approach and MCMC sampling," *IEEE Trans. Signal Processing*, vol. 54, no. 11, pp. 4133–4145, Nov. 2006.
- [3] C. Févotte and S. J. Godsill, "A Bayesian approach for blind separation of sparse sources," *IEEE Trans. Audio, Speech, Language Processing*, vol. 14, no. 6, pp. 2174–2188, Nov. 2006.
- [4] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Processing Magazine*, vol. 19, no. 1, pp. 44–57, Jan. 2002.
- [5] D. C. Heinz and C.-I. Chang, "Fully constrained least-squares linear spectral mixture analysis method for material quantification in hyperspectral imagery," *IEEE Trans. Geosci. and Remote Sensing*, vol. 29, no. 3, pp. 529–545, March 2001.
- [6] N. Dobigeon and J.-Y. Tourneret, "Spectral unmixing of hyperspectral images using a hierarchical Bayesian model," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal (ICASSP)*, Honolulu, Hawaii, USA, April 2007, pp. 1209–1212.
- [7] A. de Juan and R. Tauler, "Chemometrics applied to unravel multicomponent processes and mixtures: Revisiting latest trends in multivariate resolution," *Analytica Chimica Acta*, vol. 500, no. 1–2, pp. 195–210, Dec. 2003.
- [8] E. Punskeya, C. Andrieu, A. Doucet, and W. Fitzgerald, "Bayesian curve fitting using MCMC with applications to signal segmentation," *IEEE Trans. Signal Processing*, vol. 50, no. 3, pp. 747–758, March 2002.
- [9] C. P. Robert, "Simulation of truncated normal variables," *Statistics and Computing*, vol. 5, no. 2, pp. 121–125, June 1995.
- [10] N. Dobigeon and J.-Y. Tourneret, "Efficient sampling according to a multivariate Gaussian distribution truncated on a simplex," IRT/ENSEEIH/TéSA, Tech. Rep., March 2007. [Online]. Available: <http://www.enseiht.fr/~dobigeon>