Nonlinear Unmixing of Hyperspectral Images Using a Generalized Bilinear Model

Abderrahim Halimi, Yoann Altmann, Nicolas Dobigeon, *Member, IEEE*, and Jean-Yves Tourneret, *Senior Member, IEEE*

***Abstract*—Nonlinear models have recently shown interesting properties for spectral unmixing. This paper studies a generalized bilinear model and a hierarchical Bayesian algorithm for unmixing hyperspectral images. The proposed model is a generalization not only of the accepted linear mixing model but also of a bilinear model that has been recently introduced in the literature. Appropriate priors are chosen for its parameters to satisfy the positivity and sum-to-one constraints for the abundances. The joint posterior distribution of the unknown parameter vector is then derived. Unfortunately, this posterior is too complex to obtain analytical expressions of the standard Bayesian estimators. As a consequence, a Metropolis-within-Gibbs algorithm is proposed, which allows samples distributed according to this posterior to be generated and to estimate the unknown model parameters. The performance of the resulting unmixing strategy is evaluated via simulations conducted on synthetic and real data.**

***Index Terms*—Bayesian algorithm, bilinear model, Gibbs sampler, hyperspectral imagery, Markov chain Monte Carlo (MCMC) methods, spectral unmixing.**

I. Introduction

OVER the last few decades, spectral unmixing has been receiving considerable attention in the signal and image processing literature (see, for instance, [1] and the references therein). Unmixing hyperspectral images is based on the assumption that a pixel spectrum is a combination of pure constituent spectra, or endmembers, and a set of corresponding fractions, or abundances, that indicate the proportion of each endmember present in the pixel. The mixture model associated with spectral unmixing imagery can be linear or nonlinear, depending on the hyperspectral image under study. Linear mixtures known as macrospectral mixtures are interesting when the detected photons interact mainly with a single component on the scene before they reach the sensor. Conversely, nonlinear mixture models result from the interaction of photons with multiple components. Linear mixing models (LMMs) have motivated a lot of research works in the geoscience community [2]–[6]. However, nonlinear models constitute a new interesting field of research for hyperspectral imagery [7], [8]. In particular, nonlinear models have shown interesting properties for abundance estimation, e.g., for scenes including mixtures of minerals [9], orchards [10], or vegetation [11]–[13].

This paper studies a generalized bilinear model (GBM) for nonlinear unmixing of hyperspectral images due to multipath effects. This model is a generalization not only of the usual LMM but also of a bilinear model that has been recently introduced by Fan *et al*. in [9]. Spectral unmixing includes two main steps. The first step (referred to as *endmember extraction*) consists of extracting endmembers from the hyperspectral image. This extraction can be achieved in a supervised manner when prior information about the image is available. For instance, one might recognize classes of pure materials in the image and select the associated endmembers to create a learning set containing samples belonging to the different classes. When this information is not available, an automatic endmember extraction algorithm (EEA) has to be considered. Automatic EEA include the *pixel purity index* [14], the N-FINDR [15], and the *vertex component analysis* (VCA) [16]. They assume that the data set contains at least one pure pixel for each endmember (which is not always a realistic assumption) and extract the purest pixels from the image. The second step, called *inversion*, consists of estimating the corresponding abundances under positivity and sum-to-one constraints. Many different algorithms have been proposed in the literature to estimate the abundances for the LMM. These algorithms are based on the least square principle [3], maximum likelihood estimation [17], or Bayesian algorithms [18], [19]. Estimating the abundances for nonlinear models is a more challenging problem. Almost all algorithms for the unmixing of nonlinear models are based on the least square estimators as in [8], [9], and [13]. Some other methods based on support vector machines (SVMs) [20] and neural networks [21] have also been recently investigated.

This paper studies a Bayesian algorithm to estimate the abundance coefficients and the noise variance of the GBM. Appropriate prior distributions are chosen for the unknown GBM parameters. The joint posterior distribution of these parameters is then derived. However, the classical minimum mean square error (MMSE) and maximum *a posteriori* (MAP) estimators cannot be easily computed from this joint posterior. A classical way of alleviating this problem is to generate samples distributed according to the posterior using Markov chain Monte Carlo (MCMC) methods. These methods can also be used to estimate the model hyperparameters via hierarchical Bayesian inference. For that purpose, prior distributions for the hyperparameters have to be defined, introducing a second level of hierarchy into the Bayesian formulation. As in any

The authors are with the Institut de Recherche en Informatique de Toulouse (IRIT), École Nationale Supérieure d'Électronique, d'Électrotechnique, d'Informatique, d'Hydraulique, et des Télécommunications (ENSEEIHT) and Institut National Polytechnique de Toulouse (INP Toulouse), University of Toulouse, 31071 Toulouse, France (e-mail: Abderrahim.Halimi@enseeiht.fr; Yoann.Altmann@enseeiht.fr; Nicolas.Dobigeon@enseeiht.fr; Jean-Yves. Tourneret@enseeiht.fr).

Bayesian algorithm, the joint posterior distribution can be also used to compute confidence intervals for the parameter and hyperparameter estimates. Note that these hierarchical Bayesian algorithms have already been applied successfully to linear unmixing of hyperspectral images in [18] and [22]–[24].

This paper is structured as follows. Section II presents the linear and bilinear models considered in this study. The different components of the Bayesian algorithms associated with these models are studied in Section III. Section IV introduces the Metropolis-within-Gibbs sampler which will be used to generate samples according to the joint posterior of the unknown parameters and hyperparameters. Section V analyzes the quality of the proposed unmixing procedure. Section VI investigates the behavior of the proposed algorithm when applied to a single synthetic pixel. The simulation results obtained with synthetic data are presented in Section VI, whereas simulation results on real images are presented in Section VIII. Conclusions and future works are finally reported in Section IX.

II. Unmixing Models

A. LMM

The physical assumption underlying the LMM is that each incident photon interacts with one Earth surface component only and that the reflected spectra do not mix before entering the sensor [1], [10]. In this case, the L-spectrum $\boldsymbol{y} = [y_1, \ldots, y_L]^{\mathrm{T}}$ of a mixed pixel can be expressed as a mixture of R endmembers $\boldsymbol{m}_k$ with additive noise [18]

$$\boldsymbol{y} = \sum_{k=1}^{R} \alpha_k \boldsymbol{m}_k + \boldsymbol{n} = \boldsymbol{M}\boldsymbol{\alpha} + \boldsymbol{n} \tag{1}$$

where $\boldsymbol{M}$ is the $L \times R$ matrix whose columns are the $L \times 1$ endmember spectra $\boldsymbol{m}_k = [m_{1,k}, \ldots, m_{L,k}]^{\mathrm{T}}$, $k = 1, \ldots, R$, $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_R]^{\mathrm{T}}$ is the $R \times 1$ fractional abundance vector, and $\boldsymbol{n} = [n_1, \ldots, n_L]^{\mathrm{T}}$ is an additive white noise sequence. The additive noise is classically assumed to be an independent and identically distributed zero-mean Gaussian sequence with variance σ^2, denoted as $\boldsymbol{n} \sim \mathcal{N}(\boldsymbol{0}_L, \sigma^2 \boldsymbol{I}_L)$, where $\boldsymbol{I}_L$ is the identity matrix of dimension $L \times L$. As mentioned in the literature [1], [25], the abundances have to satisfy the following positivity and sum-to-one constraints:

$$\alpha_k \geq 0 \quad \forall k \in \{1, \ldots, R\} \quad \text{and} \quad \sum_{k=1}^{R} \alpha_k = 1. \tag{2}$$

B. Bilinear Mixing Model

Nonlinear mixture models account for the presence of multiple photon interactions by introducing additional "interaction" terms in the LMM [10]. The bilinear model considers second-order interactions between endmembers $\#i$ and $\#j$ (for i, $j = 1, \ldots, R$ and $i \neq j$) such that the observed mixed pixel $\boldsymbol{y}$ can be written as

$$\boldsymbol{y} = \boldsymbol{M}\boldsymbol{\alpha} + \sum_{i=1}^{R-1} \sum_{j=i+1}^{R} \beta_{i,j} \boldsymbol{m}_i \odot \boldsymbol{m}_j + \boldsymbol{n}$$

where $\odot$ is the Hadamard (term-by-term) product operation

$$\boldsymbol{m}_i \odot \boldsymbol{m}_j = \begin{pmatrix} m_{1,i} \\ \vdots \\ m_{L,i} \end{pmatrix} \odot \begin{pmatrix} m_{1,j} \\ \vdots \\ m_{L,j} \end{pmatrix} = \begin{pmatrix} m_{1,i} m_{1,j} \\ \vdots \\ m_{L,i} m_{L,j} \end{pmatrix}.$$

It is interesting to mention here that different bilinear models have been proposed in the literature. These models differ by the additivity constraints imposed on the abundances. For instance, the model proposed in [13], referred to as "Nascimento model," is based on the following constraints:

$$\sum_{k=1}^{R} \alpha_k + \sum_{i=1}^{R-1} \sum_{j=i+1}^{R} \beta_{i,j} = 1$$

whereas the model studied in [9], referred to as "Fan model" (FM), is defined by

$$\sum_{k=1}^{R} \alpha_k = 1 \qquad \beta_{i,j} = \alpha_i \alpha_j.$$

The GBM studied in this paper assumes that the observed mixed pixel can be written as

$$\boldsymbol{y} = \boldsymbol{M}\boldsymbol{\alpha} + \sum_{i=1}^{R-1} \sum_{j=i+1}^{R} \gamma_{i,j} \alpha_i \alpha_j \boldsymbol{m}_i \odot \boldsymbol{m}_j + \boldsymbol{n} \tag{3}$$

with the following constraints for the different parameters:

$$\begin{aligned} &\alpha_k \geq 0 \quad \forall k \in \{1, \ldots, R\} \quad \text{and} \quad \sum_{k=1}^{R} \alpha_k = 1 \\ &0 \leq \gamma_{i,j} \leq 1 \quad \forall i \in \{1, \ldots, R-1\} \quad \forall j \in \{i+1, \ldots, R\}. \end{aligned} \tag{4}$$

Note that $\gamma_{i,j}$ is a coefficient that controls the interactions between endmembers $\#i$ and $\#j$ in the considered pixel. An important property of the GBM is that it reduces to LMM for $\gamma_{i,j} = 0$ $\forall i = 1, \ldots, R-1$, $j = i+1, \ldots, R$, and to FM for $\gamma_{i,j} = 1$ $\forall i = 1, \ldots, R-1$, $j = i+1, \ldots, R$. The unknown parameter vector $\boldsymbol{\theta}$ associated with the GBM includes the abundance vector $\boldsymbol{\alpha}$, the nonlinearity coefficient vector $\boldsymbol{\gamma} = [\gamma_{1,2}, \ldots, \gamma_{R-1,R}]^{\mathrm{T}}$, and the noise variance σ^2. To understand the physics related to this model, assume that there are only two components "tree" and "soil" present in the observed pixel. The hyperspectral sensor will receive signals backscattered by the tree and the soil independently, corresponding to the terms $\alpha_1 \boldsymbol{m}_1$ and $\alpha_2 \boldsymbol{m}_2$. However, it will also receive interaction terms resulting from multiple scattering of photons between the two components "tree" and "soil" (which can be seen as multipath terms). For instance, if the signal is first backscattered by the tree and then backscattered by the soil, the hyperspectral sensor will receive an interaction term $\gamma_{12}\alpha_1\alpha_2 \boldsymbol{m}_1 \boldsymbol{m}_2$. It makes sense to assume $\gamma_{12} \leq 1$ since the path associated to the two reflections on the tree and the soil is longer than the direct paths associated to the tree only and soil only signals (resulting in smaller amplitudes). Note that higher order interaction terms are also received by the hyperspectral sensor. However, experiments conducted in [10] have shown that these higher

order terms can be neglected. The reader is invited to consult [9], [10], and [13] for more details.

III. Hierarchical Bayesian Model

This section introduces the hierarchical Bayesian model used to estimate the unknown parameter vector $\boldsymbol{\theta} = (\boldsymbol{\alpha}^T, \boldsymbol{\gamma}^{\mathrm{T}}, \sigma^2)^{\mathrm{T}}$ associated with the GBM.

A. Likelihood

The observation model defined in (3) and the Gaussian properties of the noise sequence $\boldsymbol{n}$ yield

$$f(\boldsymbol{y}|\boldsymbol{\theta}) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{L}{2}} \exp\left[-\frac{\|\boldsymbol{y} - \boldsymbol{\mu}_{\mathrm{GBM}}\|^2}{2\sigma^2}\right] \tag{5}$$

where $\boldsymbol{\mu}_{\mathrm{GBM}} = \boldsymbol{M}\boldsymbol{\alpha} + \sum_{i=1}^{R-1}\sum_{j=i+1}^{R} \gamma_{i,j}\alpha_i\alpha_j \boldsymbol{m}_i \odot \boldsymbol{m}_j$ and $\|\cdot\|$ denotes the standard l_2 norm such that $\|\boldsymbol{x}\|^2 = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}$.

B. Parameter Prior Distributions

This section details the prior distributions associated with the parameter vector $\boldsymbol{\theta}$.

1) Abundance Prior: The sum-to-one constraint can be encompassed by expressing one abundance α_{k^*} as a function of the others

$$\alpha_{k^*} = 1 - \sum_{k \neq k^*} \alpha_k.$$

As a consequence, the positivity and sum-to-one constraints (4) are satisfied if $\boldsymbol{\alpha}_{\backslash k^*} = [\alpha_1, \ldots, \alpha_{k^*-1}, \alpha_{k^*+1}, \ldots, \alpha_R]^{\mathrm{T}}$ belongs to the following simplex:

$$\mathcal{S}_{\backslash k^*} = \left\{ \boldsymbol{\alpha}_{\backslash k^*} | \alpha_k \geq 0 \quad \forall k \neq k^* \text{ and } \sum_{k \neq k^*} \alpha_k \leq 1 \right\}. \tag{6}$$

The prior chosen for $\boldsymbol{\alpha}_{\backslash k^*}$ is a uniform distribution on the simplex $\mathcal{S}_{\backslash k^*}$ since there is no additional information about this parameter vector. For any $\boldsymbol{\alpha}_{\backslash k^*} \in \mathcal{S}_{\backslash k^*}$, the abundance vector can be computed as follows:

$$\boldsymbol{\alpha} = \left[\alpha_1, \ldots, \alpha_{k^*-1}, 1 - \sum_{k \neq k^*} \alpha_k, \alpha_{k^*+1}, \ldots, \alpha_R\right]^{\mathrm{T}}$$

which shows that the abundance vector can be determined as a function of $\boldsymbol{\alpha}_{\backslash k^*} \in \mathcal{S}_{\backslash k^*}$. The resulting abundance vector has a singular probability density function (pdf) $f(\boldsymbol{\alpha})$ whose support is

$$\mathcal{S}_{\boldsymbol{\alpha}} = \left\{ \boldsymbol{\alpha} | \alpha_i \geq 0, i = 1, \ldots, R \text{ and } \sum_{i=1}^{R} \alpha_i = 1 \right\}. \tag{7}$$

2) Nonlinearity Parameter Prior: The parameters $\gamma_{i,j}$ are supposed to be positive as in [10] and [13] and less than one, reflecting the fact that the interaction abundances are always smaller than the product of the individual abundances. Since there is no other information about the nonlinearity parameters, each coefficient $\gamma_{i,j}$ is assigned a uniform prior on the interval [0, 1]. Assuming the parameters $\gamma_{i,j}$ are *a priori* independent, the prior pdf of $\boldsymbol{\gamma}$ can be written as

$$f(\boldsymbol{\gamma}) = \prod_{i=1}^{R-1} \prod_{j=i+1}^{R} \mathbb{I}_{[0,1]}(\gamma_{i,j}) \tag{8}$$

where $\mathbb{I}_{[0,1]}(.)$ is the indicator function defined on [0, 1].

3) Noise Variance Prior: It is very common to assign a conjugate inverse gamma prior to the noise variance parameter

$$\sigma^2 | \zeta_1, \zeta_2 \sim \mathcal{IG}\left(\frac{\zeta_1}{2}, \frac{\zeta_2}{2}\right) \tag{9}$$

where ζ_1 and ζ_2 are two hyperparameters. For simplicity, we set $\zeta_1 = 2$ and $\zeta_2 = \zeta$, in order to have a single adjustable hyperparameter ζ. This simplifying assumption has previously been applied successfully, e.g., in [18].

4) Hyperparameter Prior: A noninformative Jeffreys' prior is chosen for the hyperparameter ζ, which reflects the absence of knowledge about this hyperparameter [26]

$$f(\zeta) \propto \frac{1}{\zeta} \mathbb{I}_{\mathbb{R}^+}(\zeta). \tag{10}$$

C. Posterior distribution of $\boldsymbol{\theta}$

The posterior distribution of the parameter vector $\boldsymbol{\theta}$ can be computed as follows:

$$f(\boldsymbol{\theta}|\boldsymbol{y}) \propto \int f(\boldsymbol{y}|\boldsymbol{\theta}) f(\boldsymbol{\theta}|\zeta) f(\zeta)\, d\zeta \tag{11}$$

where $\propto$ means "proportional to," $f(\boldsymbol{y}|\boldsymbol{\theta})$ is the likelihood function defined in (5), and $f(\boldsymbol{\theta}|\zeta) = f(\boldsymbol{\alpha}) f(\boldsymbol{\gamma}) f(\sigma^2|\zeta)$ (assuming *a priori* independence between all the parameters). After substituting the likelihood and the priors in (11) and integrating out with respect to the hyperparameter ζ, the posterior of $\boldsymbol{\theta}|\boldsymbol{y}$ can be written as

$$f(\boldsymbol{\theta}|\boldsymbol{y}) \propto \frac{1}{\sigma^{L+2}} \exp\left[-\frac{\|\boldsymbol{y} - \boldsymbol{\mu}_{\mathrm{GBM}}\|^2}{2\sigma^2}\right] f(\boldsymbol{\gamma}) f(\boldsymbol{\alpha}). \tag{12}$$

The MMSE and MAP estimators associated with the posterior (12) are not easy to determine mainly because of the positivity and sum-to-one constraints contained in $f(\boldsymbol{\alpha})$. The next section presents a Metropolis-within-Gibbs algorithm that allows samples distributed according to the joint distribution $f(\boldsymbol{\theta}|\boldsymbol{y})$ to be generated. These samples are then used to estimate the unknown parameters.

IV. Metropolis-Within-Gibbs Algorithm

A. Generating Samples According to $f(\boldsymbol{\alpha}, \boldsymbol{\gamma}, \sigma^2|\boldsymbol{y})$

The principle of the Gibbs sampler is to generate samples according to the conditional distributions relative to the posterior distribution. When a conditional distribution cannot be sampled directly, an alternative consists of sampling according to a proposal distribution (whose support includes the support of the target distribution). The candidate is then accepted or rejected with an appropriate probability. The resulting algorithm whose sketch is described in Algorithm 1 (see the Appendix A) is classically referred to as Metropolis-within-Gibbs algorithm. In the considered unmixing problem, to simulate according to $f(\boldsymbol{\alpha}, \boldsymbol{\gamma}, \sigma^2|\boldsymbol{y})$, the Gibbs sampler iteratively generates samples distributed according to the conditional distributions:

1) $f(\alpha_k|\boldsymbol{\gamma}, \sigma^2, \boldsymbol{y}, \boldsymbol{\alpha}_{\setminus\{k,k^*\}})\ \forall k \neq k^*$;
2) $f(\gamma_{i,j}|\boldsymbol{\alpha}, \sigma^2, \boldsymbol{y}, \boldsymbol{\gamma}_{\setminus(i,j)})\ \forall i \neq R\ \forall j \in \{i+1, \ldots, R\}$;
3) $f(\sigma^2|\boldsymbol{y}, \boldsymbol{\alpha}, \boldsymbol{\gamma})$

that are derived in the Appendix A, where the following notations have been used:

$$\begin{aligned}
\boldsymbol{\alpha}_{i:j}^{(t)} &= \left(\alpha_i^{(t)}, \ldots, \alpha_j^{(t)}\right) \\
\boldsymbol{\alpha}_{\setminus\{k,k^*\}}^{(t)} &= \left(\boldsymbol{\alpha}_{1:k-1}^{(t)}, \boldsymbol{\alpha}_{k+1:k^*-1}^{(t-1)}, \boldsymbol{\alpha}_{k^*+1:R}^{(t-1)}\right) \\
\boldsymbol{\gamma}_{\setminus(i,j)}^{(t)} &= \left\{\gamma_{1,2}^{(t)}, \ldots, \gamma_{i,j-1}^{(t)}, \gamma_{i,j+1}^{(t-1)}, \ldots, \gamma_{R-1,R}^{(t-1)}\right\}.
\end{aligned}$$

The interested reader is invited to consult [27] for more details about Gibbs and Metropolis-within-Gibbs samplers, including the proofs of convergence of these algorithms.

B. Approximating the Bayesian Estimators

The Gibbs sampler introduced earlier generates a collection of N_{MC} samples

$$\mathcal{X}_{\boldsymbol{\theta}} = \left\{\sigma^{2(t)}, \boldsymbol{\alpha}_{\setminus k^*}^{(t)}, \boldsymbol{\gamma}^{(t)}\right\}_{t=1,\ldots,N_{\mathrm{MC}}}$$

that are asymptotically distributed according to the joint posterior $f(\sigma^2, \boldsymbol{\alpha}_{\setminus k^*}, \boldsymbol{\gamma}|\boldsymbol{y})$ in (12). Moreover, after a short burn-in period N_{bi}, the sequences

$$\begin{aligned}
\mathcal{X}_{\sigma^2} &= \left\{\sigma^{2(t)}\right\}_{t=1,\ldots,N_{\mathrm{MC}}} \\
\mathcal{X}_{\boldsymbol{\alpha}_{\setminus k^*}} &= \left\{\boldsymbol{\alpha}_{\setminus k^*}^{(t)}\right\}_{t=1,\ldots,N_{\mathrm{MC}}} \\
\mathcal{X}_{\boldsymbol{\gamma}} &= \left\{\boldsymbol{\gamma}^{(t)}\right\}_{t=1,\ldots,N_{\mathrm{MC}}}
\end{aligned}$$

form Markov chains whose stationary distributions are $f(\sigma^2|\boldsymbol{y})$, $f(\boldsymbol{\alpha}_{\setminus k^*}|\boldsymbol{y})$, and $f(\boldsymbol{\gamma}|\boldsymbol{y})$, respectively [27]. Consequently, the MMSE estimators of these parameters can be approximated by empirical averages over the last $N_r = N_{\mathrm{MC}} - N_{\mathrm{bi}}$ outputs of the sampler. More precisely, the MMSE estimates of the noise variance σ^2, the abundances α_k ($k = 1, \ldots, R$), and the nonlinearity parameters $\gamma_{i,j}$ ($i = 1, \ldots, R-1, j = i+1, \ldots, R$) are computed as follows:

$$\hat{x}^{(\mathrm{MMSE})} = E[x|\boldsymbol{y}] \approx \frac{1}{N_r} \sum_{t=1}^{N_{\mathrm{r}}} x^{(N_{\mathrm{bi}}+t)} \tag{13}$$

where x is the parameter of interest.

V. Unmixing Performance

This section introduces the criteria used to evaluate the quality of the unmixing strategy. The quality of the unmixing strategy for synthetic images can be measured by comparing the estimated and actual abundances by using the root mean square error (RMSE) [28]

$$\mathrm{RMSE} = \sqrt{\frac{1}{n\,R} \sum_{p=1}^{n} \|\boldsymbol{\alpha}(p) - \hat{\boldsymbol{\alpha}}(p)\|^2} \tag{14}$$

where $\boldsymbol{\alpha}(p)$ and $\hat{\boldsymbol{\alpha}}(p)$ are the actual and estimated abundance vectors of the pth pixel of the image and n is the number of pixels. The relative RMSE (RRMSE) can also be used by computing for each material

$$\mathrm{RRMSE}_j = \frac{\sqrt{\frac{1}{n} \sum_{p=1}^{n} (\alpha_j(p) - \hat{\alpha}_j(p))^2}}{\frac{1}{n} \sum_{p=1}^{n} \alpha_j(p)}.$$

In the case of real hyperspectral images, the reconstruction error (RE) is classically used to evaluate the quality of an unmixing method [1], [21]

$$\mathrm{RE} = \sqrt{\frac{1}{n\,L} \sum_{p=1}^{n} \|\hat{\boldsymbol{y}}(p) - \boldsymbol{y}(p)\|^2}$$

where L is the number of spectral bands and $\boldsymbol{y}(p)$ and $\hat{\boldsymbol{y}}(p)$ are the measured and estimated spectra for the pixel $\#p$. The following spectral angle mapper (SAM) is also used to estimate the performance of the unmixing procedure [1]:

$$\mathrm{SAM} = \frac{1}{n} \sum_{p=1}^{n} \theta\left[\boldsymbol{y}(p), \hat{\boldsymbol{y}}(p)\right]$$

where

$$\theta\left[\boldsymbol{y}(p), \hat{\boldsymbol{y}}(p)\right] = \arccos\left(\frac{\langle \boldsymbol{y}(p), \hat{\boldsymbol{y}}(p)\rangle}{\|\boldsymbol{y}(p)\|\ \|\hat{\boldsymbol{y}}(p)\|}\right)$$

and $\arccos(\cdot)$ is the inverse cosine operator.

VI. Simulation Results on a Synthetic Pixel

Many simulations have been conducted to validate the proposed unmixing algorithm. The first experiment considers a synthetic pixel defined as a GBM combination of three pure components (green grass, olive green paint, and galvanized steel metal) extracted from the ENVI software library. The abundances have been fixed to $\alpha_1 = 0.3$, $\alpha_2 = 0.6$, and $\alpha_3 = 0.1$ and the nonlinearity coefficients as $\gamma_{1,2} = 2/3$, $\gamma_{1,3} = 1/3$, and $\gamma_{2,3} = 2/3$. The observed spectrum has been corrupted by additive white Gaussian noise with variance $\sigma^2 = 2.8 \times 10^{-3}$ corresponding to a signal-to-noise ratio $SNR = 15$ dB, with $SNR = L^{-1}\sigma^{-2}\|\boldsymbol{y} - \boldsymbol{n}\|^2$. The unmixing algorithm has been run using $N_{\mathrm{bi}} = 300$ burn-in iterations and $N_r = 700$ iterations to compute the different estimates following the MMSE principle (13). The MMSE estimates of the abundances and the

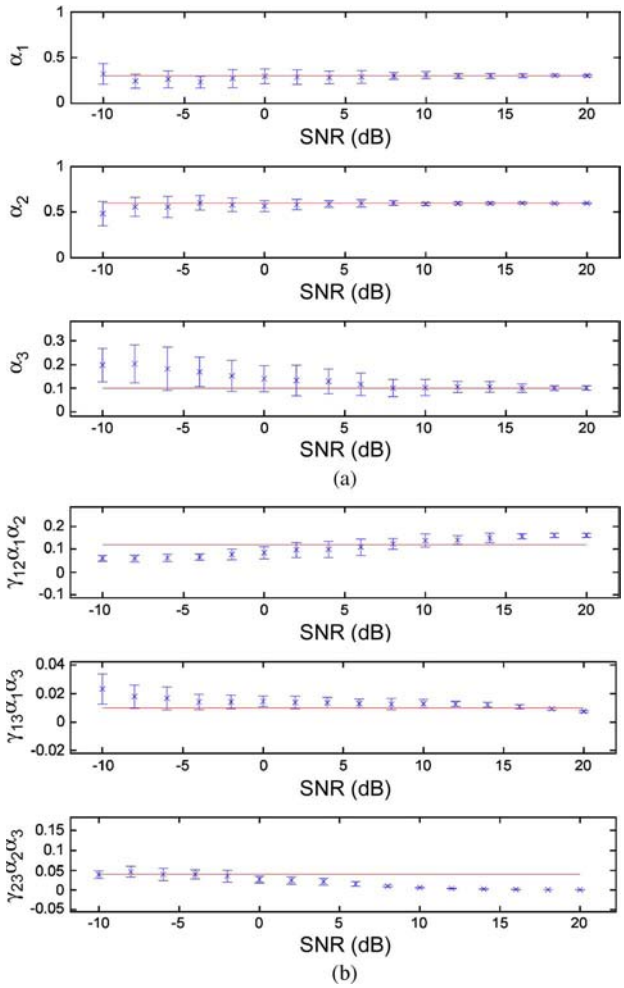

Fig. 1. (Cross) MMSE estimates and (vertical bars) standard deviations of (top) α_1, α_2, α_3 and (bottom) $\gamma_{1,2}\alpha_1\alpha_2$, $\gamma_{1,3}\alpha_1\alpha_3$ and $\gamma_{2,3}\alpha_2\alpha_3$ versus SNR. (a) Abundances of endmembers. (b) Abundances of endmember products.

corresponding standard deviations are represented as functions of the SNR in Fig. 1. These results have been obtained by running 30 Markov chains for each value of SNR. They are in good agreement with the actual values of abundances (red lines), particularly for high SNRs [note that the actual spectrometers, like Airborne Visible/Infrared Imaging Spectrometer (AVIRIS), provide images with SNR levels that are higher than 20 dB when the water absorption bands have been removed]. The small deviations between the MMSE estimates and the actual values of the nonlinear abundances $\gamma_{1,2}\alpha_1\alpha_2$, $\gamma_{1,3}\alpha_1\alpha_3$, and $\gamma_{2,3}\alpha_2\alpha_3$ are mainly due to the resemblance between the pure endmembers and some of the "endmember products" $\boldsymbol{m}_i \odot \boldsymbol{m}_j$. However, the effect of these deviations can be easily neglected since they are of small amplitudes.[1] The convergence issues related to this single pixel unmixing have been addressed in [29].

VII. Comparisons

Four synthetic images (10×10 pixels) have been generated using linear and/or nonlinear models to evaluate the performance of the proposed method. The three endmembers considered before have been used to generate mixtures according to the different models reported in Table I. Note that the 4th image has been generated using a hybrid model, i.e., half of the pixels have been generated with the standard LMM and half of the pixels have been generated with the GBM. The abundance vectors $\boldsymbol{\alpha}(p)$ $(p = 1, \ldots, 100)$ have been generated uniformly on the simplex with a cutoff threshold fixed to 0.8, i.e., $\alpha_k(p) \leq 0.8$, $k = 1, \ldots, R$. This means that there is no pure pixel in the generated images. The nonlinearity coefficients are uniformly drawn in the set [0, 1] for the GBM. All images have been corrupted by an additive Gaussian noise of variance $\sigma^2 = 2.8 \times 10^{-3}$. To evaluate the influence of the endmember extraction step on the unmixing algorithm accuracy, two scenarios have been considered.

1) The first scenario considers that the actual endmembers are used to unmix the synthetic images.
2) The second scenario uses an EEA to identify spectra used in the inversion step (here, the VCA [16]). This scenario will allow one to measure the influence of the absence of pure pixels in the images.

Note that VCA is an algorithm that looks for the simplex of largest volume inside the data. It relies on the standard LMM and exploits the fact that the endmembers are the vertices of a simplex. However, it is possible to assume that the endmembers are still extreme spectra when small nonlinearities occur (see [28] and [29] for more details).

The performance of the proposed algorithm is compared with the performance of other unmixing strategies specifically dedicated to other linear and/or nonlinear models. The algorithm proposed in [9] for a bilinear model denoted FM and introduced in Section II-B is applied on the four synthetic images. Two linear unmixing algorithms, proposed in [3] and [18], have been also considered. The strategy proposed in [18] defines priors for the abundances satisfying additivity and positivity constraints and estimates the unknown model parameters using Bayesian inference. Finally, the fully constrained least square (FCLS) algorithm [3] formulates the linear unmixing of hyperspectral images as a constrained optimization problem.

TABLE I
Mixing Models Used to Generate the Synthetic Images

Synthetic images	Mixing models	Parameters
image 1 (I_1)	LMM	$\alpha_1, \alpha_2, \alpha_3$
image 2 (I_2)	FM	$\alpha_1, \alpha_2, \alpha_3$
image 3 (I_3)	GBM	$\alpha_1, \alpha_2, \alpha_3, \gamma_{1,2}, \gamma_{1,3}, \gamma_{2,3}$
image 4 (I_4)	hybrid model	$\alpha_1, \alpha_2, \alpha_3, \gamma_{1,2}, \gamma_{1,3}, \gamma_{2,3}$

A. Performance With Actual Endmembers (First Scenario)

Table II reports the REs and SAMs obtained by unmixing the four images with the different algorithms. Obviously, the most appropriate algorithm to unmix a given image is the algorithm based on the corresponding model. However, it can be also noticed that the proposed GBM coupled with the proposed algorithm is able to handle all mixing models efficiently. Indeed, it provides good results for all images. Another means of evaluating the performance of the unmixing for synthetic images is to compute the RMSE between the estimated and

[1] Note that different scales have been used in the plots of Fig. 1, which makes the errors more visible for the nonlinear abundances.

TABLE II
UNMIXING ALGORITHM PERFORMANCES WITH ACTUAL AND ESTIMATED ENDMEMBERS (FIRST AND SECOND SCENARIOS): RE AND SAM

	1st scenario								2nd scenario							
	RE ($\times 10^{-2}$)				SAM ($\times 10^{-2}$)				RE ($\times 10^{-2}$)				SAM ($\times 10^{-2}$)			
	LMM		FM	GBM	LMM		FM	GBM	LMM		FM	GBM	LMM		FM	GBM
	Bay	FCLS			Bay	FCLS			Bay	FCLS			Bay	FCLS		
I_1	5.48	8.65	5.75	5.48	15.55	18.68	16.12	15.56	6.39	5.62	7.04	5.64	12.59	11.57	12.45	11.57
I_2	6.81	7.21	5.44	5.57	15.55	16.76	13.93	14.08	6.95	5.64	6.22	5.52	12.97	11.46	12.19	11.39
I_3	6.01	7.67	5.55	5.50	15.40	17.39	14.87	14.75	6.93	5.93	7.37	6.00	13.65	12.10	12.82	12.05
I_4	5.81	7.82	5.65	5.51	15.49	17.68	15.42	15.11	7.19	5.84	6.84	5.85	14.57	12.53	12.84	12.31

TABLE III
UNMIXING ALGORITHM PERFORMANCES WITH ACTUAL ENDMEMBERS (FIRST SCENARIO)

	RMSE ($\times 10^{-2}$)			
	LMM		FM	GBM
	Bay	FCLS		
I_1	0.99	24.68	12.93	1.86
I_2	15.82	26.16	0.80	7.73
I_3	9.52	25.73	6.89	4.02
I_4	7.31	23.97	9.74	3.42

TABLE IV
UNMIXING ALGORITHM PERFORMANCES WITH ESTIMATED ENDMEMBERS (SECOND SCENARIO)

	RMSE ($\times 10^{-2}$)			
	LMM		FM	GBM
	Bay	FCLS		
I_1	13.19	9.65	22.71	9.46
I_2	16.84	13.12	22.97	12.16
I_3	15.21	11.14	22.10	10.63
I_4	19.68	19.17	24.40	18.65

actual abundances, as defined by (14). These RMSE values are depicted in Table III, which shows that the GBM is a good compromise between all of the models.

B. Performance With Estimated Endmembers (Second Scenario)

The influence of the endmember extraction step has been evaluated by unmixing the synthetic images with endmembers recovered by VCA. Note that there is no pure pixel in these images, as explained previously. Consequently, the inversion procedures applied on the images rely on approximated endmembers. The REs and SAMs obtained with the different algorithms on the four images are reported in Table II. The corresponding abundance RMSE values are shown in Table IV. Since the endmembers used to estimate the abundances are approximated, a degradation with respect to Tables II and III can be observed. However, the GBM and its associated Bayesian algorithm are more robust to the imperfect knowledge of endmembers. This robustness can be explained by the intrinsic flexibility of the proposed GBM when compared to the other models.

Table V shows the RRMSEs between the actual and estimated abundances (using the GBM) for image 4. The RRMSEs obtained for the abundances $\alpha_1, \ldots, \alpha_3$ are smaller than those obtained for the nonlinearity coefficients $\gamma_{1,2}, \ldots, \gamma_{2,3}$. This can be explained by the fact that each $\gamma_{i,j}$ is multiplied by the factor $\alpha_i \alpha_j$ to obtain the corresponding abundance of the endmember product $\boldsymbol{m}_i \odot \boldsymbol{m}_j$. Consequently, when the abundance product $\hat{\alpha}_i \hat{\alpha}_j$ is small, the estimated nonlinearity coefficient $\hat{\gamma}_{i,j}$ can differ from its actual value without introducing a significant difference in the RE. Table VI shows the computational times of MATLAB implementations on a 1.66-GHz dual-core PC of the studied algorithms. The proposed algorithm exhibits a higher computational complexity when compared to the other algorithms. However, it provides better unmixing performance than the other unmixing strategies.

TABLE V
RELATIVE RMSE FOR THE GBM-BASED UNMIXING OF IMAGE 4

	Relative RMSE ($\times 10^{-2}$)	
	1st scenario (known endmembers)	2nd scenario (VCA endmembers)
α_1	9.74	17.08
α_2	4.99	21.05
α_3	9.63	22.76
$\gamma_{1,2}$	64.26	84.08
$\gamma_{1,3}$	67.33	93.34
$\gamma_{2,3}$	96.53	97.51

TABLE VI
COMPUTATIONAL TIMES OF THE UNMIXING ALGORITHMS WITH ACTUAL AND APPROXIMATED ENDMEMBERS (IN SECONDS)

	1st scenario				2nd scenario			
	LMM		FM	GBM	LMM		FM	GBM
	Bay	FCLS (10^{-2})			Bay	FCLS (10^{-2})		
I_1	35	7	10	236	49	6	14	279
I_2	46	3	9	258	48	2	13	268
I_3	38	3	10	240	48	3	12	280
I_4	36	3	10	241	53	3	13	273

VIII. SPECTRAL UNMIXING OF AVIRIS IMAGES

This section illustrates the performance of the proposed algorithm when applied to a real hyperspectral data set. The real images used in this section are described in Section VIII-A. The endmembers that are present in the two considered scenes are extracted in Section VIII-B. The abundance estimation is finally presented in Section VIII-C, whereas the quality of the unmixing is studied in Section VIII-D.

A. Description of the Hyperspectral Data

The first hyperspectral image has received much attention in the remote sensing community [25], [30]. This image was

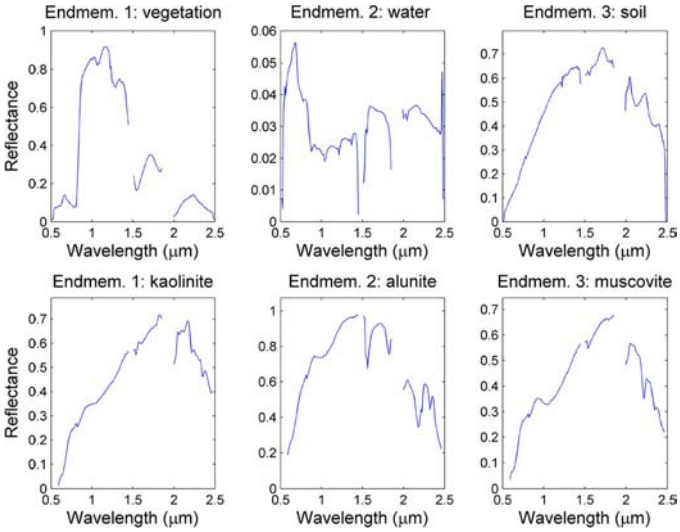

Fig. 2. $R = 3$ endmember spectra obtained by VCA for (top) Moffett Field and (bottom) Cuprite.

acquired over Moffett Field, CA, in 1997 by the AVIRIS. The proposed unmixing algorithm has been applied on a 50×50 subimage. This area of interest has $L = 203$ spectral bands after removing water absorption bands and is mainly composed of three components: water, soil, and vegetation [18].

The second image was acquired also by AVIRIS over the Cuprite mining site, Nevada, in 1997. The image of interest (which is of size 50×50) has 189 spectral bands. The geologic characteristics of this area, dominated by muscovite, alunite, and cuprite, have been investigated in [31].

B. Endmember Determination

The first step in the unmixing procedure is endmember determination. The endmember spectra can be determined by averaging the pixel spectra on appropriate regions of interest when preliminary knowledge is available. However, when no *a priori* knowledge is available, a fully automatic procedure is required. The extraction of endmembers from the AVIRIS images has been performed by VCA. As explained in Section VII, using an EEA that looks for the vertices of a simplex is allowed to perform nonlinear unmixing. The obtained spectra for the Moffett Field and Cuprite images are shown in Fig. 2. These spectra have been identified using prior knowledge regarding the Moffett field image (which is known to be composed of vegetation, soil, and water) and by using the U.S. Geological Survey spectral library when analyzing the Cuprite image.

C. Abundance Estimation

The abundances of each image have been estimated by running the proposed Bayesian unmixing method on each pixel of the AVIRIS images (using the endmember spectra resulting from Section VIII-B). The unmixing procedure has been performed by considering the $R = 3$ corresponding endmember spectra shown in Fig. 2. The MMSE estimates of the abundances have been computed by averaging the $N_r = 700$ last generated samples obtained after $N_{\text{bi}} = 300$ burn-in iterations. The image fraction maps, estimated by the proposed method and relative to the linear contribution of the endmembers, are

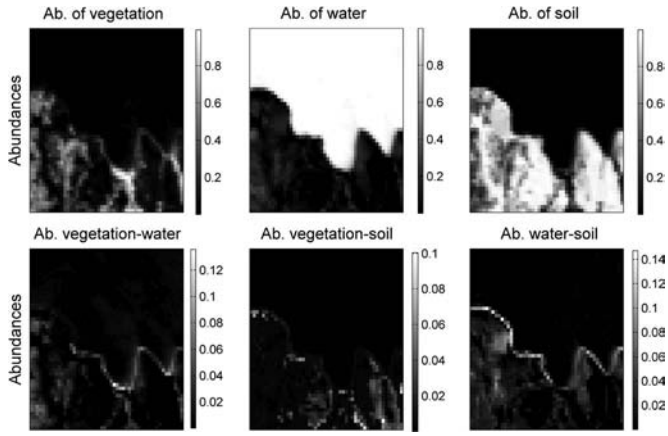

Fig. 3. Fraction maps estimated according to GBM for Moffett Field.

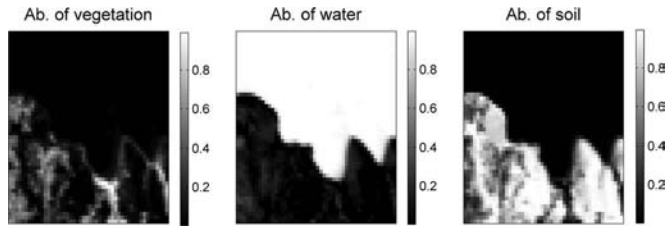

Fig. 4. Fraction maps estimated according to LMM for Moffett Field.

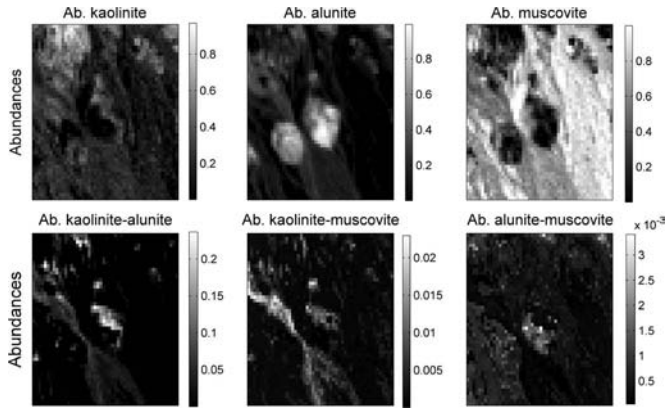

Fig. 5. Fraction maps estimated according to GBM for Cuprite.

shown in Fig. 3 (top). Note that a white (black) pixel indicates a large (small) proportion of the corresponding materials. These pictures are in good agreement with the abundances obtained when considering a linear model, as shown in Fig. 4.

The proposed algorithm also provides maps for the possible interactions between the materials. These maps are shown in Fig. 3 (bottom). The interactions between vegetation and water are located in the vegetation area, whereas the interactions between vegetation and soil occur in the soil area. The interactions between soil and water are essentially located in the coastal area, which is the zone of contact between these two components.

Figs. 5 and 6 show the results obtained with the Cuprite image by considering GBM and LMM unmixing algorithms. The maps corresponding to the material abundances are similar

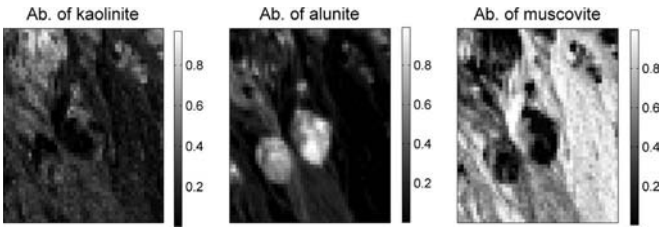

Fig. 6. Fraction maps estimated according to LMM for Cuprite.

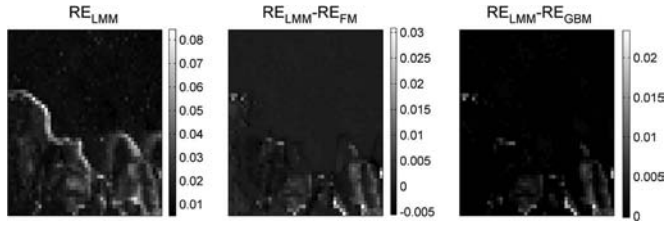

Fig. 7. RE maps obtained for Moffett Field.

TABLE VII
UNMIXING PERFORMANCE FOR AVIRIS IMAGES OBTAINED WITH DIFFERENT MODELS

		Moffett Field		Cuprite	
		RE ($\times 10^{-2}$)	SAM ($\times 10^{-1}$)	RE ($\times 10^{-2}$)	SAM ($\times 10^{-2}$)
LMM	FCLS	1.88	1.843	2.11	3.130
	Bay.	1.92	1.842	2.14	3.140
FM		1.82	1.848	3.05	5.087
GBM		1.82	1.837	1.92	2.952

for the two models. However, the GBM provides three other maps for nonlinearity coefficients. The interaction terms are located in regions where two components are present.

D. Quality of the Unmixing

The RE and SAM criteria introduced in Section V have been used to evaluate the quality of unmixing. Table VII shows the results obtained when unmixing the AVIRIS images using the LMM, FM and GBM. The results obtained for Moffett Field indicate that the GBM is well adapted to this image. Comparing these results with those obtained for synthetic images leads to the following conclusions.

1) Moffett Field is composed of linear and nonlinear mixtures of the endmembers.
2) The nonlinearity present in the image is captured well by the FM and the GBM.
3) The similarity between the results of Tables II and VII shows that the actual and estimated endmembers are similar which indicates the presence of pure pixels in the Moffett Field image.
4) The best unmixing model for this image is the GBM.

Unmixing the Cuprite image by the GBM provides better results than with LMM. These results are similar to those obtained in Table II (second scenario) for image 3, which leads to the following conclusions.

1) The nonlinearities present in Cuprite follow the GBM.
2) The endmembers are not well extracted, which can be explained by the absence of pure pixels in the image.
3) The best unmixing model for this image is the GBM.

Finally, to illustrate the importance of considering nonlinearities when performing unmixing, three maps of RE related to the Moffett Field image are shown in Fig. 7. The first RE map has been obtained when considering the LMM. The two others have been obtained by computing the difference between this LMM-based RE map and the corresponding REs obtained when considering FM and GBM. The last two RE maps show that FM and GBM significantly reduce the REs in the areas corresponding to the interactions between vegetation and soil. This means that they allow the mixture to be better approximated when nonlinearity occurs. The obtained figure also shows that the GBM outperforms the LMM for all pixels in the RE sense.

IX. CONCLUSION

This paper studied a generalized bilinear model (GBM) to model the interactions between the macroscopic components of a hyperspectral image. A hierarchical Bayesian algorithm was proposed to estimate the abundances and nonlinearity coefficients of this nonlinear model. Appropriate priors were chosen to ensure that the abundances satisfy positivity and sum-to-one constraints. The posterior distribution of the unknown parameter vector was then derived. The corresponding Bayesian estimators were approximated from samples generated using MCMC methods. A detailed comparative analysis of this bilinear model with other linear and nonlinear models showed that the proposed mixture model better approximates the existing interactions between pure spectral components. Future investigations include the consideration of spatial correlation between pixels of the hyperspectral image using the proposed GBM.

APPENDIX A
GIBBS SAMPLER FOR THE GENERALIZED BILINEAR MODEL

A. Generating Samples According to $f(\gamma_{i,j}|\boldsymbol{\alpha}, \sigma^2, \boldsymbol{y}, \boldsymbol{\gamma}_{\setminus(i,j)})$

The conditional distribution of $\gamma_{i,j}|\boldsymbol{\alpha}$, σ^2, $\boldsymbol{y}$, $\boldsymbol{\gamma}_{\setminus(i,j)}$ for $i = 1, \ldots, R-1$ and $j = i+1, \ldots, R$ is

$$\gamma_{i,j}|\boldsymbol{\alpha}, \sigma^2, \boldsymbol{y}, \boldsymbol{\gamma}_{\setminus(i,j)} \sim \mathcal{N}_{[0,1]}\left(\frac{\boldsymbol{p}_{i,j}^{\mathrm{T}}\boldsymbol{e}_{i,j}}{\|\boldsymbol{p}_{i,j}\|^2}, \frac{\sigma^2}{\|\boldsymbol{p}_{i,j}\|^2}\right)$$

where

$$\begin{cases} \boldsymbol{p}_{i,j} = \boldsymbol{m}_i \odot \boldsymbol{m}_j \\ \boldsymbol{e}_{i,j} = \boldsymbol{y} - \boldsymbol{M}\boldsymbol{\alpha} - \sum_{l=1, l\neq i}^{R-1} \sum_{p=l+1, p\neq j}^{R} \gamma_{lp}\alpha_l\alpha_p \boldsymbol{m}_l \odot \boldsymbol{m}_p \end{cases}$$

and $\mathcal{N}_{[0,1]}(\cdot,\cdot)$ denotes the Gaussian distribution truncated on the set [0, 1]. The simulation of samples according to this truncated Gaussian distribution can be performed efficiently by using the method proposed in [32] and detailed in [29].

B. Generating Samples According to $f(\alpha_k|\boldsymbol{\gamma}, \sigma^2, \boldsymbol{y}, \boldsymbol{\alpha}_{\setminus\{k,k^*\}})$

The conditional pdf of $\alpha_k|\boldsymbol{\gamma}, \sigma^2, \boldsymbol{y}, \boldsymbol{\alpha}_{\setminus\{k,k^*\}}$ is given by

$$f(\alpha_k|\boldsymbol{\gamma}, \sigma^2, \boldsymbol{y}, \boldsymbol{\alpha}_{\setminus\{k,k^*\}}) \propto \exp\left[-\frac{\left\|\boldsymbol{g}_{k,k^*} - \alpha_k \boldsymbol{h}_{k,k^*} + \alpha_k^2 \boldsymbol{q}_{k,k^*}\right\|^2}{2\sigma^2}\right] \mathbb{I}_{[0,\alpha_k^+]}(\alpha_k) \tag{15}$$

where

$$\begin{cases} \alpha_k^+ = 1 - \sum_{j\neq k} \alpha_j \\ \boldsymbol{g}_{k,k^*} = \boldsymbol{y} - \boldsymbol{s}_{k,k^*} - \boldsymbol{u}'_{k,k^*} \\ \boldsymbol{h}_{k,k^*} = \boldsymbol{m}_k \odot \boldsymbol{v}_{k,k^*} - \boldsymbol{m}_{k^*} \odot \boldsymbol{w}_{k,k^*} - \boldsymbol{m}_{k^*} \\ \boldsymbol{q}_{k,k^*} = \boldsymbol{\gamma}_{k,k^*} \boldsymbol{p}_{k,k^*} \end{cases}$$

with

$$\begin{cases} \boldsymbol{s}_{k,k^*} = \boldsymbol{t}_{k,k^*} + (1 - \lambda_{k,k^*})\boldsymbol{m}_{k^*} + \boldsymbol{m}_{k^*} \odot \boldsymbol{w}_{k,k^*} + \boldsymbol{u}_{k,k^*} \\ \boldsymbol{t}_{k,k^*} = \sum_{i\notin\{k,k^*\}} \alpha_i \boldsymbol{m}_i \\ \boldsymbol{u}_{k,k^*} = \sum_{i\notin\{R,k,k^*\}} \sum_{j\in\{i+1,\ldots,R\}\setminus\{k,k^*\}} \boldsymbol{\gamma}_{i,j} \alpha_i \alpha_j \boldsymbol{p}_{i,j} \\ \boldsymbol{u}'_{k,k^*} = \sum_{i\notin\{k,k^*\}} \sum_{j\notin\{k,k^*\}} \boldsymbol{\gamma} i, k^* \alpha_i \alpha_j \boldsymbol{p}_{i,k^*} \end{cases}$$

$$\begin{cases} \boldsymbol{w}_{k,k^*} = \sum_{i\notin\{k,k^*\}} \boldsymbol{\gamma}_{i,k^*} \alpha_i \boldsymbol{m}_i \\ \boldsymbol{z}_{k,k^*} = \sum_{i\notin\{k,k^*\}} \boldsymbol{\gamma}_{k,i} \alpha_i \boldsymbol{m}_i \\ \lambda_{k,k^*} = \sum_{i\notin\{k,k^*\}} \alpha_i \\ \boldsymbol{v}_{k,k^*} = \boldsymbol{1}_L + \boldsymbol{z}_{k,k^*} + \boldsymbol{\gamma}_{k,k^*} \boldsymbol{m}_{k^*} - \boldsymbol{\gamma}_{k,k^*} \lambda_{k,k^*}\ \boldsymbol{m}_{k^*} \\ \boldsymbol{1}_L = [1, \ldots, 1]^{\mathrm{T}} \text{ of size } L \times 1. \end{cases}$$

The conditional distribution (15) is not easy to sample. As a consequence, at iteration t of the sampler, the kth component of the vector $\boldsymbol{\alpha}_{\setminus\{k^*\}}$ is updated using a Metropolis–Hastings step. A candidate $\alpha_k^{(\star)}$ is generated according to a proposal distribution $\pi(\alpha_k^{(\star)})$. This candidate is accepted or rejected with the probability ρ defined in Algorithm 2. The proposal used in this paper is a Gaussian distribution truncated on [0, 1] whose mean ω_{k,k^*} is the mode of the pdf $f(\alpha_k|\boldsymbol{y}, \sigma^2, \boldsymbol{\alpha}_{\setminus\{k,k^*\}})$ (see [29] for more details). The variance of the proposal distribution has to ensure an appropriate exploration of the target distribution support. Our experiments have shown that an interesting value for this variance is $(2\sigma^2/\boldsymbol{h}_k^{\mathrm{T}}\boldsymbol{h}_k)$, leading to

$$\alpha_k^{(\star)} \sim \mathcal{N}_{[0,1]}\left(\omega_{k,k^*}, \frac{2\sigma^2}{\boldsymbol{h}_k^{\mathrm{T}}\boldsymbol{h}_k}\right). \tag{16}$$

The resulting algorithm sketch is summarized in Algorithm 2.

Algorithm 1 Gibbs sampler
1: Initialization
2: Initialize parameters $k^* = R$, $\sigma^{2(0)}$, $\alpha_k^{(0)}$ $(\forall k \neq k^*)$
3: $\alpha_{k^*}^{(0)} = 1 - \sum_{i=1}^{R-1} \alpha_i^{(0)}$
4: Sample generation
5: **for** $t = 1 : N_{\mathrm{bi}} + N_r$ **do**
6: Choose a random k^* from $\{1, \ldots, R\}$
7: **for** $k = 1 : R$ and $k \neq k^*$ **do**
8: Sample $\alpha_k^{(t)} \sim f(\alpha_k|\sigma^{2(t-1)}, \boldsymbol{y}, \boldsymbol{\gamma}, \boldsymbol{\alpha}_{\setminus\{k,k^*\}}^{(t)})$ using Algorithm 2
9: **end for**
10: Set $\alpha_{k^*}^{(t)} = 1 - \sum_{i\neq k^*} \alpha_i^{(t)}$
11: Sample $\gamma_{1,2} \sim f(\gamma_{1,2}|\sigma^{2(t-1)}, \boldsymbol{y}, \boldsymbol{\gamma}_{\setminus(1,2)}^{(t)}, \boldsymbol{\alpha}^{(t)})$
12: $\vdots$ $\vdots$ $\vdots$
13: Sample $\gamma_{R-1,R} \sim f(\gamma_{R-1,R}|\sigma^{2(t-1)}, \boldsymbol{y}, \boldsymbol{\gamma}_{\setminus(R-1,R)}^{(t)}, \boldsymbol{\alpha}^{(t)})$
14: Sample $\sigma^{2(t)} \sim f(\sigma^2|\boldsymbol{y}, \boldsymbol{\gamma}^{(t)}, \boldsymbol{\alpha}^{(t)})$
15: **end for**

Algorithm 2 Sampling α_k with "Metropolis-within-Gibbs"
1: draw a candidate $\alpha_k^{(\star)}$ from the proposal $\pi(\alpha_k^{(\star)})$ in (16),
2: compute the acceptance probability

$$\rho = \min\left\{1, \frac{f\left(\alpha_k^{(\star)}|\sigma^{2(t-1)}, \boldsymbol{y}, \boldsymbol{\gamma}, \boldsymbol{\alpha}_{\setminus\{k,k^*\}}^{(t)}\right)}{f\left(\alpha_k^{(t-1)}|\sigma^{2(t-1)}, \boldsymbol{y}, \boldsymbol{\gamma}, \boldsymbol{\alpha}_{\setminus\{k,k^*\}}^{(t)}\right)} \frac{\pi\left(\alpha_k^{(t-1)}\right)}{\pi\left(\alpha_k^{(\star)}\right)}\right\}$$

3: sample $w \sim U_{[0,1]}$
4: **if** $w < \rho$ **then**
5: $\alpha_k^{(t)} = \alpha_k^{(\star)}$
6: **else**
7: $\alpha_k^{(t)} = \alpha_k^{(t-1)}$
8: **end if**

C. Generating Samples According to $f(\sigma^2|\boldsymbol{y}, \boldsymbol{\alpha}, \boldsymbol{\gamma})$

Looking carefully at the joint posterior distribution (12), the conditional distribution of $\sigma^2|\boldsymbol{y}, \boldsymbol{\alpha}, \boldsymbol{\gamma}$ can be determined

$$\sigma^2|\boldsymbol{y}, \boldsymbol{\alpha}, \boldsymbol{\gamma} \sim \mathcal{IG}\left(\frac{L}{2}, \frac{\|\boldsymbol{y} - \boldsymbol{\mu}_{\mathrm{GBM}}\|^2}{2}\right).$$

ACKNOWLEDGMENT

The authors would like to thank Prof. X. Jia from the University of New South Wales, Australian Defence Force Academy, Canberra, Australia, for the interesting discussion regarding this work and Prof. S. McLaughlin from the University of Edinburgh, Edinburgh, U.K., for helping them fix the English grammar.

REFERENCES

[1] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 44–57, Jan. 2002.
[2] M. Craig, "Minimum volume transforms for remotely sensed data," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 3, pp. 542–552, May 1994.
[3] 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. Remote Sens.*, vol. 39, no. 3, pp. 529–545, Mar. 2001.
[4] D. Manolakis, C. Siracusa, and G. Shaw, "Hyperspectral subpixel target detection using the linear mixing model," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 7, pp. 1392–1409, Jul. 2001.
[5] J. M. P. Nascimento and J. M. Bioucas-Dias, "Does independent component analysis play a role in unmixing hyperspectral data?" *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 1, pp. 175–187, Jan. 2005.
[6] A. Zare and P. Gader, "PCE: Piece-wise convex endmember detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 6, pp. 2620–2632, Jun. 2010.
[7] B. W. Hapke, "Bidirectional reflectance spectroscopy. I. Theory," *J. Geophys. Res.*, vol. 86, no. B4, pp. 3039–3054, 1981.

[8] J.-P. Combe, P. Launeau, V. Carrère, D. Despan, V. Méléder, L. Barillé, and C. Sotin, "Mapping microphytobenthos biomass by non-linear inversion of visible-infrared hyperspectral images," *Remote Sens. Environ.*, vol. 98, no. 4, pp. 371–387, Oct. 2005.
[9] W. Fan, B. Hu, J. Miller, and M. Li, "Comparative study between a new nonlinear model and common linear model for analysing laboratory simulated-forest hyperspectral data," *Int. J. Remote Sens.*, vol. 30, no. 11, pp. 2951–2962, Jun. 2009.
[10] B. Somers, K. Cools, S. Delalieux, J. Stuckens, D. V. der Zande, W. W. Verstraeten, and P. Coppin, "Nonlinear hyperspectral mixture analysis for tree cover estimates in orchards," *Remote Sens. Environ.*, vol. 113, no. 6, pp. 1183–1193, Jun. 2009.
[11] T. Ray and B. Murray, "Nonlinear spectral mixing in desert vegetation," *Remote Sens. Environ.*, vol. 55, no. 1, pp. 59–64, Jan. 1996.
[12] L. Zhang, D. Li, Q. Tong, and L. Zheng, "Study of the spectral mixture model of soil and vegetation in PoYang Lake area, China," *Int. J. Remote Sens.*, vol. 19, no. 11, pp. 2077–2084, 1998.
[13] J. M. Bioucas-Dias and J. M. P. Nascimento, "Nonlinear mixture model for hyperspectral unmixing," in *Proc. SPIE Image Signal Process. Remote Sens. XV*, L. Bruzzone, C. Notarnicola, and F. Posa, Eds., 2009, vol. 7477, no. 1, p. 747 70I.
[14] J. Boardman, "Automating spectral unmixing of AVIRIS data using convex geometry concepts," in *Proc. Summ. 4th Annu. JPL Airborne Geosci. Workshop*, 1993, vol. 1, pp. 11–14.
[15] M. Winter, "Fast autonomous spectral end-member determination in hyperspectral data," in *Proc. 13th Int. Conf. Appl. Geologic Remote Sens.*, Vancouver, BC, Canada, Apr. 1999, vol. 2, pp. 337–344.
[16] J. M. Nascimento and J. M. Bioucas-Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 898–910, Apr. 2005.
[17] J. Settle, "On the relationship between spectral unmixing and subspace projection," *IEEE Trans. Geosci. Remote Sens.*, vol. 34, no. 4, pp. 1045–1046, Jul. 1996.
[18] N. Dobigeon, J.-Y. Tourneret, and C.-I Chang, "Semi-supervised linear spectral unmixing using a hierarchical Bayesian model for hyperspectral imagery," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 2684–2695, Jul. 2008.
[19] O. Eches, N. Dobigeon, C. Mailhes, and J.-Y. Tourneret, "Bayesian estimation of linear mixtures using the normal compositional model. Application to hyperspectral imagery," *IEEE Trans. Image Process.*, vol. 19, no. 6, pp. 1403–1413, Jun. 2010.
[20] J. Plaza, A. Plaza, P. Martinez, and R. Perez, "Nonlinear mixture models for analyzing laboratory simulated-forest hyperspectral data," *Proc. SPIE Image Signal Process. Remote Sens. IX*, vol. 5238, pp. 480–487, 2004.
[21] J. Plaza, P. Martinez, R. Perez, and A. Plaza, "Nonlinear neural network mixture models for fractional abundance estimation in AVIRIS hyperspectral images," in *Proc. XIII NASA/Jet Propulsion Lab. Airborne Earth Sci. Workshop*, Pasadena, CA, 2004.
[22] 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 Process.*, vol. 54, no. 11, pp. 4133–4145, Nov. 2006.
[23] S. Moussaoui, H. Hauksdottir, F. Schmidt, C. Jutten, J. Chanussot, D. Brie, S. Douté, and J. A. Benediktsson, "On the decomposition of mars hyperspectral data by ICA and Bayesian positive source separation," *Neurocomputing*, vol. 71, no. 10–12, pp. 2194–2208, Jun. 2008.
[24] O. Eches, N. Dobigeon, and J.-Y. Tourneret, "Estimating the number of endmembers in hyperspectral images using the normal compositional model and a hierarchical Bayesian algorithm," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 3, pp. 582–591, Jun. 2010.
[25] N. Dobigeon, S. Moussaoui, M. Coulon, J.-Y. Tourneret, and A. O. Hero, "Joint Bayesian endmember extraction and linear unmixing for hyperspectral imagery," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4355–4368, Nov. 2009.
[26] H. Jeffreys, "An invariant form for the prior probability in estimation problems," *Proc. R. Soc. Lond. A, Math. Phys. Sci.*, vol. 186, no. 1007, pp. 453–461, Sep. 1946.
[27] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*. New York: Springer-Verlag, 1999.
[28] J. Plaza, A. Plaza, R. Pérez, and P. Martinez, "Joint linear/nonlinear spectral unmixing of hyperspectral image data," in *Proc. IEEE IGARSS*, 2007, pp. 4037–4040.
[29] A. Halimi, Y. Altmann, N. Dobigeon, and J.-Y. Tourneret, Nonlinear unmixing of hyperspectral images using a generalized bilinear model, Univ. Toulouse, Toulouse, France, Tech. Rep. [Online]. Available: http://dobigeon.perso.enseeiht.fr/
[30] E. Christophe, D. Léger, and C. Mailhes, "Quality criteria benchmark for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 9, pp. 2103–2114, Sep. 2005.
[31] R. N. Clark, G. A. Swayze, and A. Gallagher, "Mapping minerals with imaging spectroscopy, U.S. Geological Survey," *Office Mineral Resour. Bull.*, vol. 2039, pp. 141–150, 1993.
[32] C. P. Robert, "Simulation of truncated normal variables," *Statist. Comput.*, vol. 5, no. 2, pp. 121–125, Jun. 1995.

Abderrahim Halimi was born in Algiers, Algeria, in 1987. He received the Engineering degree in electronics from the National Polytechnic School of Algiers, Algeria, in 2009 and the M.Sc. degree in signal processing from the Institut National Polytechnique de Toulouse, University of Toulouse, Toulouse, France, in 2010. He is currently working toward the Ph.D. degree in the Signal and Communication group of the IRIT Laboratory, University of Toulouse, France. His research interest are centered around hyperspectral imaging and altimetry.

Yoann Altmann was born in Toulouse, France, in 1987. He received the Engineering degree in electrical engineering from ENSEEIHT, University of Toulouse, Toulouse, and the M.Sc. degree in signal processing from Institut National Polytechnique de Toulouse, University of Toulouse, both in June 2010. He is currently working toward the Ph.D. degree in the Signal and Communication group of the IRIT Laboratory, University of Toulouse, France. His research interest are centered around hyperspectral imaging.

Nicolas Dobigeon (S'05–M'08) was born in Angoulême, France, in 1981. He received the Engineering degree in electrical engineering from the ENSEEIHT, University of Toulouse, Toulouse, France, in 2004 and the M.Sc. and Ph.D. degrees in signal processing from the Institut National Polytechnique de Toulouse, University of Toulouse, in 2004 and 2007, respectively.

From 2007 to 2008, he was a Postdoctoral Research Associate with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor. Since 2008, he has been an Assistant Professor with the Institut National Polytechnique de Toulouse, University of Toulouse, within the Signal and Communication Group, Institut de Recherche en Informatique de Toulouse. His research interests are focused on statistical signal and image processing, with particular interest in Bayesian inference and Markov chain Monte Carlo methods.

Jean-Yves Tourneret (SM'08) received the Ingénieur degree in electrical engineering from the École Nationale Supérieure d'Électronique, d'Électrotechnique, d'Informatique, d'Hydraulique, et des Télécommunications (ENSEEIHT), Institut National Polytechnique de Toulouse, University of Toulouse, Toulouse, France, in 1989 and the Ph.D. degree from the Institut National Polytechnique de Toulouse, University of Toulouse, in 1992.

He is currently a Professor with the University of Toulouse and a member of the Institut de Recherche en Informatique de Toulouse (UMR 5505 of the CNRS). His research activities are centered around statistical signal processing with a particular interest to Markov chain Monte Carlo methods.

Dr. Tourneret was the Program Chair of the European Conference on Signal Processing (EUSIPCO). He was also a member of the organizing committee of the IEEE International Conference on Acoustics, Speech and Signal Processing, which was held in Toulouse in 2006. He has been a member of different technical committees, including the Signal Processing Theory and Methods committee of the IEEE Signal Processing Society (in 2001–2007 and 2010–present). He is currently an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING.