BAYESIAN ALGORITHM FOR UNSUPERVISED UNMIXING OF HYPERSPECTRAL IMAGES USING A POST-NONLINEAR MODEL

Yoann Altmann, Nicolas Dobigeon and Jean-Yves Tourneret

University of Toulouse
IRIT/INP-ENSEEIHT/TéSA
Toulouse, France

ABSTRACT

This paper presents a nonlinear mixing model for hyperspectral image unmixing. The proposed model assumes that the pixel reflectances are post-nonlinear functions of unknown pure spectral components contaminated by an additive white Gaussian noise. The nonlinear effects are approximated by a polynomial leading to a polynomial post-nonlinear mixing model. A Bayesian algorithm is proposed to estimate the parameters involved in the model yielding an unsupervised nonlinear unmixing algorithm. Due to the large number of parameters to be estimated, an efficient constrained Hamiltonian Monte Carlo algorithm is investigated. The performance of the unmixing strategy is finally evaluated on synthetic data.

Index Terms— Hyperspectral imagery, unsupervised spectral unmixing, Hamiltonian Monte Carlo, post-nonlinear model.

1. INTRODUCTION

Identifying macroscopic materials and quantifying the proportions of these materials are major issues when analyzing hyperspectral images. This spectral unmixing (SU) problem has been widely studied for the applications where the pixel reflectances are linear combinations of pure component spectra. However, as explained in [1], the linear mixing model (LMM) can be inappropriate for some hyperspectral images. Nonlinear mixing models provide an interesting alternative for overcoming the inherent limitations of the LMM. Several models have been studied in the literature to handle specific kinds of nonlinearity. In particular, the bilinear models recently studied in [2–5] address the problem of scattering effects, mainly observed in vegetation or urban areas. Other more flexible unmixing techniques have been also proposed to handle wider class of nonlinearity, including radial basis function networks and kernel-based models. In this paper, we study a polynomial post-nonlinear mixing model (PPNMM) that has recently shown interesting properties for the SU of hyperspectral images [6].

More precisely, this paper presents a fully unsupervised unmixing algorithm based on the PPNMM, i.e., assumes that the endmembers are unknown. In the Bayesian framework, appropriate prior distributions are chosen for the unknown PPNMM parameters. However, the classical Bayesian estimators cannot be easily computed from the resulting joint posterior. To alleviate this problem, a Markov chain Monte Carlo (MCMC) method is used to generate samples according to the posterior of interest. Due to the large number of parameters to be estimated we propose to use a Hamiltonian Monte Carlo (HMC) method to sample according to the posterior. HMCs are powerful simulation strategies based on Hamiltonian dynamics which can improve the convergence and mixing properties of classical MCMC methods [7]. More recently, new HMCs methods have been proposed to handle constrained variables [7, Chap. 5] which allows HMCs to be applied to our Bayesian model.

The paper is organized as follows. Section 2 introduces the PPNMM for hyperspectral image analysis. Section 3 presents the hierarchical Bayesian model associated with the proposed PPNMM and its posterior distribution. Constrained HMC (CHMC) methods are coupled with a standard Gibbs sampler presented in Section 4. Simulation results conducted on synthetic data are shown and discussed in Section 5. Conclusions are finally reported in Section 6.

2. PROBLEM FORMULATION

2.1. Polynomial Post-Nonlinear Mixing Model

This section recalls the nonlinear mixing model used in [6] for hyperspectral image SU. We consider a set of N observed spectra $\mathbf{y}_n = [y_{n,1}, \ldots, y_{n,L}]^T, n \in \{1, \ldots, N\}$ where L is the number of spectral bands. Each of these spectra is defined as a nonlinear transformation $\boldsymbol{g}_n$ of a linear mixture of R spectra $\mathbf{m}_r$ contaminated by additive noise

$$\mathbf{y}_n = \boldsymbol{g}_n \left(\sum_{r=1}^{R} a_{r,n} \mathbf{m}_r \right) + \mathbf{e}_n = \boldsymbol{g}_n \left(\mathbf{M} \boldsymbol{a}_n \right) + \mathbf{e}_n \qquad (1)$$

Part of this work has been supported by Direction Generale de l'armement, French Ministry of defence, and by the Hypanema ANR Project ANR Project n∘ANR- 12-BS03-003.

where $\mathbf{m}_r = [m_{r,1}, \ldots, m_{r,L}]^T$ is the spectrum of the rth material present in the scene, $a_{r,n}$ is its corresponding proportion in the nth pixel, R is the number of endmembers contained in the image and $\boldsymbol{g}_n$ is a nonlinear function associated with the nth pixel. Moreover, $\mathbf{e}_n$ is an additive independent and identically distributed (i.i.d) zero-mean Gaussian noise sequence with variance σ^2, denoted as $\mathbf{e}_n \sim \mathcal{N}\left(\mathbf{0}_L, \sigma^2 \mathbf{I}_L\right)$. Note that the usual matrix and vector notations $\mathbf{M} = [\mathbf{m}_1, \ldots, \mathbf{m}_R]$ and $\boldsymbol{a}_n = [a_{1,n}, \ldots, a_{R,n}]^T$ have been used in the right hand side of (1). As in [6], the N nonlinear functions $\boldsymbol{g}_n$ are defined as second order polynomial nonlinearities defined by $\boldsymbol{g}_n(\mathbf{s}) = \mathbf{s} + b_n(\mathbf{s} \odot \mathbf{s})$, with $\mathbf{s} \in \mathbb{R}^L$, b_n a real parameter, and where $\odot$ denotes the Hadamard (termwise) product. An interesting property of the resulting PPNMM is that it reduces to the classical LMM for $b_n = 0$. Motivations for considering polynomial nonlinearities have been discussed in [6]. Straightforward computations allow the PPNMM observation matrix to be expressed as follows

$$\mathbf{Y} = \mathbf{MA} + [(\mathbf{MA}) \odot (\mathbf{MA})] \operatorname{diag}(\boldsymbol{b}) + \mathbf{E} \qquad (2)$$

where $\mathbf{A} = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_N]$ is an $R \times N$ matrix, $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_N]$ and $\mathbf{E} = [\mathbf{e}_1, \ldots, \mathbf{e}_N]$ are $L \times N$ matrices, and $\boldsymbol{b} = [b_1, \ldots, b_N]^T$ is an $N \times 1$ vector containing the nonlinearity parameters. Moreover, $\operatorname{diag}(\boldsymbol{b})$ is an $N \times N$ diagonal matrix containing the elements of the vector $\boldsymbol{b}$.

2.2. Abundance reparametrization

Due to physical considerations, the abundance vectors $\boldsymbol{a}_n$ satisfy the following positivity and sum-to-one constraints

$$\sum_{r=1}^{R} a_{r,n} = 1, \quad a_{r,n} > 0, \forall r \in \{1, \ldots, R\}. \qquad (3)$$

To handle these constraints, we propose to reparameterize the abundance vectors belonging to the set

$$\mathcal{S} = \left\{ \boldsymbol{a} = [a_1, \ldots, a_R]^T \,\middle|\, a_r > 0, \sum_{r=1}^{R} a_r = 1 \right\}$$

using the following transformation

$$a_{r,n} = \left(\prod_{k=1}^{r-1} z_{k,n} \right) \times \begin{cases} 1 - z_{r,n} & \text{if } r < R \\ 1 & \text{if } r = R \end{cases}. \qquad (4)$$

This transformation has been recently suggested in [8]. The main motivation for using the latent variables $z_{r,n}$ instead of $a_{r,n}$ is the fact that the constraints (3) for the nth abundance vector $\boldsymbol{a}_n$ express as

$$0 < z_{r,n} < 1, \quad \forall r \in \{1, \ldots, R-1\} \qquad (5)$$

for the nth coefficient vector $\boldsymbol{z}_n = [z_{1,n}, \ldots, z_{R-1,n}]^T$. As a consequence, the constraints (5) are much easier to handle for the sampling procedure than (3). The next section presents the Bayesian model associated with the PPNMM for SU.

3. BAYESIAN MODEL

This section generalizes the hierarchical Bayesian model introduced in [6] in order to jointly estimate the abundances and endmembers. The unknown parameter vector associated with the PPNMM contains the reparameterized abundances $\mathbf{Z} = [\boldsymbol{z}_1, \ldots, \boldsymbol{z}_N]$, the endmember matrix $\mathbf{M}$, the nonlinearity parameter vector $\boldsymbol{b}$ and the additive noise variance σ^2. This section summarizes the likelihood and the parameters priors introduced to perform unsupervised hyperspectral unmixing.

3.1. Likelihood

Assuming prior independence between the observed pixels and using (2), the joint likelihood of the observation matrix $\mathbf{Y}$ can be expressed as

$$f(\mathbf{Y}|\mathbf{M}, \mathbf{Z}, \boldsymbol{b}, \sigma^2) \propto \sigma^{-NL} \operatorname{etr}\left[-\frac{(\mathbf{Y} - \boldsymbol{X})^T (\mathbf{Y} - \boldsymbol{X})}{2\sigma^2} \right] \qquad (6)$$

where $\propto$ means "proportional to", $\operatorname{etr}(\cdot)$ denotes the exponential trace and $\boldsymbol{X} = \mathbf{MA} + [(\mathbf{MA}) \odot (\mathbf{MA})] \operatorname{diag}(\boldsymbol{b})$ is an $L \times N$ matrix.

3.2. Parameter priors

To reflect the lack of prior knowledge about the abundances, we propose to assign prior distributions for the coefficient vector $\boldsymbol{z}_n$ that correspond to noninformative prior distributions for $\boldsymbol{a}_n$. More precisely, assigning the beta priors $z_{n,r} \sim \mathcal{B}e(R - r, 1) \quad r \in \{1, \ldots, R-1\}$ and assuming prior independence between the elements of $\boldsymbol{z}_n$ yield an abundance vector $\boldsymbol{a}_n$ uniformly distributed in set $\mathcal{S}$ (see [8] for details). Assuming prior independence between the coefficient vectors $\{\boldsymbol{z}_n\}_{n=1,\ldots,N}$ leads to

$$f(\mathbf{Z}) = \prod_{r=1}^{R-1} \left\{ \frac{1}{B(R-r,1)^N} \prod_{n=1}^{N} z_{n,r}^{R-r-1} \right\} \qquad (7)$$

where $B(\cdot, \cdot)$ is the Beta function.

Each endmember $\mathbf{m}_r = [m_{r,1}, \ldots, m_{r,L}]^T$ is a reflectance vector satisfying the following constraints

$$0 \leq m_{r,\ell} \leq 1, \forall r \in \{1, \ldots, R\}, \forall \ell \in \{1, \ldots, L\}. \qquad (8)$$

For each endmember $\mathbf{m}_r$, we propose to use a Gaussian prior

$$\mathbf{m}_r \sim \mathcal{N}_{[0,1]^L}(\bar{\mathbf{m}}_r, s^2 \mathbf{I}_L), \qquad (9)$$

truncated on $[0, 1]^L$ to satisfy the constraints (8). In this paper, we propose to select the mean vectors $\bar{\mathbf{m}}_r$ as the pure components previously identified by the nonlinear EEA studied in [9] and referred to as "Heylen". The variance s^2 reflects the degree of confidence given to this prior information. When no additional knowledge is available, this variance is fixed to a large value ($s^2 = 50$ in our simulations).

The PPNMM reduces to the LMM for $b_n = 0$. Since the LMM is probably relevant for most observed pixels it makes sense to assign prior distributions to the nonlinearity parameters that enforce sparsity for the vector $\boldsymbol{b}$. Consequently, the following conjugate Bernoulli-Gaussian prior is assigned to each parameter b_n

$$f(b_n|w,\sigma_b^2) = (1-w)\delta(b_n) + w\frac{1}{\sqrt{2\pi\sigma_b^2}}\exp\left(-\frac{b_n^2}{2\sigma_b^2}\right) \quad (10)$$

where $\delta(\cdot)$ denotes the Dirac delta function. Note that the prior distributions for $\{b_n\}_{n=1,\ldots,N}$ share the same hyperparameters $w \in [0,1]$ and $\sigma_b^2 \in]0,+\infty[$. Moreover, the weight w is the prior probability of having a nonlinearly mixed pixel in the image. Assuming prior independence between the nonlinearity parameters $\{b_n\}_{n=1,\ldots,N}$, the joint prior distribution of the nonlinearity parameter vector $\boldsymbol{b}$ is given by

$$f(\boldsymbol{b}|w,\sigma_b^2) = \prod_{n=1}^{N} f(b_n|w,\sigma_b^2). \quad (11)$$

A Jeffreys' prior is chosen for the noise variance σ^2

$$f(\sigma^2) \propto \frac{1}{\sigma^2}\mathbf{I}_{\mathbb{R}^+}(\sigma^2) \quad (12)$$

which reflects the absence of knowledge for this parameter.

3.3. Hyperparameter priors

The performance of the proposed Bayesian model for spectral unmixing depends on the values of the hyperparameters σ_b^2 and w. When the hyperparameters are difficult to adjust, it is classical to include them in the unknown parameter vector, resulting in a hierarchical Bayesian model [6, 10]. A conjugate inverse-Gamma prior is assigned to σ_b^2, i.e., $\sigma_b^2 \sim \mathcal{IG}(\gamma,\nu)$ where (γ,ν) are real parameters fixed to obtain a flat prior, reflecting the absence of knowledge about the variance σ_b^2 ((γ,ν) will be set to $(10^{-1},10^{-1})$ in the simulation section). A uniform prior distribution is assigned to the hyperparameter w, i.e., $w \sim \mathcal{U}_{[0,1]}(w)$ since there is no a priori information regarding the proportions of linearly and nonlinearly mixed pixels in the image.

3.4. Joint Posterior distribution

The joint posterior distribution of the unknown parameters $\boldsymbol{\theta} = \left\{\mathbf{Z},\mathbf{M},\boldsymbol{b},\sigma^2,\sigma_b^2,w\right\}$ can be computed using the following hierarchical structure

$$f(\boldsymbol{\theta}|\mathbf{Y}) \propto f(\mathbf{Y}|\boldsymbol{\theta})f(\boldsymbol{\theta}) \quad (13)$$

where $f(\mathbf{Y}|\boldsymbol{\theta})$ has been defined in (6). By assuming *a priori* independence between the parameters $\mathbf{Z}$, $\mathbf{M}$, $\boldsymbol{b}$ and σ^2 and between the hyperparameters σ_b and w, the joint prior distribution of the $\boldsymbol{\theta}$ can be expressed as

$$f(\boldsymbol{\theta}) = f(\mathbf{Z})f(\mathbf{M})f(\sigma^2)f(\boldsymbol{b}|\sigma_b^2,w)f(\sigma_b^2)f(w). \quad (14)$$

Unfortunately, it is difficult to obtain closed form expressions for the standard Bayesian estimators associated with (13). In this paper, we propose to use efficient MCMC methods to generate samples asymptotically distributed according to (13). Due to the large number of parameters to be sampled, we use an HMC algorithm which allows the number of sampling steps to be reduced and which improves the mixing properties of the sampler. The basic principles of the HMC methods that will be used to sample asymptotically from (13) can be found in [11]. The generated samples are then used to compute the MMSE estimator of $\boldsymbol{\theta}$. The next section summarizes the Gibbs sampler including constrained HMC methods used to sample from (13).

4. GIBBS SAMPLER

The principle of the Gibbs sampler is to sample according to the conditional distributions of the posterior of interest [12, Chap. 10]. Due to the large number of parameters to be estimated, it makes sense to use a block Gibbs sampler to improve the convergence of the sampling procedure. More precisely, we propose to sample sequentially $\mathbf{M},\mathbf{Z},\boldsymbol{b},\sigma^2,\sigma_b^2$ and w using six moves that are detailed in the next sections.

4.1. Sampling the coefficient matrix $\mathbf{Z}$

Sampling from $f(\mathbf{Z}|\mathbf{Y},\mathbf{M},\boldsymbol{b},\sigma^2,\sigma_b^2,w)$ is difficult due to the complexity of this distribution. In this case, it is classical to use an accept/reject procedure to update the coefficient matrix $\mathbf{Z}$ (leading to a hybrid Metropolis-Within-Gibbs sampler). It can be shown that

$$f(\mathbf{Z}|\mathbf{Y},\mathbf{M},\boldsymbol{b},\sigma^2,\sigma_b,w) = \prod_{n=1}^{N} f(\boldsymbol{z}_n|\mathbf{y}_n,\mathbf{M},b_n,\sigma^2), \quad (15)$$

i.e., the N coefficients vectors $\{\boldsymbol{z}_n\}_{n=1,\ldots,N}$ are a posteriori independent and can be sampled independently in a parallel manner. Straightforward computations lead to

$$f(\boldsymbol{z}_n|\mathbf{y}_n,\mathbf{M},b_n,\sigma^2) \propto \exp\left(-\frac{\|\mathbf{y}_n-\boldsymbol{x}_n\|^2}{2\sigma^2}\right) \times \mathbf{1}_{(0,1)^{R-1}}(\boldsymbol{z}_n)\prod_{r}^{R-1} z_{n,r}^{R-r-1} \quad (16)$$

where $\boldsymbol{x}_n = \boldsymbol{g}_n(\mathbf{M}\boldsymbol{a}_n)$, $\mathbf{1}_{(0,1)^{R-1}}(\cdot)$ denotes the indicator function over $(0,1)^{R-1}$. The distribution (16) can be related to a potential energy that is then used within the CHMC method presented in [11] to update the vector $\boldsymbol{z}_n$ (see [11] for details).

4.2. Sampling the endmember matrix $\mathbf{M}$

From (13) and (14), it can be seen that

$$f(\mathbf{M}|\mathbf{Y},\mathbf{Z},\boldsymbol{b},\sigma^2,s^2,\widetilde{\mathbf{M}}) = \prod_{\ell=1}^{L} f(\mathbf{m}_{\ell,:}|\mathbf{y}_{\ell,:},\mathbf{Z},\boldsymbol{b},\sigma^2,s^2,\bar{\mathbf{m}}_{\ell,:})$$

where $\mathbf{m}_{\ell,:}$ (resp. $\bar{\mathbf{m}}_{\ell,:}$ and $\mathbf{y}_{\ell,:}$) is the ℓth row of $\mathbf{M}$ (resp. of $\widetilde{\mathbf{M}}$ and $\mathbf{Y}$) and

$$f(\mathbf{m}_{\ell,:}|\mathbf{y}_{\ell,:},\mathbf{Z},\boldsymbol{b},\sigma^2,s^2,\bar{\mathbf{m}}_{\ell,:}) \propto \exp\left(-\frac{\|\mathbf{y}_{\ell,:}-\boldsymbol{t}_\ell\|^2}{2\sigma^2}\right) \times \exp\left(-\frac{\|\mathbf{m}_{\ell,:}-\bar{\mathbf{m}}_{\ell,:}\|^2}{2s^2}\right)\mathbf{1}_{(0,1)^R}(\mathbf{m}_{\ell,:}) \quad (17)$$

with $\boldsymbol{t}_\ell = \mathbf{A}^T\mathbf{m}_{\ell,:} + \text{diag}(\boldsymbol{b})\left[\left(\mathbf{A}^T\mathbf{m}_{\ell,:}\right)\odot\left(\mathbf{A}^T\mathbf{m}_{\ell,:}\right)\right]$. Consequently, the rows of the endmember matrix $\mathbf{M}$ can be sampled independently similarly to the CHMC procedure described in the previous section to sample $\mathbf{Z}$ by introducing the L potential energies associated with each $\mathbf{m}_{\ell,:}$ (see [11] for details).

4.3. Sampling the nonlinearity parameter vector b

Using (13) and (14), it can be easily shown that the conditional distribution of $b_n|\mathbf{y}_n,\mathbf{M},\boldsymbol{z}_n,\sigma^2,w,\sigma_b^2$ is the following Bernoulli-Gaussian distribution

$$b_n|\mathbf{y}_n,\mathbf{M},\boldsymbol{z}_n,\sigma^2,w,\sigma_b^2 \sim (1-w_n^*)\delta(b_n) + w_n^*\mathcal{N}\left(\mu_n,s_n^2\right) \quad (18)$$

where

$$\mu_n = \frac{\sigma_b^2\left(\mathbf{y}_n-\mathbf{M}\boldsymbol{a}_n\right)^T\boldsymbol{h}_n}{\sigma_b^2\boldsymbol{h}_n^T\boldsymbol{h}_n+\sigma^2}, \quad s_n^2 = \frac{\sigma_b^2\sigma^2}{\sigma_b^2\boldsymbol{h}_n^T\boldsymbol{h}_n+\sigma^2}$$

and $\boldsymbol{h}_n = (\mathbf{M}\boldsymbol{a}_n)\odot(\mathbf{M}\boldsymbol{a}_n)$. Moreover,

$$w_n^* = \frac{w}{\beta_n+w(1-\beta_n)}, \quad \beta_n = \frac{\sigma_b}{s_n}\exp\left(-\frac{\mu_n^2}{2s_n^2}\right). \quad (19)$$

For each b_n, the conditional distribution (18) does not depend on $\{b_k\}_{k\neq n}$. Consequently, the nonlinearity parameters $\{b_n\}_{n=1,\ldots,N}$ can be sampled independently in a parallel manner.

4.4. Sampling the noise variance σ^2

By considering the posterior distribution (13), it can be shown that $\sigma^2|\mathbf{Y},\mathbf{M},\mathbf{Z},\boldsymbol{b}$ is distributed according to the following inverse-gamma distribution

$$\sigma^2|\mathbf{Y},\mathbf{M},\mathbf{Z},\boldsymbol{b} \sim \mathcal{IG}\left(\frac{NL}{2},\frac{\text{tr}\left((\mathbf{Y}-\boldsymbol{X})^T(\mathbf{Y}-\boldsymbol{X})\right)}{2}\right) \quad (20)$$

with $\text{tr}(\cdot)$ the matrix trace, from which it is easy to sample.

4.5. Sampling the hyperparameters σ_b^2 and w

Looking carefully at the posterior distribution (13), it can be seen that $\sigma_b^2|\boldsymbol{b},\gamma,\nu$ is distributed according to the following inverse-gamma distribution

$$\sigma_b^2|\boldsymbol{b},\gamma,\nu \sim \mathcal{IG}\left(\frac{n_1}{2}+\gamma,\sum_{n\in I_1}\frac{b_n^2}{2}+\nu\right) \quad (21)$$

with $I_1=\{n|b_n\neq 0\}$, $n_0=\|\boldsymbol{b}\|_0$ (where $\|\cdot\|_0$ is the ℓ_0 norm, i.e., the number of elements of $\boldsymbol{b}$ that are different from zero) and $n_1=N-n_0$, from which it is easy to sample. Similarly, we obtain

$$w|\boldsymbol{b} \sim \mathcal{B}e(n_1+1,n_0+1). \quad (22)$$

The small number of sampling steps is due to the high parallelization properties of the proposed sampling procedure, i.e., the generation of the N coefficient vectors $\{\boldsymbol{z}_n\}_{n=1,\ldots,N}$, the N nonlinearity parameters $\{b_n\}_{n=1,\ldots,N}$ and the L reflectance vectors $\{\mathbf{m}_{\ell,:}\}_{\ell=1,\ldots,L}$. After generating N_{MC} samples using the procedures detailed above, the MMSE estimator of the unknown parameters can be approximated by computing the empirical averages of these samples, after an appropriate burn-in period[1]. The next section studies the performance of the proposed algorithm for synthetic hyperspectral images.

5. SIMULATIONS

The performance of the proposed nonlinear SU algorithm is first evaluated by unmixing 3 synthetic images I_1 to I_3 of size $N=2500$ pixels. The $R=3$ endmembers observed at $L=207$ different spectral bands and contained in these images have been extracted from the spectral libraries provided with the ENVI software. The first image I_1 has been generated using the LMM. The image I_2 has been generated according to the PPNMM and I_3 has been generated according to the generalized bilinear mixing model (GBM) presented in [5]. For each image, the abundance vectors have been randomly generated according to a uniform distribution in the admissible set defined by $\mathcal{S}_{\text{t}}=\left\{\boldsymbol{a}\left|0<a_r<0.9,\sum_{r=1}^R a_r=1\right.\right\}$ to ensure that there is no pure pixel in the images. All images have been corrupted by an i.i.d Gaussian noise of variance $\sigma^2=10^{-4}$, corresponding to an average signal-to-noise ratio $\text{SNR}\simeq 21$dB for the three images. The nonlinearity coefficients are uniformly drawn in the set $[0,1]$ for the GBM. The parameters b_n have been generated uniformly in the set $[-0.3,0.3]$ for the PPNMM.

Different estimation procedures have been considered for the three mixing models. For the LMM, two unmixing algorithms have been considered for the LMM. The first strategy extracts the endmembers using the N-FINDR algorithm [14] and estimates the abundances using the FCLS algorithm [15] (it is referred to as "SLMM" for supervised LMM). The second strategy is a Bayesian algorithm which jointly estimates the endmembers and the abundance matrix [10] (it is referred to as "ULMM" for unsupervised LMM). Two approaches have also been considered for the PPNMM. The first strategy uses the nonlinear EEA studied in [9] and the gradient-based approach based on the PPNMM studied in [6]

[1] The length of the burn-in period has been determined using appropriate convergence diagnoses [13].

for estimating the abundances and the nonlinearity parameter. This strategy is referred to as "SPPNMM" (supervised PPNMM). The second strategy is the proposed unmixing procedure referred to as "UPPNMM" (unsupervised PPNMM). The unmixing strategy used for the GBM is the EEA studied in [9] and the Bayesian algorithm presented in [5] for abundance estimation.

The quality of the unmixing procedures can be measured by the root normalized mean square error (RNMSE) defined by RNMSE $= \sqrt{\sum_{n=1}^{N} \|\hat{\boldsymbol{a}}_n - \boldsymbol{a}_n\|^2/(NR)}$, where $\boldsymbol{a}_n$ and $\hat{\boldsymbol{a}}_n$ are the actual and estimated abundance vectors for the nth pixel of the image. Table 1 shows the RNMSEs associated with the images I_1 to I_3 for the different estimation methods. These results show that the UPPNMM performs better (in term of RNMSE) than the other considered unmixing methods for the three images. Moreover, the proposed method provides similar results when compared with the ULMM for the linearly mixed image I_1. The quality of endmember estimation is evaluated by the spectral angle mapper (SAM) defined as SAM $= \arccos\left(\frac{\langle \hat{\mathbf{m}}_r, \mathbf{m}_r \rangle}{\|\hat{\mathbf{m}}_r\| \, \|\mathbf{m}_r\|}\right)$, where $\mathbf{m}_r$ is the rth actual endmember and $\hat{\mathbf{m}}_r$ its estimate. Table 2 compares the performance of the different endmember estimation algorithms using the SAM (averaged over the $R = 3$ endmembers (ASAM)). This table shows that the proposed UPPNMM provides more accurate endmembers estimates than the others methods. Moreover, these results illustrates the robustness of the PPNMM regarding model mis-specification. Note that the ULMM and the UPPNMM provide similar results (in term of ASAMs) for the image I_1 generated according to the LMM. Additional simulation results including reconstruction performance and simulations with different number of endmembers can be found in [11].

Table 1. Abundance RNMSEs ($\times 10^{-2}$): synthetic images.

		I_1 (LMM)	I_2 (PPNMM)	I_3 (GBM)
LMM	SLMM	3.78	13.21	6.83
	ULMM	0.66	10.87	4.21
PPNMM	SPPNMM	4.18	6.04	4.13
	UPPNMM	**0.37**	**0.81**	**1.38**
GBM		4.18	11.15	5.02

Table 2. SAMs ($\times 10^{-2}$): synthetic images.

	N-Findr	ULMM	Heylen	UPPNMM
I_1	4.95	0.52	6.38	**0.42**
I_2	7.44	8.23	7.92	**0.39**
I_3	7.46	4.66	7.19	**1.63**

6. CONCLUSIONS AND FUTURE WORK

We proposed a new hierarchical Bayesian algorithm for unsupervised nonlinear spectral unmixing of hyperspectral images. This algorithm assumed that each pixel of the image is a post-nonlinear mixture of the endmembers contaminated by additive Gaussian noise. The physical constraints for the abundances and endmembers was included in the Bayesian framework through appropriate prior distributions. Due to the complexity of the resulting joint posterior distribution and the number of parameters to be estimated, a constrained Hamiltonian Monte Carlo method was used to approximate the MMSE estimator of the unknown model parameters. Simulations conducted on synthetic data illustrated the flexibility of the proposed model for linear and nonlinear spectral unmixing and provided promising results. An important advantage of the proposed algorithm is its flexibility regarding the absence of pure pixels in the image. Future work includes the estimation of number of endmembers, which was assumed to be known in this paper.

7. REFERENCES

[1] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observations Remote Sensing*, vol. 5, no. 2, pp. 354–379, April 2012.

[2] B. Somers, K. Cools, S. Delalieux, J. Stuckens, D. Van der Zande, W. W. Verstraeten, and P. Coppin, "Nonlinear hyperspectral mixture analysis for tree cover estimates in orchards," *Remote Sensing of Environment*, vol. 113, no. 6, pp. 1183–1193, 2009.

[3] J. M. P. Nascimento and J. M. Bioucas-Dias, "Nonlinear mixture model for hyperspectral unmixing," in *Proc. SPIE Image and Signal Processing for Remote Sensing XV*, L. Bruzzone, C. Notarnicola, and F. Posa, Eds. 2009, vol. 7477, p. 74770I, SPIE.

[4] 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," *Remote Sensing of Environment*, vol. 30, no. 11, pp. 2951–2962, June 2009.

[5] A. Halimi, Y. Altmann, N. Dobigeon, and J.-Y. Tourneret, "Nonlinear unmixing of hyperspectral images using a generalized bilinear model," *IEEE Trans. Geosci. and Remote Sensing*, vol. 49, no. 11, pp. 4153–4162, nov. 2011.

[6] Y. Altmann, A. Halimi, N. Dobigeon, and J. Tourneret, "Supervised nonlinear spectral unmixing using a postnonlinear mixing model for hyperspectral imagery," *IEEE Trans. Image Processing*, vol. 21, no. 6, pp. 3017–3025, June 2012.

[7] S. Brooks, *Handbook of Markov Chain Monte Carlo*, Chapman & Hall/CRC Handbooks of Modern Statistical Methods. Taylor & Francis, 2011.

[8] M. J. Betancourt, "Cruising The Simplex: Hamiltonian Monte Carlo and the Dirichlet Distribution," *ArXiv e-prints*, Oct. 2010.

[9] R. Heylen, D. Burazerovic, and P. Scheunders, "Non-linear spectral unmixing by geodesic simplex volume maximization," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 3, pp. 534 –542, june 2011.

[10] 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. 2657–2669, Nov. 2009.

[11] Yoann Altmann, Nicolas Dobigeon, and Jean-Yves Tourneret, "Bayesian algorithm for unsupervised unmixing of hyperspectral images using a post-nonlinear model.," Tech. Rep., University of Toulouse, France, March 2013.

[12] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*, Springer-Verlag, New York, second edition, 2004.

[13] C. P. Robert and D. Cellier, "Convergence control of MCMC algorithms," in *Discretization and MCMC Convergence Assessment*, C. P. Robert, Ed., pp. 27–46. Springer Verlag, New York, 1998.

[14] M.E. Winter, "Fast autonomous spectral end-member determination in hyperspectral data," in *Proc. 13th Int. Conf. on Applied Geologic Remote Sensing*, Vancouver, April 1999, vol. 2, pp. 337–344.

[15] 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.