

Estimating the number of endmembers in hyperspectral images using the normal compositional model and a hierarchical Bayesian algorithm

Olivier Eches, Nicolas Dobigeon and Jean-Yves Tourneret

Abstract

This paper studies a semi-supervised Bayesian unmixing algorithm for hyperspectral images. This algorithm is based on the normal compositional model recently introduced by Eismann and Stein. The normal compositional model assumes that each pixel of the image is modeled as a linear combination of an unknown number of pure materials, called *endmembers*. These endmembers are supposed to be random in order to model uncertainties regarding their knowledge. More precisely, we model endmembers as Gaussian vectors whose means belong to a known library. However, all endmembers of this library do not concur to the linear mixture associated to a given pixel. This paper proposes to estimate the mixture coefficients (referred to as *abundances*) as well as their number using a reversible jump Bayesian algorithm that will solve the model selection problem. Suitable priors are assigned to the abundances in order to satisfy positivity and sum-to-one constraints whereas conjugate priors are chosen for the other parameters. However, the resulting posterior distribution is too complicated to yield closed-form expressions for the usual Bayesian estimators (including the maximum a posteriori and the minimum mean square error estimators). This paper proposes to generate samples distributed according to the posterior of interest using an appropriate Markov chain Monte Carlo method. The generated samples are then used to compute the Bayesian estimators of the unknown parameters. As the number of endmembers is unknown, the sampler has to jump between spaces of different dimensions. This is achieved by a reversible jump Markov chain Monte Carlo method that allows one to handle the model order selection problem. The performance of the proposed methodology is evaluated thanks to simulations conducted on synthetic and real AVIRIS images.

Index Terms

Bayesian inference, Monte Carlo methods, spectral unmixing, hyperspectral images, normal compositional model, reversible jump.

The authors are with University of Toulouse, IRIT/INP-ENSEEIH/TéSA, 2 rue Charles Camichel, BP 7122, 31071 Toulouse cedex 7, France (e-mail: {olivier.eches, nicolas.dobigeon, jean-yves.tourneret}@enseeiht.fr).

I. INTRODUCTION

For several decades, hyperspectral imagery has been receiving growing interest in the signal and image processing literature (see [1] and references therein). This interest can be easily explained by the high spectral resolution of the images provided by the hyperspectral sensors such as AVIRIS [2] and Hyperion [3]. Spectral unmixing is a crucial step in the analysis of these images [4]. It consists of decomposing the measured pixel reflectances into a mixture of pure spectra whose fractions, referred to as abundances, have to be estimated. Most unmixing procedures for hyperspectral images assume that the image pixels are linear combinations of pure materials. More precisely, according to the linear mixing model (LMM) presented in [4], the L -spectrum $\mathbf{y} = [y_1, \dots, y_L]^T$ of a mixed pixel is modeled as a linear combination of R spectra \mathbf{m}_r corrupted by additive white Gaussian noise

$$\mathbf{y} = \sum_{r=1}^R \mathbf{m}_r \alpha_r + \mathbf{n} \quad (1)$$

where $\mathbf{m}_r = [m_{r,1}, \dots, m_{r,L}]^T$ denotes the spectrum of the r^{th} material (referred to as endmember), α_r is the fraction of the r^{th} material in the pixel (referred to as abundance), R is the number of pure materials present in the observed scene and L is the number of available spectral bands for the image. In addition, due to obvious physical considerations, the abundances satisfy the following positivity and sum-to-one constraints

$$\begin{cases} \alpha_r \geq 0, \quad \forall r = 1, \dots, R \\ \sum_{r=1}^R \alpha_r = 1. \end{cases} \quad (2)$$

Supervised algorithms assume that the R endmember spectra \mathbf{m}_r are known, e.g., extracted from a spectral library. In practical applications, they can be obtained by an endmember extraction procedure such as the well-known N-finder (N-FINDR) algorithm developed by Winter [5] or the Vertex Component Analysis (VCA) proposed by Nascimento *et al.* [6]. However, the LMM has some limitations when applied to real images [4]. For instance, the endmember extraction procedures based on the LMM can be inefficient when the image does not contain enough pure pixels. This problem, also outlined by Nascimento in [6], is illustrated in Fig. 1. This figure shows i) the projections on the two most discriminant axes identified by a principal component analysis (PCA) of $R = 3$ endmembers (red stars corresponding to the vertices of the red triangle), ii) the domain containing all linear combinations of the $R = 3$ endmembers (i.e. the red triangle),

and iii) the simplex estimated by the N-FINDR algorithm using the black pixels. As there is no pixel close to the vertices of the red triangle, the N-FINDR estimates a much smaller simplex (in blue) than the actual one (in red).

An alternative model referred to as normal compositional model (NCM) has been recently proposed in [7]. The NCM allows one to alleviate the problems mentioned above by assuming that the pixels of the hyperspectral image are linear combinations of random endmembers (as opposed to deterministic for the LMM) with known means (e.g., resulting from the N-FINDR algorithm). This model provides more flexibility regarding the observed pixels and the endmembers. In particular, the endmembers are allowed to be further from the observed pixels which is clearly an interesting property for the problem illustrated in Fig. 1. The NCM assumes that the spectrum of a pixel is defined by the following mixture

$$\mathbf{y} = \sum_{r=1}^R \boldsymbol{\varepsilon}_r \alpha_r \quad (3)$$

where the $\boldsymbol{\varepsilon}_r$ are independent Gaussian vectors with known means extracted from a spectral library or estimated by the N-FINDR algorithm. Note that there is no additive noise in (3) since the random nature of the endmembers already models some kind of uncertainty regarding the endmembers. This paper assumes that the covariance matrix of each endmember can be written $\sigma^2 \mathbf{I}_L$, where \mathbf{I}_L is the $L \times L$ identity matrix and σ^2 is the endmember variance in any spectral band. More sophisticated models with different variances in the spectral bands could be investigated¹. For simplicity, this paper assumes a common variance in all spectral bands that has been considered successfully in many studies [10]–[13].

A supervised spectral unmixing strategy has been proposed in [14] when the number of endmembers R participating in the mixture model (3) is known. This paper addresses the important problem of estimating this number of endmembers R and the nature of these endmembers. It is a model selection problem since different values of R provide abundance vectors with different sizes, i.e., different models. The proposed algorithm concentrates on the NCM because of the interesting properties mentioned above. However, it could be used with appropriate modifications to determine the number of component of the LMM (1) (see [15] for more details). We assume

¹The case of a structured covariance matrix having K different variances has been considered in [8]. More recently, the results have been generalised to any colored noise [9].

that the endmember means in (3) belong to a given spectral library. However, contrary to the approach in [14], the algorithm does not know the nature and the number of endmember means from this library participating in the NCM. The algorithm is referred to as “semi-supervised” to reflect the partial knowledge about the endmembers present in the NCM mixture. The proposed strategy relies on a hierarchical Bayesian model. Appropriate prior distributions are chosen for the NCM abundances to satisfy the positivity and sum-to-one constraints, as in [15]. A vague conjugate inverse Gamma distribution is defined for the endmember variance reflecting the lack of knowledge regarding this parameter. A prior distribution for the number of endmembers R is also introduced. The proposed algorithm is hierarchical since it allows one to estimate the hyperparameter of the NCM. A vague prior is assigned to this hyperparameter. The parameters and hyperparameter of the resulting hierarchical Bayesian model are then jointly estimated using the full posterior distribution. Unfortunately the joint posterior distribution for the NCM is too complex to derive closed-form expressions for the standard minimum mean square error (MMSE) or maximum *a posteriori* (MAP) estimators. In particular, the dimension of the abundance vector depends on the unknown number R of endmembers involved in the mixture. Therefore, we propose to use a reversible jump Markov Chain Monte Carlo (MCMC) algorithm [16] to generate samples distributed according to the posterior of interest. This dimension matching strategy proposes moves between parameter spaces with different dimensions. It has been successfully used in many signal and image processing applications including segmentation [17], [18], analysis of musical audio data [19] and spectral analysis [20].

The paper is organized as follows. Section II derives the posterior distribution of the unknown parameter vector resulting from the NCM. Section III studies the reversible jump MCMC sampling strategy that is used to generate samples distributed according to this posterior and to solve the model selection problem. Simulation results conducted on synthetic and real data as well as a comparison with other model selection techniques are presented in Sections IV and V. Conclusions are reported in Section VI.

II. HIERARCHICAL BAYESIAN MODEL

This section studies the likelihood and the priors inherent to the proposed NCM for the spectral unmixing of hyperspectral images. A particular attention is devoted to the unknown number R of endmembers, and to the positivity and sum-to-one constraints about the abundances.

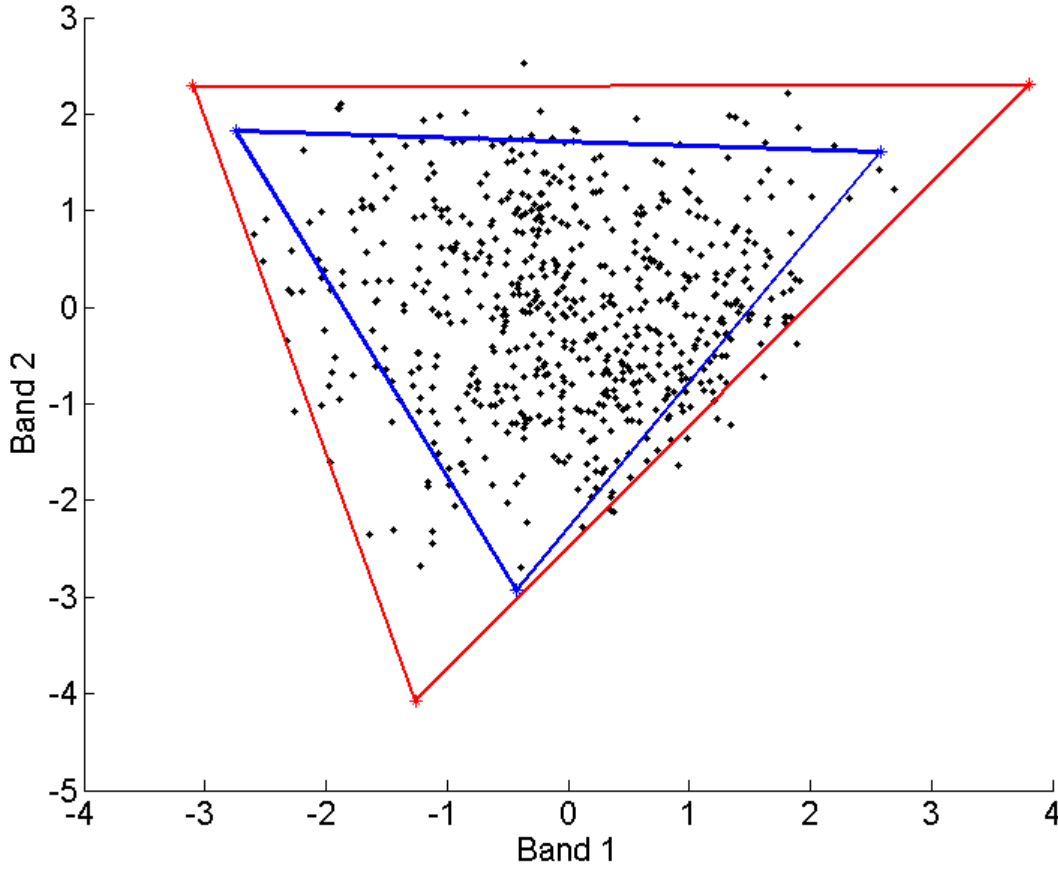


Fig. 1. Scatterplot of dual-band correct (red) and incorrect (blue) results of the N-FINDR algorithm.

A. Likelihood

The NCM assumes that the endmembers ε_r , $r = 1, \dots, R$, are independent Gaussian vectors whose means are assumed to belong to a known spectral library $\mathcal{S} = \{\mathbf{s}_1, \dots, \mathbf{s}_{R_{\max}}\}$ (where \mathbf{s}_r represents the L -spectrum $[s_{r,1}, \dots, s_{r,L}]^T$ of the r th endmember). Moreover, this paper assumes that the endmember spectrum components are independent from one band to another, i.e.

$$\varepsilon_r | \mathbf{m}_r, \sigma^2 \sim \mathcal{N}(\mathbf{m}_r, \sigma^2 \mathbf{I}_L)$$

where $\mathbf{m}_r = [m_{r,1}, \dots, m_{r,L}]^T$ is the mean vector of ε_r , $\sigma^2 \mathbf{I}_L$ is its covariance matrix and $\mathcal{N}(\mathbf{m}, \Sigma)$ denotes the Gaussian distribution with mean vector \mathbf{m} and covariance matrix Σ . However, the number of components R , as well as the spectra involved in the mixture are

unknown. Using the NCM definition (3) and the independence between the endmembers, the likelihood of the observed pixel \mathbf{y} can be written as

$$f(\mathbf{y} | \boldsymbol{\alpha}^+, \sigma^2, R, \mathbf{M}) = \frac{1}{[2\pi\sigma^2c(\boldsymbol{\alpha}^+)]^{\frac{L}{2}}} \exp\left(-\frac{\|\mathbf{y} - \mu(\boldsymbol{\alpha}^+)\|^2}{2\sigma^2c(\boldsymbol{\alpha}^+)}\right) \quad (4)$$

where $\|\mathbf{x}\| = \sqrt{\mathbf{x}^T \mathbf{x}}$ is the standard ℓ^2 norm, $\boldsymbol{\alpha}^+ = [\alpha_1, \dots, \alpha_R]^T$, $\mathbf{M} = [\mathbf{m}_1, \dots, \mathbf{m}_R]$ contains the endmember defining the mixture (3) and

$$\mu(\boldsymbol{\alpha}^+) = \sum_{r=1}^R \mathbf{m}_r \alpha_r, \quad c(\boldsymbol{\alpha}^+) = \sum_{r=1}^R \alpha_r^2. \quad (5)$$

Note that the mean and variance of this Gaussian distribution depend both on the abundance vector $\boldsymbol{\alpha}^+$ contrary to the classical LMM. Note also that the dimensions of $\boldsymbol{\alpha}^+$ and \mathbf{M} (via the quantities $\mu(\boldsymbol{\alpha}^+)$ and $c(\boldsymbol{\alpha}^+)$) depend on the unknown number of endmembers R .

B. Parameter priors

1) *Prior distributions for the number and mean of the endmembers:* A discrete uniform distribution on $[1, \dots, R_{\max}]$ is chosen as prior distribution for the number of mixture components R

$$P(R = k) = \frac{1}{R_{\max}}, \quad k = 1, \dots, R_{\max} \quad (6)$$

where R_{\max} is the maximum number of pure elements that can be present in a pixel. Note that this uniform prior distribution does not favor any model order among $[1, \dots, R_{\max}]$.

All combinations of R spectra belonging to the library \mathcal{S} are assumed to be equiprobable conditionally upon the number of endmembers R leading to

$$P(\mathbf{M} = [\mathbf{s}_{i_1}, \dots, \mathbf{s}_{i_R}] | R) = \binom{R_{\max}}{R}^{-1} = \frac{\Gamma(R+1)\Gamma(R_{\max}-R+1)}{\Gamma(R_{\max}+1)} \quad (7)$$

where i_1, \dots, i_R is an unordered sequence of R distinct elements of $[1, \dots, R_{\max}]$.

2) *Abundance prior:* Because of the sum-to-one constraint inherent to the mixing model, the abundance vector can be split into two parts as $\boldsymbol{\alpha}^+ = [\boldsymbol{\alpha}^T, \alpha_R]^T$ with $\alpha_R = 1 - \sum_{r=1}^{R-1} \alpha_r$. Moreover, to satisfy the positivity constraint, the abundance vector $\boldsymbol{\alpha}$ belongs to a simplex defined by

$$\mathbb{S} = \left\{ \boldsymbol{\alpha} \mid \alpha_r \geq 0, \forall r = 1, \dots, R-1, \sum_{r=1}^{R-1} \alpha_r \leq 1 \right\}. \quad (8)$$

A uniform distribution on this simplex is chosen as prior distribution for the partial abundance vector α

$$f(\alpha|R) \propto \mathbf{1}_{\mathbb{S}}(\alpha) \quad (9)$$

where \propto means “proportional to” and $\mathbf{1}_{\mathbb{S}}(\cdot)$ is the indicator function defined on the set \mathbb{S} . This prior ensures the positivity and sum-to-one constraints of the abundance coefficients and reflects the absence of other prior knowledge about this parameter. Note that any abundance could be removed from α^+ and not only the last one α_R . For symmetry reasons, the algorithm proposed in the next section will remove one abundance coefficient from α^+ uniformly drawn in $\{1, \dots, R\}$. Here, this component is supposed to be α_R to simplify notations. Moreover, for sake of conciseness, the notations $\mu(\alpha)$ and $c(\alpha)$ will be used in the sequel to denote the quantities in (5) where α_R has been replaced by $1 - \sum_{r=1}^{R-1} \alpha_r$.

3) *Endmember variance prior*: The prior distribution for the variance σ^2 is chosen as a conjugate inverse Gamma distribution

$$\sigma^2|\delta \sim \mathcal{IG}(\nu, \delta) \quad (10)$$

where ν and δ are the shape and scale parameters [21]. This paper classically assumes $\nu = 1$ (as in [17]) and estimates δ within a hierarchical Bayesian framework. Such approach require to define a prior distribution for the hyperparameter δ . This paper assumes that the prior of δ is a non-informative Jeffreys’ prior [22] defined by

$$f(\delta) \propto \frac{1}{\delta} \mathbf{1}_{\mathbb{R}^+}(\delta). \quad (11)$$

This prior reflects the lack of knowledge regarding the hyperparameter δ .

C. Posterior distribution

The posterior distribution of the unknown parameter vector $\theta = \{\alpha, \sigma^2, \mathbf{M}, R\}$ can be easily computed after marginalizing over the hyperparameter δ

$$f(\theta|\mathbf{y}) \propto \int f(\mathbf{y}|\alpha, \sigma^2, \mathbf{M}, R) f(\alpha|R) P(\mathbf{M}|R) P(R) f(\sigma^2|\delta) f(\delta) d\delta \quad (12)$$

Straightforward computations lead to the following posterior distribution

$$f(\theta|\mathbf{y}) \propto \frac{\Gamma(R+1)\Gamma(R_{\max}-R+1)}{R_{\max}\Gamma(R_{\max}+1)} \times \frac{1}{\sigma^{2R+L} [c(\alpha)]^{L/2}} \exp\left(-\frac{\|\mathbf{y} - \mu(\alpha)\|^2}{2\sigma^2 c(\alpha)}\right) \mathbf{1}_{\mathbb{S}}(\alpha). \quad (13)$$

The posterior distribution (13) is too complex to derive the MMSE or MAP estimators of θ . In particular, the dimension of θ is unknown and depends on the number of endmembers R . In such case, it is usual to resort to MCMC methods to generate samples distributed according to the posterior distribution and to use these samples to approximate the MMSE or MAP estimators. To handle the model selection problem resulting from the unknown value of R , we propose a reversible jump MCMC method that generate samples according to $f(\theta | \mathbf{y})$ defined in (3).

III. REVERSIBLE JUMP MCMC ALGORITHM

A stochastic simulation algorithm will be employed to sample according to $f(\alpha, \sigma^2, \mathbf{M}, R | \mathbf{y})$. The vectors to be sampled belong to a space whose dimension depends on R , requiring to use a dimension matching strategy as in [23], [24]. More precisely, the proposed algorithm, detailed in Algo. 1, consists of four different moves

- 1) updating the means of endmembers contained in the matrix \mathbf{M} ,
- 2) updating the abundance vector α ,
- 3) updating the variance σ^2 ,
- 4) updating the hyperparameter δ .

These steps are scanned systematically as in [23] and are detailed below.

ALGORITHM 1:

Hybrid Metropolis-within-Gibbs sampler for semi-supervised hyperspectral image unmixing according to the NCM

- Initialization:
 - Sample parameter $R^{(0)}$,
 - Choose $R^{(0)}$ spectra in the library \mathcal{S} to build $\mathbf{M}^{(0)}$,
 - Sample parameters $\sigma^{2(0)}$ and $\alpha^{(0)}$,
 - Set $t \leftarrow 1$,
- Iterations: For $t = 1, 2, \dots$, do
 - Update $\mathbf{M}^{(t)}$:
 - draw $u_1 \sim \mathcal{U}_{[0,1]}$,
 - if** $u_1 \leq b_{R^{(t-1)}}$ **then**
 - propose a *birth* move (see Algo. 2),
 - else if** $b_{R^{(t-1)}} < u_1 \leq b_{R^{(t-1)}} + d_{R^{(t-1)}}$ **then**
 - propose a *death* move see Algo. 3),
 - else if** $u_1 > b_{R^{(t-1)}} + d_{R^{(t-1)}}$ **then**

```

    propose a switch move see Algo. 4),
  end if
  draw  $u_2 \sim \mathcal{U}_{[0,1]}$ ,
  if  $u_2 < \rho$  (see Eq. (18)) then
    set  $(\boldsymbol{\alpha}^{(t)}, \mathbf{M}^{(t)}, R^{(t)}) = (\boldsymbol{\alpha}^*, \mathbf{M}^*, R^*)$ ,
  else
    set  $(\boldsymbol{\alpha}^{(t)}, \mathbf{M}^{(t)}, R^{(t)}) = (\boldsymbol{\alpha}^{(t-1)}, \mathbf{M}^{(t-1)}, R^{(t-1)})$ ,
  end if,
  - Sample  $\boldsymbol{\alpha}^{(t)}$  from the pdf in Eq. (20),
  - Sample  $\sigma^{2(t)}$  from the pdf in Eq. (22) ,
  - Sample  $\delta^{(t)}$  from the pdf in Eq. (23) ,
  - Set  $t \leftarrow t + 1$ ,

```

A. Updating the endmember matrix \mathbf{M}

Three types of moves, referred to as “birth”, “death” and “switch” (as in [15]), enable one to update the endmember spectrum means involved in the mixture. The two first moves consist of increasing or decreasing the number of pure components R by 1. Therefore, they require the use of the reversible jump MCMC method introduced by Green [16]. In the third move, the dimension of R is not changed, requiring the use of a standard Metropolis-Hastings (MM) acceptance procedure. Assume that at iteration t , the current model is defined by $(\boldsymbol{\alpha}^{(t)}, \mathbf{M}^{(t)}, R^{(t)})$. The “birth”, “death” and “switch” moves are defined as follows (the step-by-step algorithms are respectively presented in Algo. 2, Algo. 3 and Algo. 4)

- birth: a *birth* move $R^* = R^{(t)} + 1$ is proposed with the probability $b_{R^{(t)}}$. A new spectrum \mathbf{s}^* is randomly chosen among the available endmembers means of the library \mathcal{S} to build $\mathbf{M}^* = [\mathbf{M}^{(t)}, \mathbf{s}^*]$. As the number of pure components has increased by 1, a new abundance coefficient vector $\boldsymbol{\alpha}^{+*}$ is proposed according to a rule inspired by [23]:
 - draw a new abundance coefficient w^* from the Beta distribution $\mathcal{Be}(1, R^{(t)})$,
 - re-scale the existing weights so that all weights sum to 1, using $\alpha_r^* = \alpha_r^{(t)}(1 - w^*)$, $r = 1, \dots, R^{(t)}$,
 - build $\boldsymbol{\alpha}^{+*} = [\alpha_1^*, \dots, \alpha_{R^{(t)}}^*, w^*]^T$,

ALGORITHM 2:

birth move

- 1: set $R^* = R^{(t)} + 1$,
- 2: choose \mathbf{s}^* in \mathcal{S} such as $\mathbf{s}^* \neq \mathbf{m}_r^{(t)}$, $r = 1, \dots, R^{(t)}$,
- 3: set $\mathbf{M}^* = [\mathbf{m}_1^{(t)}, \dots, \mathbf{m}_{R^{(t)}}^{(t)}, \mathbf{s}^*]$,
- 4: draw $w^* \sim \mathcal{Be}(1, R^{(t)})$,
- 5: add w^* to $\boldsymbol{\alpha}^{+(t)}$ and re-scale the others coefficients, i.e.

$$\boldsymbol{\alpha}^{+*} = \left[\frac{\alpha_1^{(t)}}{C}, \dots, \frac{\alpha_{R^{(t)}}^{(t)}}{C}, w^* \right]^T$$

$$\text{with } C = \frac{1}{(1-w^*)}.$$

- death: a *death* move $R^* = R^{(t)} - 1$ is proposed with the probability $d_{R^{(t)}}$. One of the spectra of $\mathbf{M}^{(t)}$ is removed, as well as the corresponding abundance coefficient. The remaining abundances coefficients are re-scaled to sum to 1,

ALGORITHM 3:

death move

- 1: set $R^* = R^{(t)} - 1$,
- 2: draw $j \sim \mathcal{U}_{\{1, \dots, R^{(t)}\}}$,
- 3: remove $\mathbf{m}_j^{(t)}$ from $\mathbf{M}^{(t)}$, i.e. set

$$\mathbf{M}^* = [\mathbf{m}_1^{(t)}, \dots, \mathbf{m}_{(j-1)}^{(t)}, \mathbf{m}_{(j+1)}^{(t)}, \dots, \mathbf{m}_{R^{(t)}}^{(t)}],$$

- 4: remove $\alpha_j^{(t)}$ from $\boldsymbol{\alpha}^{+(t)}$ and re-scale the remaining coefficients, i.e. set

$$\boldsymbol{\alpha}^{+*} = \left[\frac{\alpha_1^{(t)}}{C}, \dots, \frac{\alpha_{(j-1)}^{(t)}}{C}, \frac{\alpha_{(j+1)}^{(t)}}{C}, \dots, \frac{\alpha_{R^{(t)}}^{(t)}}{C} \right]^T,$$

$$\text{with } C = \sum_{r \neq j} \alpha_r^{(t)}.$$

- switch: a *switch* move is proposed with the probability $u_{R^{(t)}}$. A spectrum randomly chosen in $\mathbf{M}^{(t)}$ is replaced by another spectrum randomly chosen in the library \mathcal{S} . This move allows one to improve the mixing properties of the sampler.

ALGORITHM 4:

switch move

- 1: **if** $R_{\max} - R^{(t)} \neq 0$ **then**
- 2: draw $j \sim \mathcal{U}_{\{1, \dots, R^{(t)}\}}$,
- 3: choose \mathbf{s}^* in \mathcal{S} such as $\mathbf{s}^* \neq \mathbf{m}_r^{(t)}$, $r = 1, \dots, R^{(t)}$,
- 4: replace $\mathbf{m}_j^{(t)}$ in $\mathbf{M}^{(t)}$ by \mathbf{s}^* , i.e. set

$$\mathbf{M}^* = [\mathbf{m}_1^{(t)}, \dots, \mathbf{m}_{(j-1)}^{(t)}, \mathbf{s}^*, \mathbf{m}_{(j+1)}^{(t)}, \dots, \mathbf{m}_{R^{(t)}}^{(t)}],$$

4: set $\alpha^{+*} = \alpha^{+(t)}$ and $R^* = R^{(t)}$.

5: **end if**

At each iteration, one of these moves is randomly chosen with probabilities $b_{R^{(t)}}, d_{R^{(t)}}$ and $u_{R^{(t)}}$. These probabilities follow three conditions²

- $b_{R^{(t)}} + d_{R^{(t)}} + u_{R^{(t)}} = 1$,
- $d_1 = 0$ (a *death* move is not allowed for $R = 1$),
- $b_{R_{\max}} = 0$ (a *birth* move is impossible for $R = R_{\max}$),

As a result, $b_{R^{(t)}} = d_{R^{(t)}} = u_{R^{(t)}} = \frac{1}{3}$ for $R \in \{2, R_{\max} - 1\}$ and $b_1 = d_{R_{\max}} = u_1 = u_{R_{\max}} = \frac{1}{2}$.

The abundance coefficient vector α has a uniform prior distribution on the simplex defined in (8), which is equivalent to choosing a Dirichlet distribution $\mathcal{D}_R(1, \dots, 1)$ as prior for the full abundance vector α^+ . Therefore, the acceptance probability for the *birth* move is $\rho = \min\{1, A_b\}$ with:

$$A_b = \left[\frac{c(\alpha^{(t)})}{c(\alpha^*)} \right]^{\frac{L}{2}} \frac{d_{R^{(t)}+1}}{b_{R^{(t)}}} \frac{R^{(t)}}{g_{1,R^{(t)}}(w^*)} (1 - w^*)^{R^{(t)}-1} \times \exp \left[\frac{\|\mathbf{y} - \mu(\alpha^{(t)})\|^2}{2c(\alpha^{(t)})} - \frac{\|\mathbf{y} - \mu(\alpha^*)\|^2}{2c(\alpha^*)} \right] \quad (14)$$

where $g_{a,b}$ denotes the probability density function (pdf) of the Beta distribution $\mathcal{B}e(a, b)$ (see Appendix A).

Regarding the acceptance probability $\rho = \min\{1, A_d\}$ for the *death* move, two cases have to be considered. When the current model order is such that $R^{(t)} \neq 2$, the acceptance probability is

$$A_d = \exp \left[\frac{\|\mathbf{y} - \mu(\alpha^{(t)})\|^2}{2c(\alpha^{(t)})} - \frac{\|\mathbf{y} - \mu(\alpha^*)\|^2}{2c(\alpha^*)} \right] \times \left[\frac{c(\alpha^{(t)})}{c(\alpha^*)} \right]^{\frac{L}{2}} \frac{b_{R^{(t)}-1}}{d_R} \frac{g_{1,R^{(t)}-1}(\alpha_{R^{(t)}}^*)}{R^{(t)} - 1} (1 - \alpha_{R^{(t)}}^*)^{R^{(t)}-2}. \quad (15)$$

When $R^{(t)} = 2$, the death move yields $R^* = 1$, i.e. one single pure element is present in the proposed model, which leads obviously to $\alpha^* = 1$. Thus, α^* is deterministic, and the acceptance rate is

$$A_d = \frac{f(\mathbf{M}^*, R^* | \mathbf{y})}{f(\mathbf{M}^{(t)}, R^{(t)} | \mathbf{y})} \frac{p_{R^* \rightarrow R^{(t)}}}{p_{R^{(t)} \rightarrow R^*}} \frac{q(\mathbf{M}^{(t)} | \mathbf{M}^*)}{q(\mathbf{M}^* | \mathbf{M}^{(t)})}. \quad (16)$$

²The case $R = 1$ has been accepted for the semi-supervised algorithm as a pixel can be spectrally pure.

In this case, the Jacobian equals to 1 as α^* is independent from $\alpha^{(t)}$. Consequently, the previous equation yields

$$A_d = \left[\frac{c(\alpha^{(t)})}{c(\alpha^*)} \right]^{L/2} \frac{b_{R^{(t)}-1}}{d_R^{(t)}} \exp \left[\frac{\|\mathbf{y} - \boldsymbol{\mu}(\alpha^{(t)})\|^2}{2c(\alpha^{(t)})} - \frac{\|\mathbf{y} - \boldsymbol{\mu}(\alpha^*)\|^2}{2c(\alpha^*)} \right]. \quad (17)$$

Concerning the *switch* move, the acceptance probability is the standard MM ratio $\rho = \min\{1, A_s\}$ with:

$$A_s = \left[\frac{c(\alpha^{(t)})}{c(\alpha^*)} \right]^{L/2} \exp \left[\frac{\|\mathbf{y} - \boldsymbol{\mu}(\alpha^{(t)})\|^2}{2c(\alpha^{(t)})} - \frac{\|\mathbf{y} - \boldsymbol{\mu}(\alpha^*)\|^2}{2c(\alpha^*)} \right]. \quad (18)$$

Note that the proposal ratio associated to the *switch* move is 1, since in each direction the probability of selecting one spectrum from the library is $1/(R_{\max} - R^{(t)})$. Once the endmember matrix \mathbf{M} has been obtained, the abundances, the endmember variances and the hyperparameter δ are generated following the procedures detailed below.

B. Generating samples according to $f(\alpha|\mathbf{y}, R, \sigma^2, \mathbf{M})$

The Bayes' theorem yields

$$f(\alpha|\mathbf{y}, R, \sigma^2, \mathbf{M}) \propto f(\mathbf{y}|\boldsymbol{\theta})f(\alpha|R) \quad (19)$$

which easily leads to

$$f(\alpha|\mathbf{y}, R, \sigma^2, \mathbf{M}) \propto \frac{1}{[\sigma^2 c(\alpha)]^{L/2}} \times \exp \left(-\frac{\|\mathbf{y} - \boldsymbol{\mu}(\alpha)\|^2}{2\sigma^2 c(\alpha)} \right) \mathbf{1}_{\mathbb{S}}(\alpha). \quad (20)$$

Thus the conditional distribution of α is defined on the simplex \mathbb{S} ensuring the abundance vector α^+ satisfies the positivity and sum-to-one constraints (2) imposed to the model. The generation of α according to (20) can be achieved using a Metropolis-within-Gibbs move with a uniform prior distribution (9) as proposal distribution.

C. Generating samples according to $f(\sigma^2|\mathbf{y}, R, \alpha, \mathbf{M}, \delta)$

The conditional distribution of the variance σ^2 can be determined as follows

$$f(\sigma^2|\mathbf{y}, R, \alpha, \mathbf{M}, \delta) \propto f(\mathbf{y}|\boldsymbol{\theta})f(\sigma^2|\delta). \quad (21)$$

Consequently, the conditional distribution of the noise variance is the following inverse Gamma distribution

$$\sigma^2|\mathbf{y}, R, \alpha, \mathbf{M}, \delta \sim \mathcal{IG} \left(\frac{L}{2} + 1, \frac{\|\mathbf{y} - \boldsymbol{\mu}(\alpha)\|^2}{2c(\alpha)} + \delta \right). \quad (22)$$

D. Generating samples according to $f(\delta|\sigma^2, R)$

The conditional distribution of δ is the following Gamma distribution

$$\delta|\sigma^2, R \sim \mathcal{G}\left(R, \frac{R}{\sigma^2}\right) \quad (23)$$

where $\mathcal{G}(a, b)$ is the Gamma distribution with shape parameter a and scale parameter b [21, p. 581].

IV. SIMULATION RESULTS

A. Synthetic pixel

This section studies the accuracy of the semi-supervised algorithm presented in Sections II and III for unmixing a synthetic pixel. We first consider a pixel resulting from the combination of $R = 3$ random endmembers with variance $\sigma^2 = 0.002$ and abundance vector $\alpha^+ = [0.5, 0.3, 0.2]^T$. The endmember means are the spectra of construction concrete, green grass and micaceous loam. The results have been obtained for $N_{MC} = 20000$ iterations, including $N_{bi} = 1500$ burn-in iterations, with the algorithm summarized in Algo. 1. The spectrum library \mathcal{S} used in this simulation, depicted in Fig. 2, contains six elements defined as construction concrete, green grass, micaceous loam, olive green paint, bare red brick and galvanized steel metal. These spectra have been extracted from the spectral libraries provided with the ENVI software [25]. The first step of the analysis estimates the number of components R . The posterior distribution of R depicted in Fig. 3 is clearly in good agreement with the actual value of R since its maximum is obtained for $R = 3$. The second step of the analysis estimates the *a posteriori* probabilities of all spectrum combinations, conditioned to $R = 3$. For this simulation example, 100% of these sampled spectrum combinations are composed of the first three spectra of the library which are the actual spectra defining the mixture. The posterior distributions of the corresponding abundance coefficients and variance σ^2 are finally estimated and depicted in Fig. 4 and Fig. 5. The posteriors are in good agreement with the actual values of these parameters, i.e., $\alpha^+ = [0.5 \ 0.3 \ 0.2]^T$ and $\sigma^2 = 0.002$.

B. Comparison with other model selection strategies

This section compares the Bayesian algorithm developed in this paper with five other model selection strategies that have been applied to mixing models. First, it is important to note that

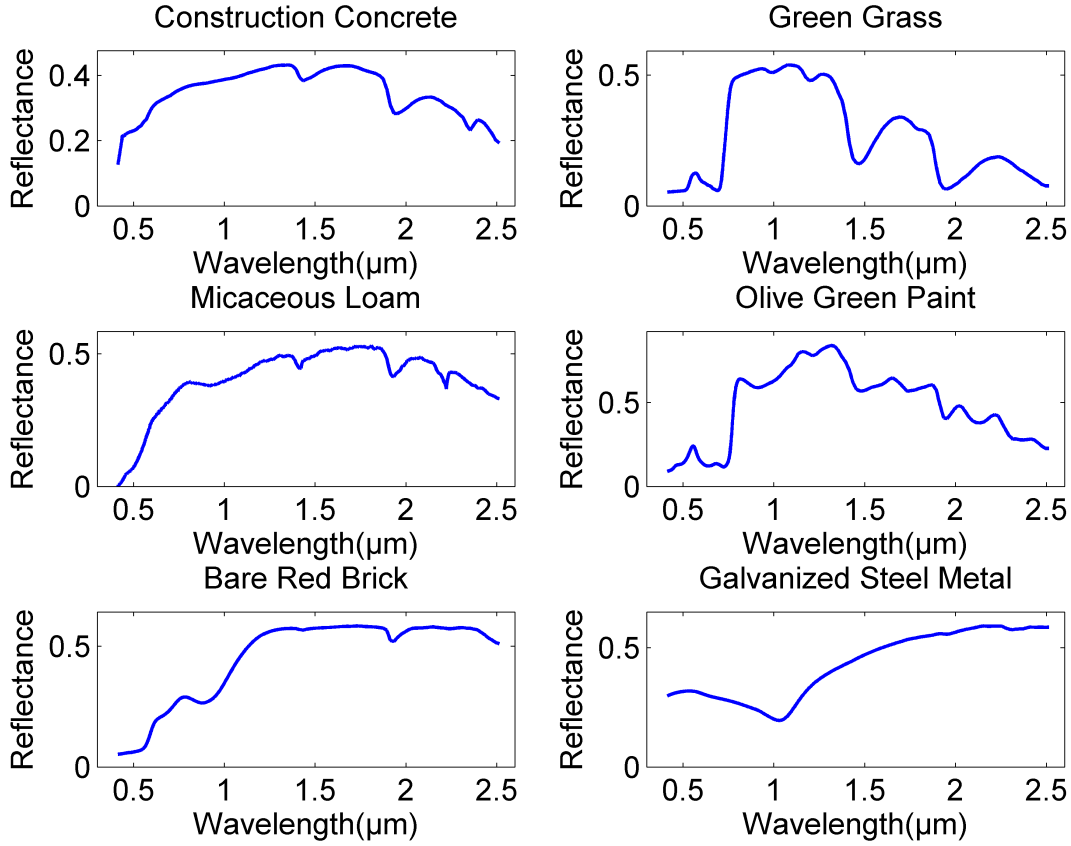


Fig. 2. Endmember spectra of the library.

all model selection methods can be used only when several pixels share the same endmembers, contrary to the proposed algorithm. Moreover, these methods cannot be used to determine which endmembers from the library are actually present in the observed mixture (they just allow one to estimate the number of endmembers R defining (3)). We suppose here that several synthetic pixels are sharing the same R endmember spectra. The first algorithm introduced in [31] studies a maximum likelihood estimator (MLE) of the intrinsic dimension of the observed data. The five others methods mainly rely on the sample covariance matrix of the observed pixels. More precisely, we consider the enhanced versions of the well-known Akaike information criterion (AIC) and minimum distance length (MDL) initially derived in [32], [33], [34] and re-considered in [27] for detecting signals embedded in white noise. The new AIC-based criterion constructed

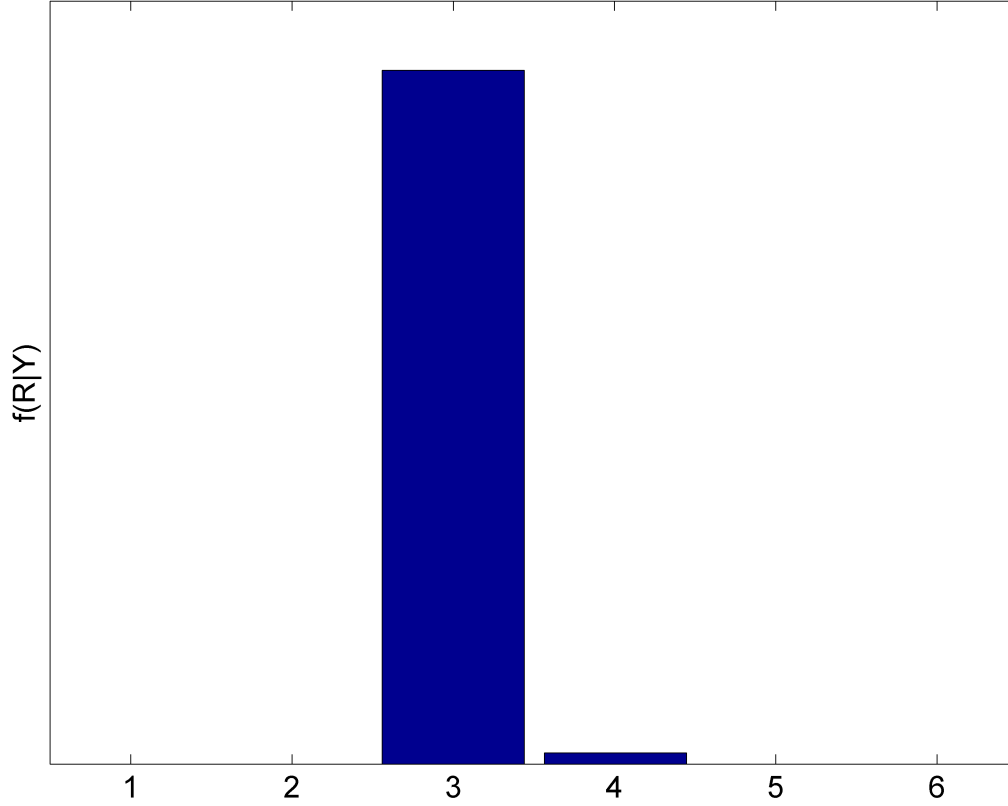


Fig. 3. Posterior distribution of the estimated number of component R .

from random matrix theory (RMT) in [28] (referred to as RMT estimator) and the Malinowski's F-test advocated in [29] are also investigated for comparison. The performance of the PCA-based dimension reduction technique described in [31] and widely used in the hyperspectral imagery community [4] has also been evaluated. Finally, the hyperspectral signal subspace identification by minimum error (HySime) drawn in [30] is considered for this comparison.

To provide a dimension estimator common to all data-sets, several pixels sharing the same R endmember spectra have to be considered. For the algorithms described above, the number of observations (i.e., the number pixels) has to be larger than the observation dimension (i.e., the number of bands). The simulation parameters have been fixed to $L = 207$ bands and $P = 225$ pixels. The $P = 225$ pixels have been simulated according to the NCM in (3) with two different

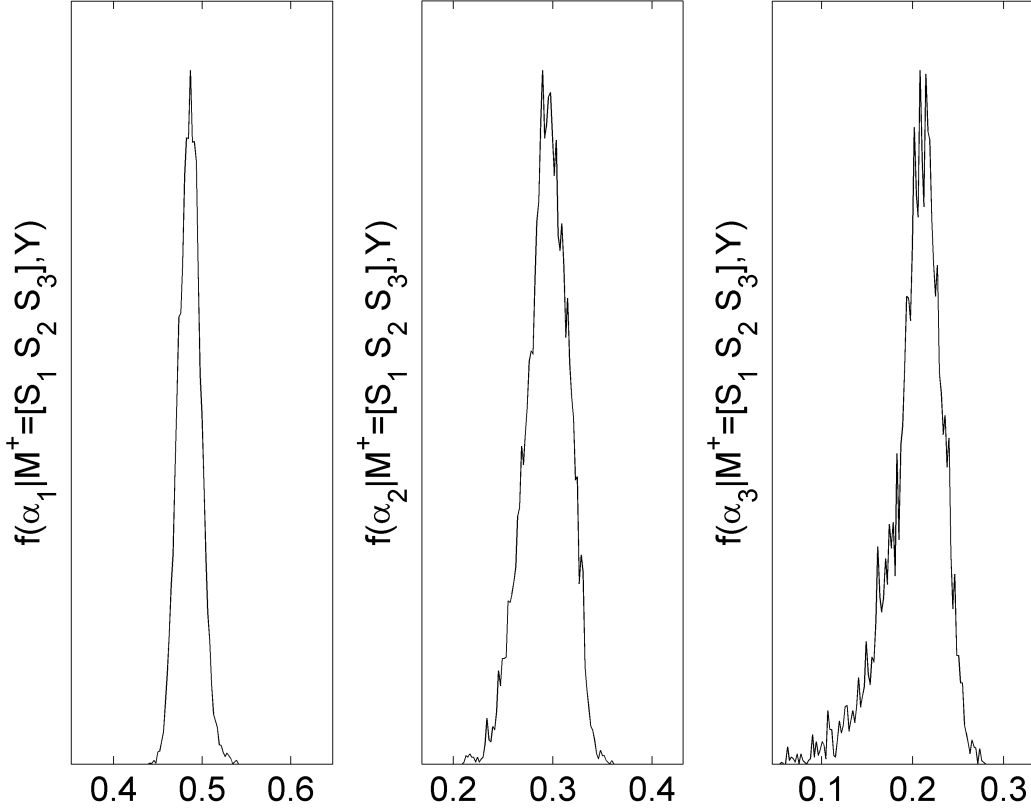


Fig. 4. Posterior distribution of the estimated abundances.

variances ($\sigma^2 = 2.10^{-5}$ and $\sigma^2 = 0.01$) and different numbers of endmember signatures ($R = 3$, $R = 4$ and $R = 5$). For each scenario, the abundance vectors α^+ for each pixel have been generated according to Gaussian distributions truncated to the simplex (8) with the following mean vectors

- 3 endmembers : $\alpha_{R=3} = [0.4, 0.25]$,
- 4 endmembers : $\alpha_{R=4} = [0.3, 0.15, 0.2]$,
- 5 endmembers: $\alpha_{R=5} = [0.3, 0.15, 0.1, 0.1]$.

The threshold default values recommended in [31] have been used for the PCA-based method and the Malinowski's F-test. The estimated number of components R obtained by the proposed reversible-jump MCMC method is compared to the other model selection algorithms in Table I

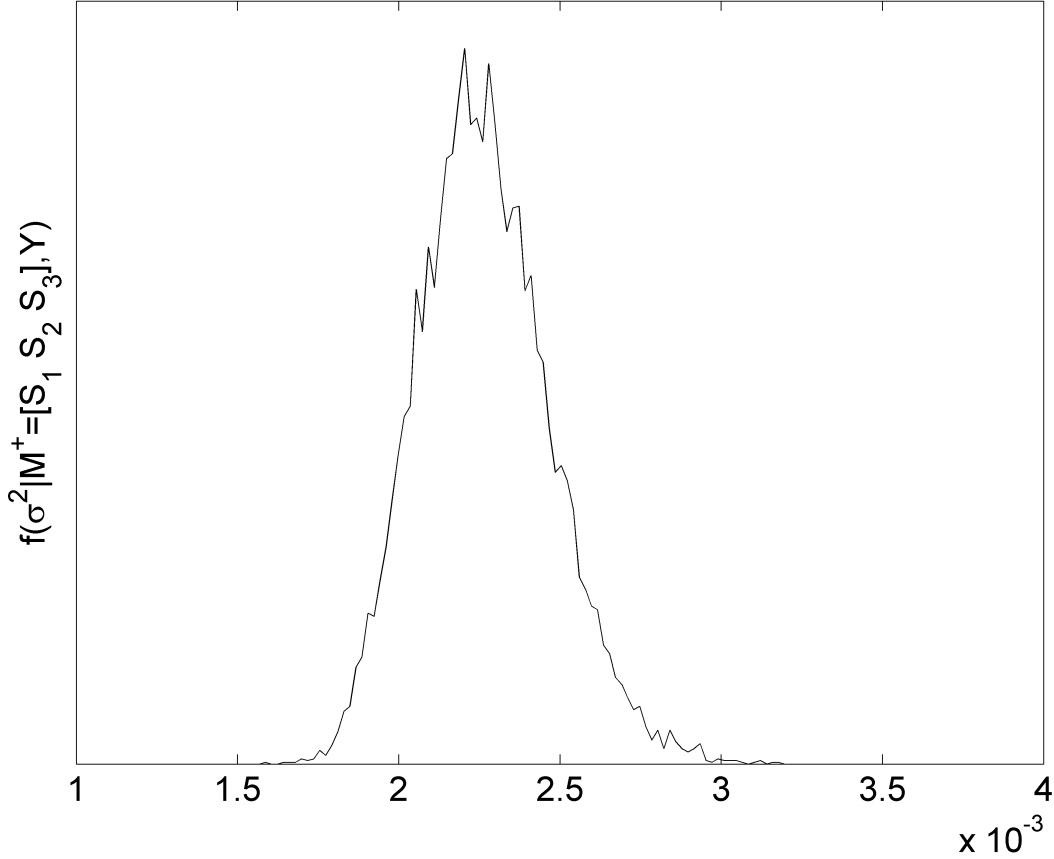


Fig. 5. Posterior distribution of the estimated variance σ^2 .

for the different simulation settings. The reversible-jump MCMC method performs significantly better than the other model selection algorithms especially for larger noise variance (i.e., $\sigma^2 = 0.01$). This result can be explained by the fact the proposed Bayesian algorithm fully takes advantage of the NCM model structure, contrary to the other model selection techniques.

V. SPECTRAL UNMIXING OF REAL AVIRIS DATA

A. Moffet field

This section considers a real hyperspectral image of size 50×50 depicted in Fig. 6 to evaluate the performance of the different algorithms. This image has been extracted from a larger image acquired in 1997 by the Airborne Visible Infrared Imaging Spectrometer (AVIRIS) over Moffett Field, CA, and has received much attention in hyperspectral imagery literature (see for instance

TABLE I

ESTIMATED NUMBERS OF PURE COMPONENTS (IN %) OBTAINED BY THE REVERSIBLE-JUMP MCMC METHOD AND OTHER MODEL SELECTION ALGORITHMS WITH $\sigma^2 = 1.10^{-2}$ AND $\sigma^2 = 2.10^{-5}$.

	$\sigma^2 = 1.10^{-2}$									$\sigma^2 = 2.10^{-5}$								
	$R = 3$			$R = 4$			$R = 5$			$R = 3$			$R = 4$			$R = 5$		
Estimated R	2	3	4	3	4	5	4	5	6	2	3	4	3	4	5	4	5	6
RJ-MCMC	0	100	0	0	100	0	0	100	0	0	100	0	0	100	0	0	100	0
MLE [26]	0	0	0	0	0	0	0	0	0	0	0	100	0	100	0	100	0	0
AIC [27]	0	100	0	0	95	5	45	55	0	0	100	0	0	99	1	0	95	5
MDL [27]	5	0	0	0	0	0	0	0	0	100	0	0	100	0	0	100	0	0
RMT [28]	32	38	20	1	8	21	0	18	20	30	34	27	0	6	17	0	12	23
F-test [29]	100	0	0	7	0	0	0	0	0	100	0	0	100	0	0	100	0	0
PCA [4]	0	0	0	0	0	0	0	0	0	100	0	0	100	0	0	83	0	0
HySime [30]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

[35] and [36]). The analyzed area is composed of a lake (top) and a coastal zone (bottom). The data set has been classically reduced from the original 224 bands to $L = 189$ bands by removing water absorption bands.

As detailed in Section III, the reversible jump MCMC algorithm requires the use of a spectral library containing the endmember spectra that might appear in the image to be analyzed. As no ground truth is available for this AVIRIS image, we have built the required library from the image. Firstly, a K-means procedure has been used to classify the image into 6 regions. Then the purest pixels related to each class, extracted with the pixel purity index algorithm [37], have been chosen as components of the library. This new spectral library whose spectra are represented in Fig.7 contains six endmembers. Three of them have been clearly identified

- endmember 1: water,
- endmember 3: vegetation,
- endmember 6: soil,

whereas the endmembers 2, 4 and 5 are more difficult to be interpreted.



Fig. 6. Real hyperspectral data: Moffett field acquired by AVIRIS in 1997 (left) and the region of interest shown in true colors (right).

The Moffett image is analyzed by the proposed reversible jump MCMC algorithm, using this spectral library. For each pixel, the abundance vector is estimated conditionally upon the endmember matrix. The fraction maps for each endmember and the presence maps (with a threshold of 15% for the abundances) obtained by the algorithm are depicted in Fig.'s 9 and 10, respectively. Fig. 8 shows the maps of the estimated number of endmembers in the considered area.

The lake area includes at least two endmembers defined as “water” and the unidentified “endmember #2”. As the water spectrum energy is very low, the algorithm has to balance the measured water pixel spectrum by adding other contributions, which explains the presence of more than one endmember in this area. The lake-shore area contains the greatest number of endmembers (up to 6), including the unidentified endmembers 4 and 5. The ground area is composed exclusively of soil and vegetation, which is clearly in good agreement with the results obtained for instance in [15] and [14].

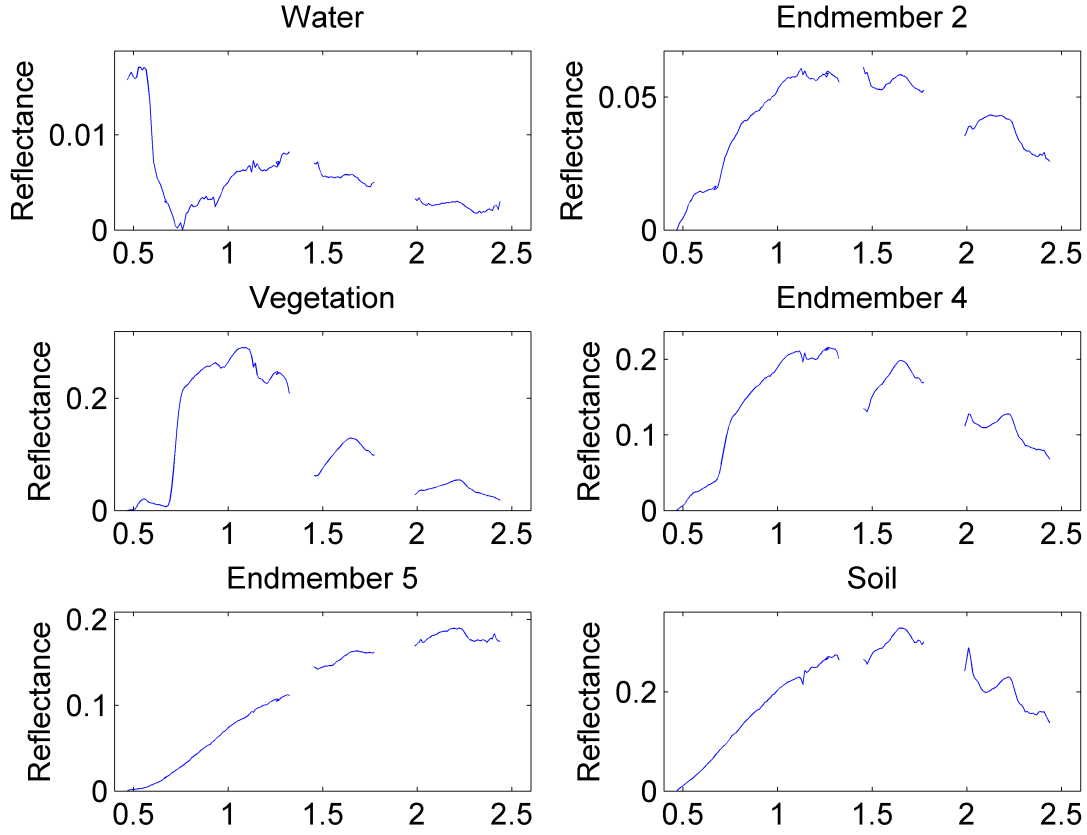


Fig. 7. The $R = 6$ endmember spectra extracted from ENVI.

B. Cuprite

The “alunite hill” that appears in the Cuprite scene, acquired by AVIRIS over the Cuprite mining site, Nevada, in 1997, has also been considered to evaluate the performance of the proposed semi-supervised algorithm. The geologic characteristics of this area of interest, represented in Fig. 11, have been mapped in [38], [39]. As a first step of the analysis, the VCA algorithm has been applied on this image to extract a spectral library depicted in Fig. 12.

The reversible-jump MCMC algorithm has been applied on a sub-image of size 28×16 pixels extracted from the Cuprite scene. The estimated numbers of endmembers and fraction maps are represented in Figs. 13 and 14.

The alunite hill has been clearly recovered: up to five endmembers are present in this center

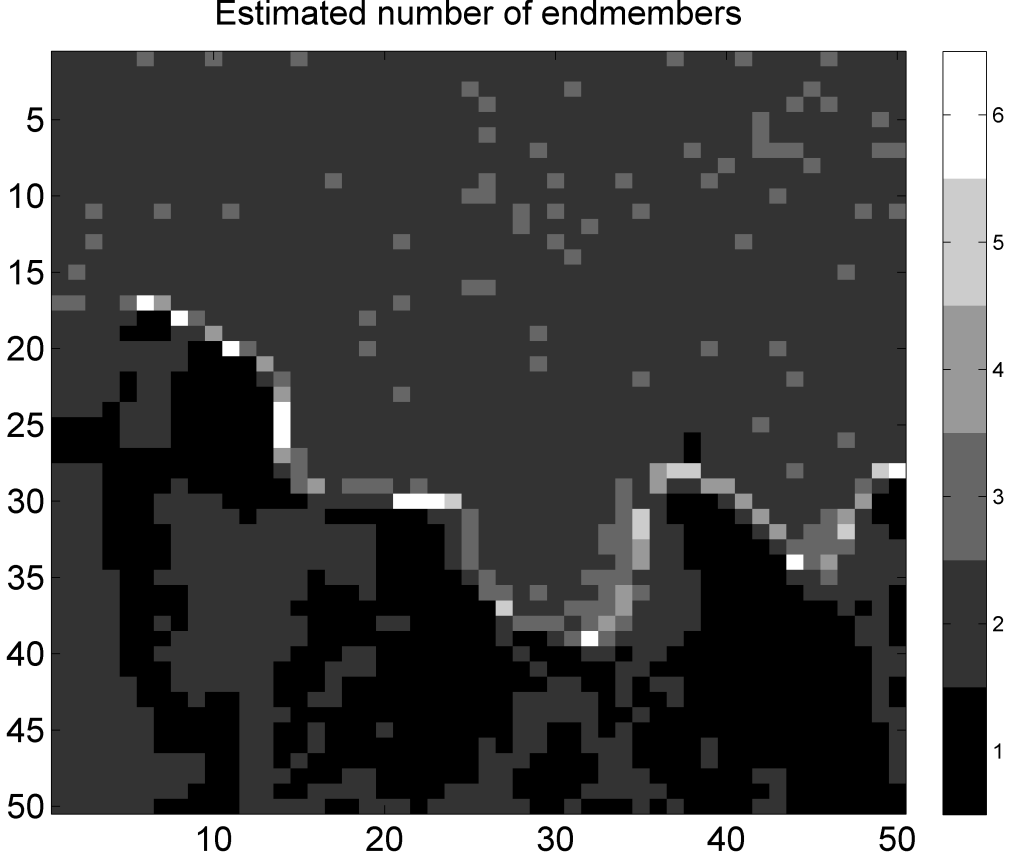


Fig. 8. Number of endmembers estimated by the proposed algorithm (darker (resp. brighter) areas means $R = 1$ (resp. $R = 6$)).

part of the image, dominated by the endmembers 3 and 4. The area surrounding the alunite hill is mainly composed of the endmember 1. By comparing these spectra with the spectral signatures extracted from the United States Geological Survey (USGS) Spectral Library [39], the endmembers 1, 3 and 4 can be identified as muscovite, alunite and kaolinite, respectively. To illustrate this point, the estimated spectra and the corresponding signatures from the USGS library are depicted in Fig. 15. These results, in good agreement with the results reported in [40], confirm the good performance of our reversible-jump algorithm.

VI. CONCLUSIONS

A new semi-supervised hierarchical Bayesian unmixing algorithm was derived for hyper-spectral images. This algorithm was based on the normal compositional model introduced by

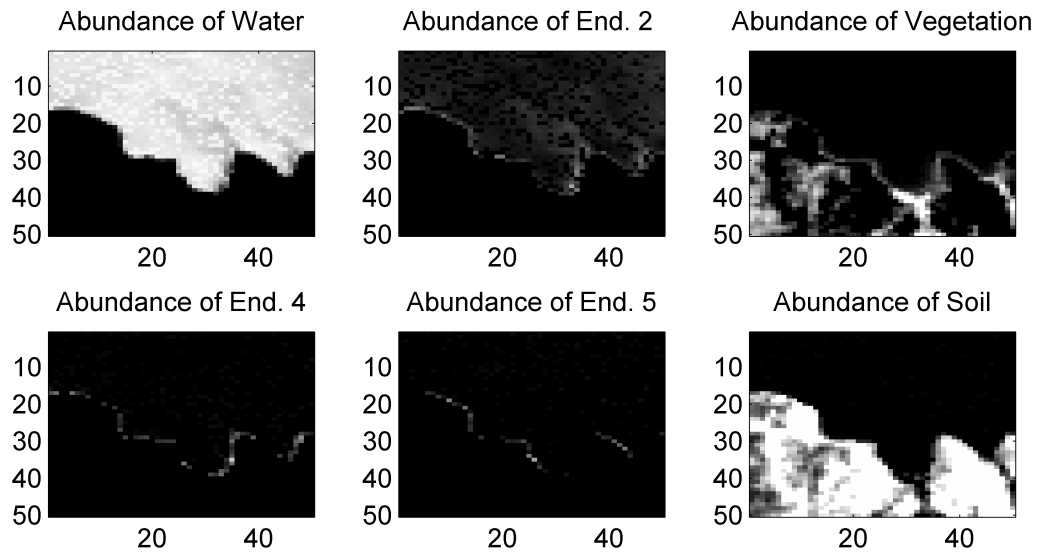


Fig. 9. Fraction maps estimated by the proposed algorithm.

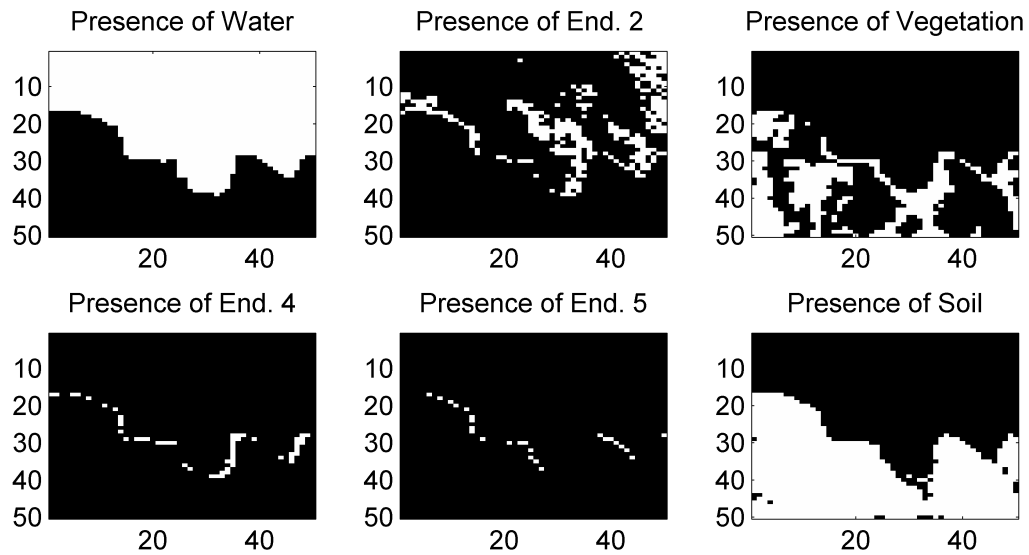


Fig. 10. Presence maps estimated by the proposed algorithm.

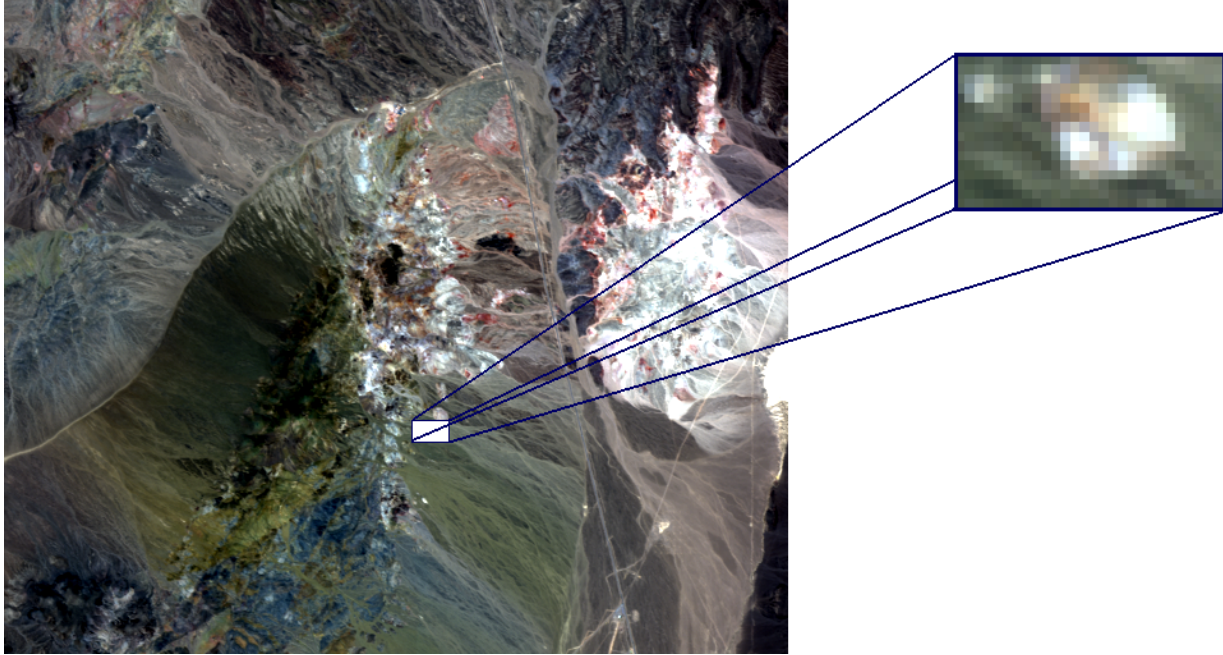


Fig. 11. Real hyperspectral data: Cuprite image acquired by AVIRIS in 1997 (left) and the region of interest shown in true colors (right).

Eismann and Stein [7]. The proposed algorithm generated samples distributed according to the joint posterior of the abundances, the endmember variances and one hyperparameter. These samples were then used to estimate the parameters of interest. The proposed algorithm showed several advantages versus the standard model selection strategies. In particular, it allows one to estimate which components from a spectral library participate in the mixture for a single given pixel. The simulation results on synthetic and real data showed interesting results.

Perspectives include the extension of the NCM algorithm to more complex scene models. For instance, the hyperspectral images could be considered as a set of homogenous regions, in which common endmembers are present. In this case, spatial correlations for the abundance vectors could be introduced between neighboring pixels to improve unmixing.

APPENDIX A

ACCEPTANCE PROBABILITIES FOR THE “BIRTH” AND “DEATH” MOVES

The acceptance probabilities for the “birth” and “death” moves introduced in Section III are derived in this appendix.

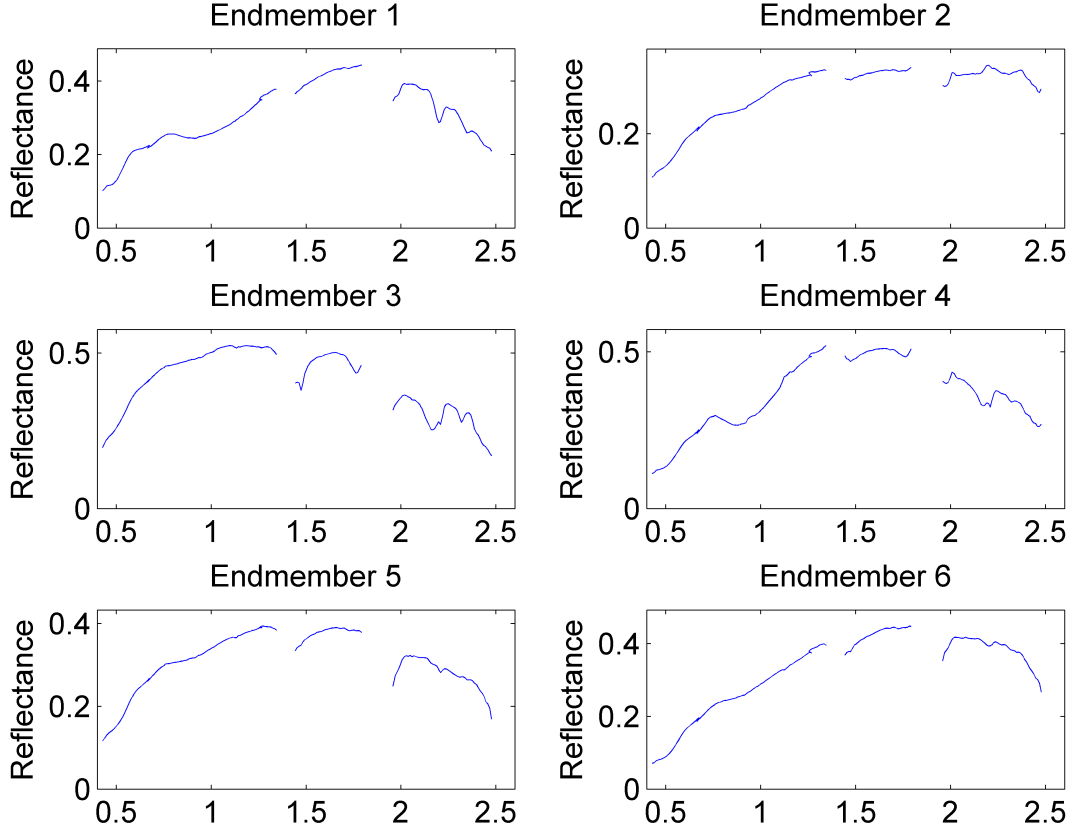


Fig. 12. The $R = 6$ endmember spectra estimated by VCA in the Cuprite image.

At iteration index t , consider the birth move from the state $\{\alpha^{(t)}, \mathbf{M}^{(t)}, R^{(t)}\}$ to the new state $\{\alpha^*, \mathbf{M}^*, R^*\}$ with $\alpha^* = [(1 - w^*)\alpha_1, \dots, (1 - w^*)\alpha_{R^{(t)}}]^T$, $\mathbf{M}^* = [\mathbf{M}^{(t)}, \mathbf{s}^*]$ and $R^* = R^{(t)} + 1$. The acceptance ratio associated to this move is

$$A_b = \frac{f(\alpha^*, \mathbf{M}^*, R^* | \mathbf{y})}{f(\alpha^{(t)}, \mathbf{M}^{(t)}, R^{(t)} | \mathbf{y})} \frac{p_{R^* \rightarrow R^{(t)}}}{p_{R^{(t)} \rightarrow R^*}} \frac{q(\mathbf{M}^{(t)}, \alpha^{(t)} | \mathbf{M}^*, \alpha^*)}{q(\mathbf{M}^*, \alpha^* | \mathbf{M}^{(t)}, \alpha^{(t)})} |\mathbf{J}(w^*)| \quad (24)$$

where $q(\cdot | \cdot)$ refers to the proposal distribution, $|\mathbf{J}(w^*)|$ is the Jacobian of the transformation and $p_{\cdot \rightarrow \cdot}$ denotes the transition probability, i.e. $p_{R^* \rightarrow R^{(t)}} = d_{R^*}$ et $p_{R^{(t)} \rightarrow R^*} = b_{R^{(t)}}$. According to the moves specified in Section III, the proposal ratio is

$$\frac{q(\mathbf{M}^{(t)}, \alpha^{(t)} | \mathbf{M}^*, \alpha^*)}{q(\mathbf{M}^*, \alpha^* | \mathbf{M}^{(t)}, \alpha^{(t)})} = \frac{1}{g_{1, R^{(t)}}(w^*)} \frac{R_{\max} - R^{(t)}}{R^{(t)} + 1} \quad (25)$$

where $g_{a,b}(\cdot)$ denotes the pdf of a Beta distribution $\mathcal{B}e(a, b)$. Indeed, the probability of choosing

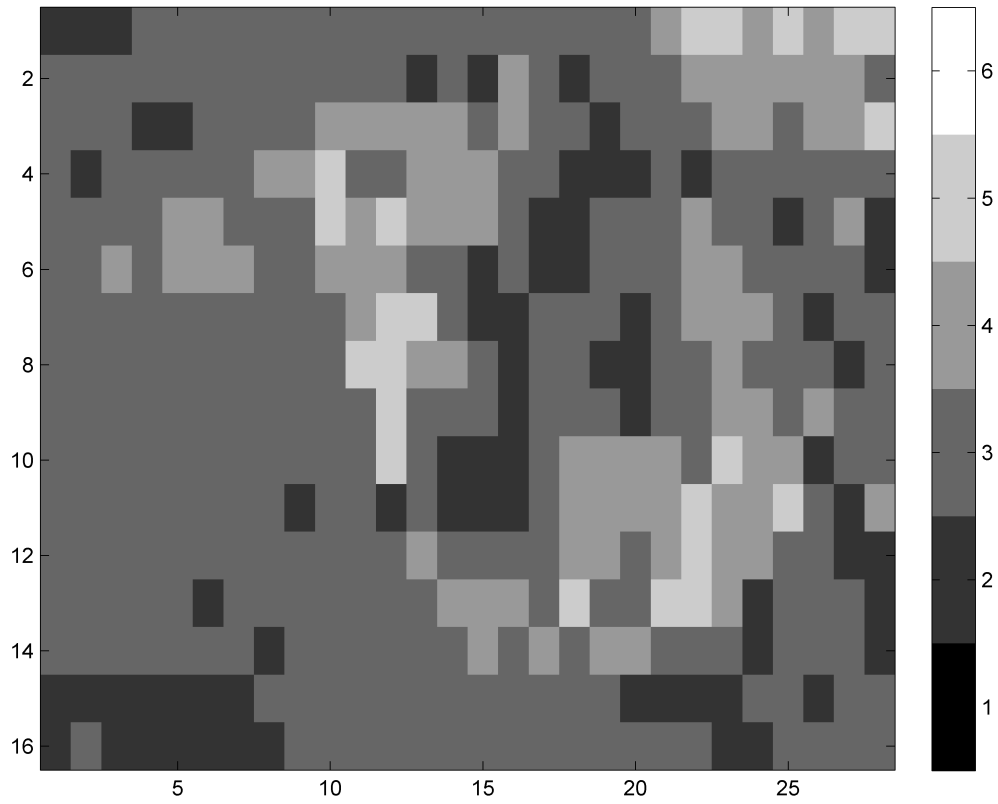


Fig. 13. Number of endmembers estimated by the proposed algorithm (darker (resp. brighter) areas means $R = 1$ (resp. $R = 6$)).

a new element in the library (“birth” move) is $1/(R_{\max} - R^{(t)})$ and the probability of removing an element (“death” move) is $1/(R^{(t)} + 1)$. The Jacobian matrix of the transformation is given

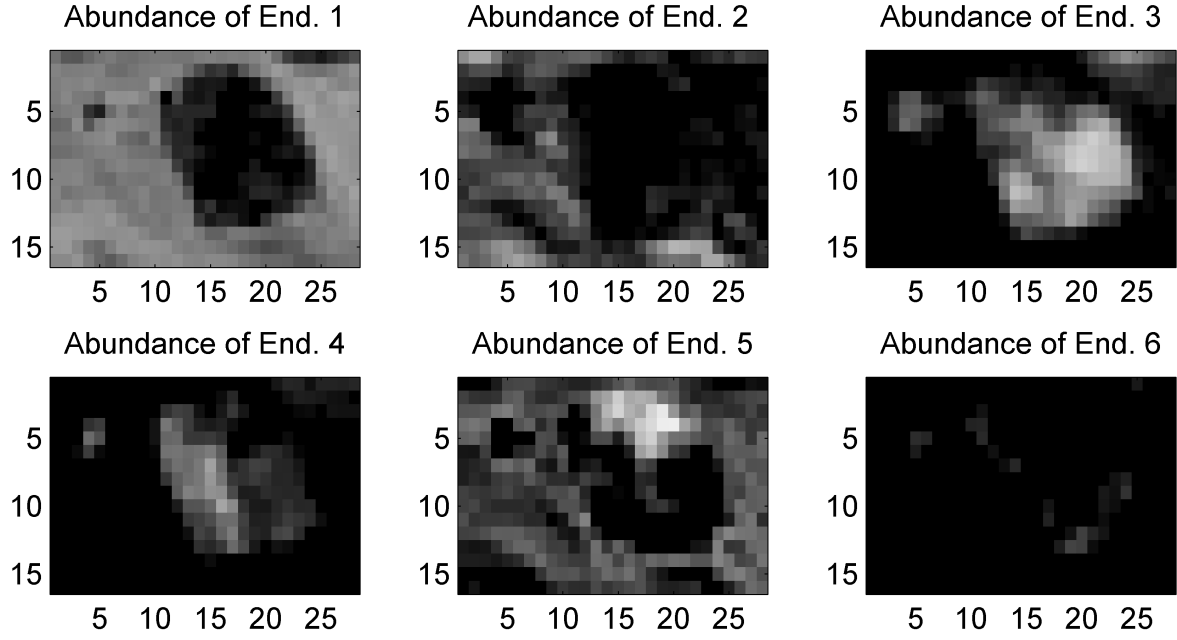


Fig. 14. Fraction maps of the Cuprite area estimated by the proposed algorithm.

as follows

$$\begin{aligned}
 \mathbf{J}(w^*) &= \begin{pmatrix} \frac{\partial \alpha_1^*}{\partial \alpha_1^{(t)}} & \cdots & \frac{\partial \alpha_1^*}{\partial \alpha_{R-1}^{(t)}} & \frac{\partial \alpha_1^*}{\partial w^*} \\ \vdots & \ddots & \vdots & \vdots \\ \frac{\partial \alpha_R^*}{\partial \alpha_1^{(t)}} & \cdots & \frac{\partial \alpha_R^*}{\partial \alpha_{R-1}^{(t)}} & \frac{\partial \alpha_R^*}{\partial w^*} \end{pmatrix} \\
 &= \begin{pmatrix} 1 - w^* & 0 & \cdots & 0 & -\alpha_1^{(t)} \\ 0 & 1 - w^* & \cdots & 0 & -\alpha_2^{(t)} \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 - w^* & -\alpha_{R-1}^{(t)} \\ -(1 - w^*) & -(1 - w^*) & \cdots & -(1 - w^*) & -\alpha_R^{(t)} \end{pmatrix}
 \end{aligned}$$

which implies

$$|\mathbf{J}(w^*)| = (1 - w^*)^{R^{(t)} - 1}. \quad (26)$$

The posterior ratio is

$$\frac{f(\boldsymbol{\alpha}^*, \mathbf{M}^*, R^* | \mathbf{y})}{f(\boldsymbol{\alpha}^{(t)}, \mathbf{M}^{(t)}, R^{(t)} | \mathbf{y})} = \frac{f(\mathbf{y} | \boldsymbol{\alpha}^*, \mathbf{M}^*, R^*)}{f(\mathbf{y} | \boldsymbol{\alpha}^{(t)}, \mathbf{M}^{(t)}, R^{(t)})} \frac{f(\boldsymbol{\alpha}^* | R^*)}{f(\boldsymbol{\alpha}^{(t)} | R^{(t)})} \frac{f(\mathbf{M}^* | R^*)}{f(\mathbf{M}^{(t)} | R^{(t)})} \frac{f(R^*)}{f(R^{(t)})}. \quad (27)$$

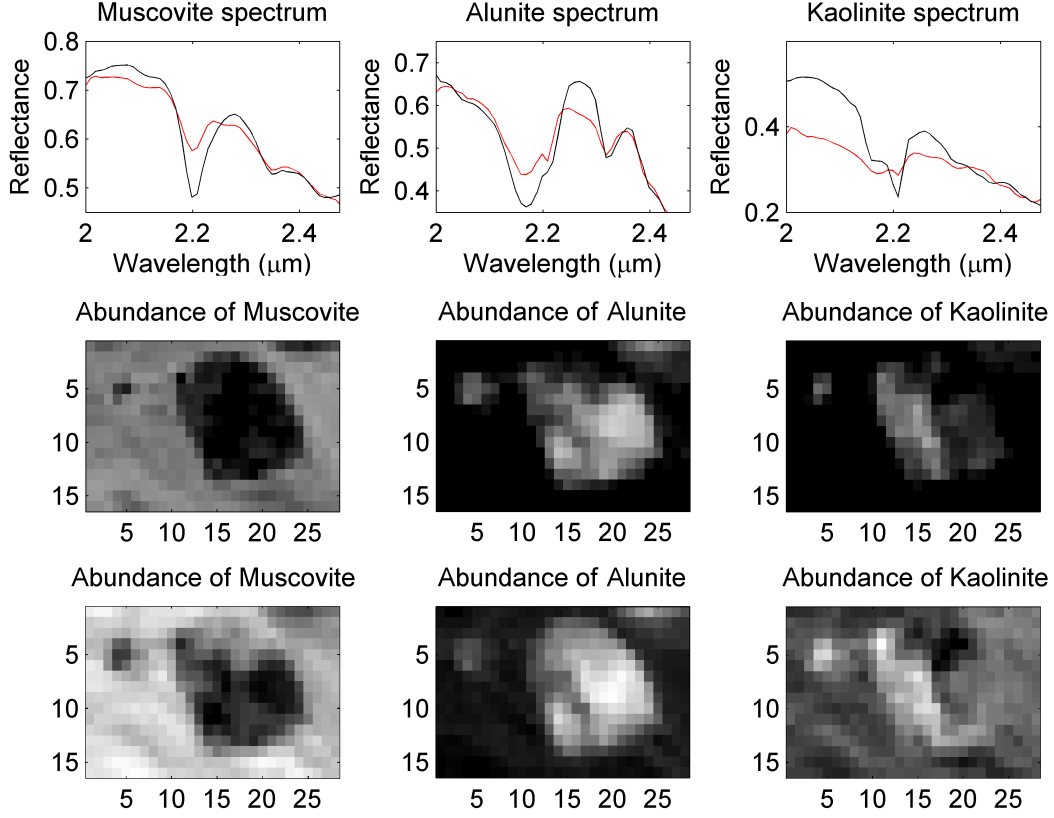


Fig. 15. Top: the 3 endmember spectra from our library recovered by VCA algorithm in the Cuprite scene (red lines), compared with the signatures extracted from the USGS spectral library (black lines). Middle: the corresponding abundance maps estimated by the proposed Bayesian reversible-jump algorithm. Bottom: the corresponding abundance maps estimated by the unsupervised Bayesian algorithm in [40].

The prior distribution of the abundance coefficient vector α^+ is a Dirichlet distribution $\mathcal{D}_R(1, \dots, 1)$, which leads to

$$\frac{f(\alpha^*|R^*)}{f(\alpha^{(t)}|R^{(t)})} = \frac{\Gamma(R^{(t)} + 1)}{\Gamma(1)\Gamma(R^{(t)})} = R^{(t)} + 1. \quad (28)$$

By choosing *a priori* equiprobable configurations for \mathbf{M} conditionally upon R , the prior ratio related to the means spectrum matrix is

$$\frac{f(\mathbf{M}^*|R^*)}{f(\mathbf{M}^{(t)}|R^{(t)})} = \frac{R^{(t)} + 1}{R_{\max} - R^{(t)}}. \quad (29)$$

The prior ratio related to the number of mixtures R reduces to 1. Finally, the acceptance ratio

for the “birth” move can be written

$$A_b = \exp \left[\frac{\|\mathbf{y} - \boldsymbol{\mu}(\boldsymbol{\alpha}^{(t)})\|^2}{2c(\boldsymbol{\alpha}^{(t)})} - \frac{\|\mathbf{y} - \boldsymbol{\mu}(\boldsymbol{\alpha}^*)\|^2}{2c(\boldsymbol{\alpha}^*)} \right] \times \left(\frac{c(\boldsymbol{\alpha}^{(t)})}{c(\boldsymbol{\alpha}^*)} \right)^{\frac{L}{2}} \frac{d_{R^{(t)}+1}}{b_{R^{(t)}}} \frac{R^{(t)}}{g_{1,R^{(t)}}(w^*)} (1 - w^*)^{R^{(t)}-1}. \quad (30)$$

REFERENCES

- [1] C.-I Chang, *Hyperspectral data exploitation: theory and applications*. Hoboken, NJ: Wiley, 2007.
- [2] R. O. Green *et al.*, “Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (AVIRIS),” *Remote Sens. Environ.*, vol. 65, no. 3, pp. 227–248, Sept. 1998.
- [3] J. S. Pearlman, P. S. Barry, C. C. Segal, J. Shepanski, D. Beiso, and S. L. Carman, “Hyperion, a space-based imaging spectrometer,” *IEEE Trans. Geosci. and Remote Sensing*, vol. 41, no. 6, pp. 1160–1173, June 2003.
- [4] N. Keshava and J. Mustard, “Spectral unmixing,” *IEEE Signal Processing Magazine*, pp. 44–56, Jan. 2002.
- [5] M. E. Winter, “Fast autonomous spectral endmember determination in hyperspectral data,” in *Proc. 13th Int. Conf. on Applied Geologic Remote Sensing*, vol. 2, Vancouver, April 1999, pp. 337–344.
- [6] J. M. Nascimento and J. M. Bioucas-Dias, “Vertex component analysis: a fast algorithm to unmix hyperspectral data,” *IEEE Trans. Geosci. and Remote Sensing*, vol. 43, no. 4, pp. 898–910, April 2005.
- [7] M. T. Eismann and D. Stein, “Stochastic mixture modeling,” in *Hyperspectral Data Exploitation: Theory and Applications*, C.-I. Chang, Ed. Wiley, 2007, ch. 5.
- [8] N. Dobigeon and J.-Y. Tourneret, “Bayesian sampling of structured noise covariance matrix for hyperspectral imagery,” University of Toulouse, France, Tech. Rep., Dec. 2008. [Online]. Available: <http://dobigeon.perso.enseeiht.fr/>
- [9] N. Dobigeon, J.-Y. Tourneret, and Alfred O. Hero III, “Bayesian linear unmixing of hyperspectral images corrupted by colored Gaussian noise with unknown covariance matrix,” in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Proc. (ICASSP)*, Las Vegas, Nevada, USA, April 2008, pp. 3433–3436.
- [10] J. Settle, “On the relationship between spectral unmixing and subspace projection,” *IEEE Trans. Geosci. and Remote Sensing*, vol. 34, no. 4, pp. 1045–1046, July 1996.
- [11] C.-I Chang, “Further results on relationship between spectral unmixing and subspace projection,” *IEEE Trans. Geosci. and Remote Sensing*, vol. 36, no. 3, pp. 1030–1032, May 1998.
- [12] C.-I Chang, X.-L. Zhao, M. L. G. Althouse, and J. J. Pan, “Least squares subspace projection approach to mixed pixel classification for hyperspectral images,” *IEEE Trans. Geosci. and Remote Sensing*, vol. 36, no. 3, pp. 898–912, May 1998.
- [13] D. Manolakis, C. Siracusa, and G. Shaw, “Hyperspectral subpixel target detection using the linear mixing model,” *IEEE Trans. Geosci. and Remote Sensing*, vol. 39, no. 7, pp. 1392–1409, July 2001.
- [14] O. Eches, N. Dobigeon, C. Mailhes, and J.-Y. Tourneret, “Unmixing hyperspectral images using a normal compositional model and MCMC methods,” University of Toulouse, Tech. Rep., March 2009. [Online]. Available: <http://eches.perso.enseeiht.fr>
- [15] N. Dobigeon, J.-Y. Tourneret, and C.-I Chang, “Semi-supervised linear spectral using a hierarchical Bayesian model for hyperspectral imagery,” *IEEE Trans. Signal Processing*, vol. 56, no. 7, pp. 2684–2696, July 2008.
- [16] P. J. Green, “Reversible jump Markov Chain Monte Carlo methods computation and Bayesian model determination,” *Biometrika*, vol. 82, no. 4, pp. 711–732, Dec. 1995.

- [17] 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.
- [18] N. Dobigeon, J.-Y. Tourneret, and M. Davy, "Joint segmentation of piecewise constant autoregressive processes by using a hierarchical model and a Bayesian sampling approach," *IEEE Trans. Signal Processing*, vol. 55, no. 4, pp. 1251–1263, April 2007.
- [19] M. Davy, S. Godsill, and J. Idier, "Bayesian analysis of polyphonic western tonal music," *J. Acoust. Soc. Am.*, vol. 119, no. 4, pp. 2498–2517, April 2006.
- [20] C. Andrieu and A. Doucet, "Joint Bayesian model selection and estimation of noisy sinusoids via reversible jump MCMC," *IEEE Trans. Signal Processing*, vol. 47, no. 10, pp. 2667–2676, Oct. 1999.
- [21] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*, 2nd ed. New York: Springer Verlag, 2004.
- [22] H. Jeffreys, "An invariant form for the prior probability in estimation problems," *Proc. of the Royal Society of London. Series A*, vol. 186, no. 1007, pp. 453–461, 1946.
- [23] S. Richardson and P. J. Green, "On Bayesian analysis of mixtures with an unknown number of components," *J. Roy. Stat. Soc. Ser. B*, vol. 59, no. 4, pp. 731–792, 1997.
- [24] —, "Corrigendum : On Bayesian analysis of mixtures with an unknown number of components," *J. Roy. Stat. Soc. Ser. B*, vol. 60, no. 3, p. 661, 1998.
- [25] RSI (Research Systems Inc.), *ENVI User's guide Version 4.0*, Boulder, CO 80301 USA, Sept. 2003.
- [26] E. Levina and P. J. Bickel, "Maximum likelihood estimation of intrinsic dimension," in *Advances in Neural Information Processing Systems 17*, L. K. Saul, Y. Weiss, and L. Bottou, Eds. Cambridge, MA: MIT Press, 2005, pp. 777–784.
- [27] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 33, no. 2, pp. 387–392, Feb. 1985.
- [28] R. R. Nadakuditi and A. Edelman, "Sample eigenvalue based detection of high-dimensional signals in white noise using relatively few samples," *IEEE Trans. Signal Processing*, vol. 56, no. 7, pp. 2625–2638, July 2008.
- [29] E. R. Malinowski, "Statistical F-tests for abstract factor analysis and target testing," *Journal of Chemometrics*, vol. 3, no. 1, pp. 49–60, 1988.
- [30] J. M. Bioucas-Dias and J. M. P. Nascimento, "Hyperspectral subspace identification," *IEEE Trans. Geosci. and Remote Sensing*, vol. 46, no. 8, pp. 2435–2445, 2008.
- [31] E. Levina, A. S. Wagaman, A. F. Callender, G. S. Mandair, and M. D. Morris, "Estimating the number of pure chemical components in a mixture by maximum likelihood," *Journal of Chemometrics*, vol. 21, no. 1–2, pp. 24–34, 2007.
- [32] H. Akaike, "Information theory and an extension of the maximum likelihood principle," in *Proc. 2nd Int. Symp. Inf. Theory*, 1973, pp. 267–281.
- [33] —, "A new look at the statistical model identification," *IEEE Trans. Autom. Contr.*, vol. 19, no. 6, pp. 716–723, Dec. 1974.
- [34] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, no. 5, pp. 465–471, 1978.
- [35] E. Christophe, D. Léger, and C. Mailhes, "Quality criteria benchmark for hyperspectral imagery," *IEEE Trans. Geosci. and Remote Sensing*, vol. 43, no. 9, pp. 2103–2114, Sept. 2005.
- [36] T. Akgun, Y. Altunbasak, and R. M. Mersereau, "Super-resolution reconstruction of hyperspectral images," *IEEE Trans. Image Processing*, vol. 14, no. 11, pp. 1860–1875, Nov. 2005.
- [37] J. Boardman, "Automating spectral unmixing of AVIRIS data using convex geometry concepts," in *Summaries 4th Annu. JPL Airborne Geoscience Workshop*, vol. 1. Washington, D.C.: JPL Pub., 1993, pp. 11–14.

- [38] R. N. Clark, G. A. Swayze, and A. Gallagher, "Mapping minerals with imaging spectroscopy, U.S. Geological Survey," *Office of Mineral Resources Bulletin*, vol. 2039, pp. 141–150, 1993.
- [39] R. N. Clark *et al.*, "Imaging spectroscopy: Earth and planetary remote sensing with the USGS Tetracorder and expert systems," *J. Geophys. Res.*, vol. 108, no. E12, pp. 5–1–5–44, Dec. 2003.
- [40] N. Dobigeon, S. Moussaoui, M. Coulon, J.-Y. Tournet, and A. O. Hero, "Joint bayesian endmember extraction and linear unmixing for hyperspectral imagery," *IEEE Trans. Signal Processing*, vol. 57, no. 11, pp. 4355–4368, Nov. 2009.