

Unmixing Multitemporal Hyperspectral Images Accounting for Smooth and Abrupt Variations

Pierre-Antoine Thouvenin, Nicolas Dobigeon and Jean-Yves Tourneret
University of Toulouse, IRIT/INP-ENSEEIH, 2 rue Camichel, 31071 Toulouse cedex 7, France
Email: `firstname.lastname@enseeiht.fr`

Abstract—A classical problem in hyperspectral imaging, referred to as hyperspectral unmixing, consists in estimating spectra associated with each material present in an image and their proportions in each pixel. In practice, illumination variations (e.g., due to declivity or complex interactions with the observed materials) and the possible presence of outliers can result in significant changes in both the shape and the amplitude of the measurements, thus modifying the extracted signatures. In this context, sequences of hyperspectral images are expected to be simultaneously affected by such phenomena when acquired on the same area at different time instants. Thus, we propose a hierarchical Bayesian model to simultaneously account for smooth and abrupt spectral variations affecting a set of multitemporal hyperspectral images to be jointly unmixed. This model assumes that smooth variations can be interpreted as the result of endmember variability, whereas abrupt variations are due to significant changes in the imaged scene (e.g., presence of outliers, additional endmembers, etc.). The parameters of this Bayesian model are estimated using samples generated by a Gibbs sampler according to its posterior. Performance assessment is conducted on synthetic data in comparison with state-of-the-art unmixing methods.

I. INTRODUCTION

Spectral unmixing consists in extracting a set of signatures associated with different materials observed in a hyperspectral (HS) image – referred to as endmembers – and their abundance fractions within each pixel of the image. Depending on the context, different mixture models allow specific environmental factors to be accounted for (e.g. multiple reflections, declivity). A linear mixing model (LMM) is usually considered when microscopic interactions between the observed materials are negligible and the relief of the scene is flat [1]. In practice, varying acquisition conditions (such as the natural evolution of the scene or illumination variations) may significantly affect the acquired signatures [2], leading to endmember variability. Even if endmember spatial variability in a given HS image has been extensively considered in the literature (using statistical [3]–[5] and deterministic [6]–[12] approaches) fewer works have considered temporal variability by exploiting multitemporal HS (MTHS) images. However, MTHS images, which are composed of images successively acquired over the same area, allow information redundancies between consecutive images to be exploited (e.g. *via* features exhibiting smooth temporal variations) [13]–[15]. While these approaches allow smooth, moderate temporal variations of the mixture parameters to be represented, they do not address abrupt spectral changes affecting HS images, either resulting from the presence of outliers or from significant time intervals between consecutive acquisitions. This commonly encountered phenomenon (e.g., in real datasets when

This work was supported in part by the Hypanema ANR Project no. ANR-12-BS03-003, by the MapInvPlnt ERA-NET MED Project no. ANR-15-NMED-0002-02, by the Thematic Trimester on Image Processing of the CIMI Labex under Grant ANR-11-LABX-0040-CIMI within the Program ANR-11-IDEX-0002-02 and by the Direction Générale de l’Armement, French Ministry of Defense.

water, snow or vegetation is present within the sensor field of view) has a major influence on the extracted endmembers, thus motivating the following study. Inspired by [14], [16], [17], we propose to simultaneously account for smooth endmember variations – in terms of endmember variability – and abrupt spectral changes – in terms of outlier contributions *via* a carefully designed hierarchical Bayesian model. Starting from a reference image analyzed individually, the proposed approach is intended to handle scenes in which most of the materials are likely to be repeatedly observed over time. However, materials not previously observed in the reference scene may later appear in a few pixels of the remaining images. On the one hand, the endmembers associated with repeatedly observed materials in each HS image are interpreted as realizations of the signatures shared by the different images, which allows information redundancies to be exploited across time. On the other hand, endmembers appearing in only a few images are considered as outliers with respect to the commonly shared materials. Based on these observations, a new instance of the perturbed linear mixing model (PLMM) [15] is investigated to account for temporal variability and additive outliers. The resulting problem is formulated within a Bayesian framework and solved with an MCMC method.

The paper is organized as follows. The mixing model, its posterior distribution and the MCMC method used to sample this posterior are introduced in Sections II and III. Section IV presents some simulation results obtained on synthetic data. The performance of the proposed method is appreciated in comparison with the VCA/FCLS algorithm [18], [19], the SISAL/FCLS algorithm [20], the robust LMM (rLMM) described in [21] and the online unmixing (OU) method of [15]. Conclusions and research perspectives are finally reported in Section V.

II. PROBLEM STATEMENT

Consider a sequence of T HS images acquired over the same area, having R endmembers in common (where R is known *a priori*). In this work, we assume that an endmember estimated using a single image can be reasonably interpreted as a smooth deviation of a reference endmember described by a PLMM [10], thus allowing information redundancies to be exploited across time. Reminiscent of [22], the PLMM represents the measurement vector $\mathbf{y}_{n,t}$ in the n th pixel of the t th image by a linear combination of the R endmembers corrupted by an additive term representing temporal variability. This model will however show severe limitations when the measurements are subject to abrupt spectral variations, e.g., due to the presence of outliers. This paper consequently introduces a new unmixing model addressing both endmember variability and abrupt changes encountered in HS images. Inspired by [17], [21], this model

considers an additive outlier term $\mathbf{x}_{n,t}$ capturing deviations from the LMM, leading to the following observation model

$$\mathbf{y}_{n,t} = \sum_{r=1}^R a_{r,n,t} (\mathbf{m}_r + \mathbf{d}\mathbf{m}_{r,t}) + \mathbf{x}_{n,t} + \mathbf{b}_{n,t} \quad (1)$$

for $n \in \{1, \dots, N\}$, $t \in \{1, \dots, T\}$, where N is the number of pixels contained in each image, T denotes the number of HS images and \mathbf{m}_r corresponds to the r th endmember. Considering the t th image of the HS sequence, $a_{r,n,t}$ is the proportion of the r th endmember in the n th pixel of the image, $\mathbf{d}\mathbf{m}_{r,t}$ is the corresponding variability with respect to the r th endmember and $\mathbf{x}_{n,t}$ represents the outlier contribution in the pixel n . Moreover, $\mathbf{b}_{n,t}$ is an additive noise accounting for the modeling and acquisition errors. The model (1) can be expressed in matrix form as

$$\mathbf{Y}_t = (\mathbf{M} + \mathbf{d}\mathbf{M}_t)\mathbf{A}_t + \mathbf{X}_t + \mathbf{B}_t \quad (2)$$

where $\mathbf{Y}_t \in \mathbb{R}^{L \times N}$ contains the pixels composing the image t , $\mathbf{M} \in \mathbb{R}^{L \times R}$ contains the endmembers, $\mathbf{A}_t \in \mathbb{R}^{R \times N}$ contains the abundance vectors $\mathbf{a}_{n,t}$, $\mathbf{d}\mathbf{M}_t \in \mathbb{R}^{L \times R}$, $\mathbf{X}_t \in \mathbb{R}^{L \times N}$ and $\mathbf{B}_t \in \mathbb{R}^{L \times N}$ contain the variability, outliers and noise terms for the t th image. The constraints related to the endmember and abundances are

$$\mathbf{A}_t \succeq \mathbf{0}_{R,N}, \mathbf{M} \succeq \mathbf{0}_{L,R}, \mathbf{M} + \mathbf{d}\mathbf{M}_t \succeq \mathbf{0}_{L,R}, \mathbf{X}_t \succeq \mathbf{0}_{L,N} \quad (3)$$

where \succeq is to be understood as a component-wise inequality. In the rest of this paper, the abundance sum-to-one constraint is relaxed for pixels corrupted by outliers, and enforced for the others (see Section III). Note that the outlier term \mathbf{X}_t may result from the local emergence of an endmember which was absent from the reference image. Thus, it is subject to a non-negativity constraint as the other endmembers, and its support is expected to be spatially sparse. The next section introduces a Bayesian model allowing the parameters of (1) (i.e., the matrices \mathbf{A}_t , \mathbf{M} , $\mathbf{d}\mathbf{M}_t$ and \mathbf{X}_t) to be estimated from the observations \mathbf{Y}_t , $t = 1, \dots, T$ while reflecting the previous observations and the constraints (3).

III. PROPOSED BAYESIAN MODEL

A. Bayesian model

Assuming the additive noise $\mathbf{b}_{n,t}$ has a zero-mean multivariate Gaussian distribution $\mathbf{b}_{n,t} \sim \mathcal{N}(\mathbf{0}_L, \sigma_t^2 \mathbf{I}_L)$ (where $\mathbf{I}_L \in \mathbb{R}^{L \times L}$ is the identity matrix and σ_t^2 is the variance of the noise in the t th image) and assuming independence between the images and between the pixels within each image, the likelihood of the observations in (2) is

$$p(\underline{\mathbf{Y}} | \Theta) \propto \prod_{t=1}^T (\sigma_t^2)^{-NL/2} \times \exp\left(-\frac{1}{2\sigma_t^2} \|\underline{\mathbf{Y}}_t - (\mathbf{M} + \mathbf{d}\mathbf{M}_t)\mathbf{A}_t - \mathbf{X}_t\|_F^2\right) \quad (4)$$

where $\Theta = \{\mathbf{M}, \mathbf{d}\mathbf{M}, \mathbf{A}, \mathbf{X}, \mathbf{Z}, \sigma^2, \Psi^2, \mathbf{s}^2\}$ summarizes the parameters to be inferred, and the underline notation stands for the overall set of the corresponding parameters. Priors specifically adapted to the model parameters and hyperparameters are introduced hereafter.

Abundance prior: We propose to promote smooth temporal variations of the abundances except when the corresponding pixel contains outliers. We thus consider a new prior defined as follows

$$\mathbf{a}_{n,1} | \mathbf{x}_{n,t} = \mathbf{0}_L \sim \mathcal{U}_{S_R}, \quad \mathbf{a}_{n,t} | \mathbf{x}_{n,t} \neq \mathbf{0}_L \sim \mathcal{U}_{\widetilde{S}_R} \quad (5)$$

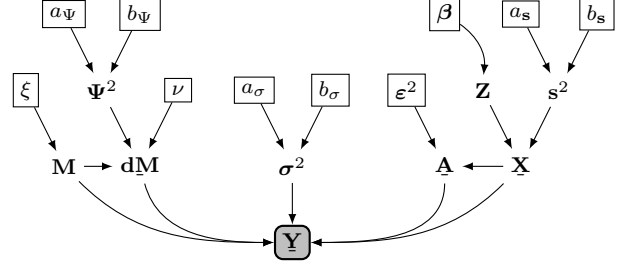


Fig. 1. Description of the proposed Bayesian model using a directed acyclic graph (fixed parameters appear in boxes).

$$p(\mathbf{a}_{n,t} | \mathbf{x}_{n,t} = \mathbf{0}_L, \mathbf{A}_{\setminus\{\mathbf{a}_{n,t}\}}) \propto \exp\left\{-\frac{1}{2\epsilon_n^2} \left(\iota_{[\mathcal{T}_{n,t} \neq \emptyset]} \|a_{n,t} - a_{n,\tau_{n,t}}\|_2^2\right)\right\} \mathbb{1}_{S_R}(\mathbf{a}_{n,t}) \quad (6)$$

where \mathcal{U}_S is a uniform distribution on the set S , $\iota_{[\mathcal{P}]}$ = 1 if the logical proposition \mathcal{P} is true and 0 otherwise, S_R denotes the unit simplex of \mathbb{R}^R , $\widetilde{S}_R = \{\mathbf{x} \in \mathbb{R}^R | \forall i, x_i \geq 0 \text{ and } \mathbf{x}^T \mathbf{1}_R \leq 1\}$ and

$$\mathcal{T}_{n,t} = \{\tau < t | z_{n,\tau} = 0\}, \quad \tau_{n,t} = \max_{\tau \in \mathcal{T}_{n,t}} \tau. \quad (7)$$

We also impose $\mathcal{T}_{n,t} = \emptyset$ if $t = 1$. More explicitly, consider an image at time t and a pixel n within this image such that $\mathbf{x}_{n,t} = \mathbf{0}$ (absence of outliers). When $t = 1$, the absence of specific prior knowledge is ensured by a uniform distribution on the unit simplex to account for the constraints (3). When $t > 1$, the proposed Markovian prior promotes smooth variations of $\mathbf{a}_{n,t}$ by penalizing the Euclidean distance between $\mathbf{a}_{n,t}$ and the corresponding abundance vector of the last outlier-free pixel, i.e., at time instant $\tau_{n,t}$. Note that the usual abundance sum-to-one constraint is relaxed ($\mathbf{a}_{n,t}^T \mathbf{1}_R \leq 1$) when outliers are present in the pixel (n, t) ($\mathbf{x}_{n,t} \neq \mathbf{0}_L$), since the prior is aimed at representing spectra which are not fully described by a linear model.

Outlier and label priors: We assume *a priori* that outliers are spatially (but not necessarily spectrally) sparse [17]. Different approaches have been proposed in the literature to promote sparsity, either relying on the use of the ℓ_1 penalty (e.g., the LASSO [23]) or on mixture probability distributions composed of a Dirac mass at zero and a continuous probability distribution [24] (e.g., the Bernoulli-Laplace [25] and Bernoulli-Gaussian distributions [26], [27], extensively used in the literature [28], [29]). To promote spatial sparsity of the outliers $\mathbf{x}_{n,t}$, we propose to consider the following prior

$$p(\mathbf{x}_{n,t} | z_{n,t}, s_t^2) = (1 - z_{n,t})\delta(\mathbf{x}_{n,t}) + z_{n,t} \mathcal{N}_{\mathbb{R}_+^L}(\mathbf{0}_L, s_t^2). \quad (8)$$

The binary variable $z_{n,t}$ reflects the absence of outliers (i.e., $\mathbf{x}_{n,t} = \mathbf{0}_L$) when $z_{n,t} = 0$, and describes outliers *a priori* via a truncated Gaussian distribution when $z_{n,t} = 1$. Indeed, the outlier term $\mathbf{x}_{n,t}$ results from the emergence of an endmember which was absent from the reference image, thus justifying the use of a truncated Gaussian distribution. In addition, we propose to take advantage of any possible spatial correlations between these outliers by modeling the binary label maps $\mathbf{z}_t \in \mathbb{R}^N$ as Ising-Markov random fields as in [30]. This correlation is natural since the apparition of a new endmember (e.g., snow or ice in winter) will be generally observed in several neighboring pixels.

To complete our model, we have to assign priors to the remaining parameters \mathbf{M} , $\mathbf{d}\mathbf{M}$, σ^2 , and hyperparameters Ψ^2 , \mathbf{s}^2 . Our choices are mainly driven by previous studies related to hyperspectral unmixing, and are briefly summarized below.

Endmember prior: The endmembers can be *a priori* considered to live in a subspace of dimension $K \ll L$ ($K = 2(R - 1)$ in the experiments presented in Section IV) [31] which can be estimated in practice by a principal component analysis (PCA) or a robust PCA (rPCA) [32]. We consider the decomposition used in [31], from which straightforward computations lead to

$$\mathbf{m}_r = (\mathbf{I}_L - \mathbf{U}\mathbf{U}^T)\bar{\mathbf{y}} + \mathbf{U}\mathbf{e}_r, \quad \mathbf{U}^T\mathbf{U} = \mathbf{I}_K \quad (9)$$

where \mathbf{U} represents a basis of the subspace and $\bar{\mathbf{y}}$ corresponds to the sample mean of \mathbf{Y} . The projected endmembers \mathbf{e}_r are then assigned a truncated multivariate Gaussian prior

$$\mathbf{e}_r \sim \mathcal{N}_{\mathcal{E}_r}(\mathbf{0}_K, \xi\mathbf{I}_K), \text{ for } r = 1, \dots, R \quad (10)$$

where \mathcal{E}_r reflects the non-negativity of the endmembers \mathbf{m}_r [31], and $\xi = 1$ ensures a non-informative prior [5], [17]. This prior is in agreement with the non-negativity constraint for the endmembers, and its conjugacy simplifies the analysis.

Variability prior: We consider a prior promoting smooth temporal variations [13], [14] while respecting the constraints (3). It is expressed as

$$dm_{\ell,r,1} | m_{\ell,r} \sim \mathcal{N}_{\mathcal{I}_{\ell,r}}(0, \nu), \quad \mathcal{I}_{\ell,r} = [-m_{\ell,r}, +\infty) \quad (11)$$

$$dm_{\ell,r,t} | m_{\ell,r}, dm_{\ell,r,(t-1)}, \psi_{\ell,r}^2 \sim \mathcal{N}_{\mathcal{I}_{\ell,r}}(dm_{\ell,r,(t-1)}, \psi_{\ell,r}^2) \quad (12)$$

Note that ν penalizes the variability energy in the first image, and that the parameters $\psi_{\ell,r}^2$ of the random walk (12) control the temporal variations of the variability from an image to another.

Hyperparameter priors: Finally, we consider the following conjugate inverse-gamma priors for the variances associated with the noise (σ^2), the variability (Ψ^2) and the outlier (s^2) respectively, i.e., for $\ell = 1, \dots, L$, $r = 1, \dots, R$ and $t = 1, \dots, T$

$$\sigma_t^2 \sim \mathcal{IG}(a_\sigma, b_\sigma), \quad \psi_{\ell,r}^2 \sim \mathcal{IG}(a_\Psi, b_\Psi), \quad s_t^2 \sim \mathcal{IG}(a_s, b_s) \quad (13)$$

where \mathcal{IG} denotes the inverse-gamma distribution, and $a_\sigma = b_\sigma = a_\Psi = b_\Psi = a_s = b_s = 10^{-3}$ ensures non-informative priors.

B. Gibbs sampler

By the Bayes' theorem, the posterior distribution of the unknown parameters Θ is given by $p(\Theta | \mathbf{Y}) \propto p(\mathbf{Y} | \Theta)p(\Theta)$. However, the complexity of the model summarized in Fig. 1 does not allow the maximum *a posteriori* (MAP) or the minimum mean square error (MMSE) estimator to be easily computed from this posterior. For instance, the presence of both continuous and discrete variables leads to a complex inference problem, for which deterministic optimization techniques are poorly adapted. Our strategy consists in sampling $p(\Theta | \mathbf{Y})$ using a Gibbs sampler to approximate the MMSE and marginal MAP (mMAP) estimators as follows

$$\widehat{\mathbf{M}}^{\text{MMSE}} \simeq \frac{1}{\tilde{N}} \sum_{q=N_{\text{bi}}+1}^{N_{\text{MC}}} \mathbf{M}^{(q)}, \quad \widehat{\mathbf{A}}^{\text{MMSE}} \simeq \frac{1}{\tilde{N}} \sum_{q=N_{\text{bi}}+1}^{N_{\text{MC}}} \mathbf{A}_t^{(q)}$$

$$\widehat{\mathbf{dM}}_t^{\text{MMSE}} \simeq \frac{1}{\tilde{N}} \sum_{q=N_{\text{bi}}+1}^{N_{\text{MC}}} \mathbf{dM}_t^{(q)}, \quad \widehat{\mathbf{X}}_t^{\text{MMSE}} \simeq \frac{1}{\tilde{N}} \sum_{q=N_{\text{bi}}+1}^{N_{\text{MC}}} \mathbf{X}_t^{(q)}$$

$$\hat{z}_{n,t}^{\text{mMAP}} \simeq \begin{cases} 0, & \text{if } \#\{q > N_{\text{bi}} : z_{n,t}^{(q)} = 0\} \leq \frac{\tilde{N}}{2} \\ 1, & \text{otherwise} \end{cases}, \quad \tilde{N} = N_{\text{MC}} - N_{\text{bi}}$$

where N_{MC} denotes the number of Monte-Carlo iterations and N_{bi} is the number of burn-in iterations. Further details on the conditional distributions of all the parameters can be found in the extended paper [33]. The proposed Gibbs sampler is summarized in Algo. 1.

Algorithm 1: Proposed Gibbs sampler.

Input: $N_{\text{bi}}, N_{\text{MC}}, \Theta^{(0)}, \beta, \xi, a_\Psi, b_\Psi, a_s, b_s, a_\sigma, b_\sigma, \nu, \varepsilon^2$.

for $q = 1$ **to** N_{MC} **do**
 for $(n, t) = (1, 1)$ **to** (N, T) **do**
 Draw $\mathbf{a}_{n,t}^{(q)} \sim p(\mathbf{a}_{n,t} | \mathbf{y}_{n,t}, \Theta_{\setminus\{\mathbf{a}_{n,t}\}})$;
 for $r = 1$ **to** R **do**
 Draw $\mathbf{e}_r^{(q)} \sim p(\mathbf{e}_r | \mathbf{Y}, \Theta_{\setminus\{\mathbf{e}_r\}})$;
 for $t = 1$ **to** T **do**
 Draw $\mathbf{dM}_t^{(q)} \sim p(\mathbf{dM}_t | \mathbf{Y}_t, \Theta_{\setminus\{\mathbf{dM}_t\}})$;
 for $(n, t) = (1, 1)$ **to** (N, T) **do**
 Draw $z_{n,t}^{(q)} \sim \mathbb{P}[z_{n,t} | \mathbf{y}_{n,t}, \Theta_{\setminus\{z_{n,t}\}}]$;
 Draw $\mathbf{x}_{n,t}^{(q)} \sim p(\mathbf{x}_{n,t} | \Theta_{\setminus\{\mathbf{x}_{n,t}\}})$;
 for $t = 1$ **to** T **do**
 Draw $s_t^{2(q)} \sim p(s_t^2 | \Theta_{\setminus\{s_t^2\}})$;
 for $t = 1$ **to** T **do**
 Draw $\sigma_t^{2(q)} \sim p(\sigma_t^2 | \Theta_{\setminus\{\sigma_t^2\}})$;
 for $(\ell, r) = (1, 1)$ **to** (L, R) **do**
 Draw $\psi_{\ell,r}^{2(q)} \sim p(\psi_{\ell,r}^2 | \Theta_{\setminus\{\psi_{\ell,r}^2\}})$;

Result: $\{\Theta^{(q)}\}_{q=1}^{N_{\text{MC}}}$.

IV. EXPERIMENTS WITH SYNTHETIC DATA

Performance assessment has been conducted on a synthetic MTHS image composed of 20 images of size 30×30 with $L = 212$ bands. The proposed scenario mimics the emergence of a previously undetected material in specific images, thus interpreted as an outlier. To this end, 9 out of the 20 images have been corrupted by spatially sparse outliers. Moreover, each HS image has been generated as the linear mixture of 3 endmembers affected by smooth time-varying variability – generated as in [10] – whose abundances vary smoothly over time except for the corrupted pixels. An additive white Gaussian noise (ensuring a signal-to-noise ratio between 25 and 30 dB) has been finally added to the mixtures.

A. Compared methods

The proposed algorithm has been compared to VCA/FCLS [18], [19], SISAL/FCLS [20], the robust LMM (rLMM) [21] applied to each HS image independently, and the online unmixing (OU) described in [15] (in the setting [15, Table I]). The endmembers and abundances of OU and the proposed algorithm have been initialized with the results of VCA/FCLS obtained from the first image. For the initialization, we assume that the observations are not affected by outliers or variability, leading to: $\mathbf{X}_t = \mathbf{0}_{L,N}$, $\mathbf{dM}_t = \mathbf{0}_{L,R}$, $z_{n,t} = 0$. Moreover, based on the values of the initial endmembers and abundances, the hyperparameters have been initialized with the following typical orders of magnitude: $\sigma_t^2 = 10^{-4}$, $\psi_{\ell,r}^2 = 10^{-6}$, $\nu = 10^{-5}$, $\beta_t = 1.9$, $s_t^2 = 5 \times 10^{-3}$, $\epsilon_n = 10^{-4}$. Finally, the number of Monte-Carlo iterations has been set to $N_{\text{MC}} = 400$, with $N_{\text{bi}} = 350$ burn-in iterations. The performance of the different algorithms has been evaluated in terms of average spectral angle mapper (aSAM), abundance and variability global mean square errors (GMSEs) and

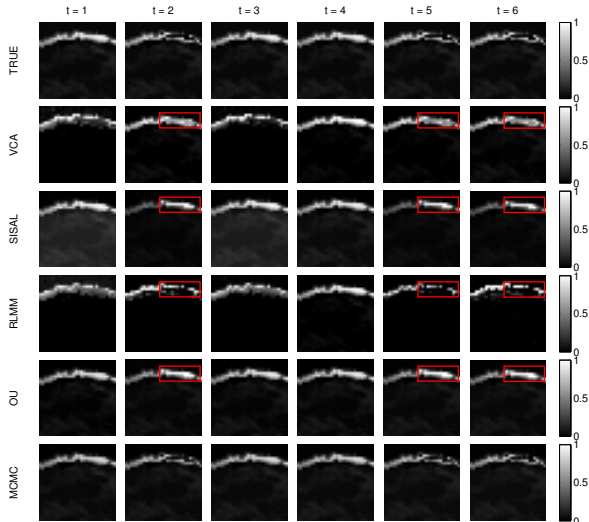


Fig. 2. Comparison of the abundance maps estimated for the third endmember by different unmixing strategies (in column) at $t = 1$ to 6 (in row). The areas corrupted by outliers are delineated in red.

quadratic reconstruction error (RE), respectively defined as

$$\begin{aligned} \text{aSAM}(\mathbf{M}) &= \frac{1}{R} \sum_{r=1}^R \arccos \left(\frac{\mathbf{m}_r^T \hat{\mathbf{m}}_r}{\|\mathbf{m}_r\|_2 \|\hat{\mathbf{m}}_r\|_2} \right) \\ \text{GMSE}(\mathbf{A}) &= \sum_{t=1}^T \frac{\|\mathbf{A}_t - \hat{\mathbf{A}}_t\|_F^2}{TRN} \\ \text{GMSE}(\mathbf{dM}) &= \sum_{t=1}^T \frac{\|\mathbf{dM}_t - \hat{\mathbf{dM}}_t\|_F^2}{TLR} \\ \text{RE} &= \sum_{t=1}^T \frac{\|\mathbf{Y}_t - \hat{\mathbf{Y}}_t\|_F^2}{TLN} \end{aligned}$$

where $\hat{\mathbf{Y}}_t$ is composed of the pixels reconstructed with the parameters estimated for the image t .

B. Results

A comparison between the abundance maps recovered by the different methods is given in Fig. 2¹, while the estimated endmembers are reported in Fig. 5. The estimated label maps are depicted with the ground truth in Fig. 3. The unmixing performance reported in Table I shows that the proposed method yields lower estimation errors, and presents an appreciable robustness to spatially sparse outliers. Indeed, the presence of outliers has significantly affected some of the endmembers estimated by VCA and SISAL, and altered the shape of the variability extracted by OU (see Figs. 5f, 5i and 5o respectively). The abundance maps recovered by VCA/FCLS and OU exhibit a notable sensitivity to the presence of outliers (see Fig. 2), since the abundance coefficients of the third endmember do not reflect the disappearance of the third endmember in the pixels where outliers have appeared instead (see Fig. 4 to appreciate the resulting abundance errors). The price to pay with the proposed algorithm is its computational complexity (significantly larger than those of the other algorithms, see Table I), which is dominated by the sampling of the variability terms.

¹Only the abundance maps associated with the 3rd endmembers have been depicted for brevity.

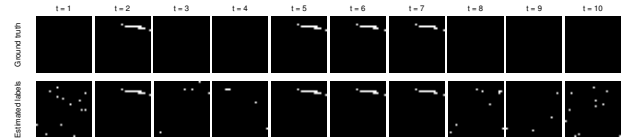


Fig. 3. Ground truth (first row) and estimated labels (second row) obtained with the proposed method for $t = 1$ to 10, where each column corresponds to a time instant [0 in black, 1 in white].



Fig. 4. Map of the re-scaled abundance estimation errors for the third endmember at time $t = 2$ (from left to right: true abundances, estimation error of VCA/FCLS, SISAL/FCLS, rLMM, OU and the proposed method). Except for the proposed method, the results exhibit notable errors in pixels corrupted by outliers (area in red).

TABLE I
SIMULATION RESULTS OBTAINED ON SYNTHETIC DATA
($\text{GMSE}(\mathbf{A}) \times 10^{-2}$, $\text{GMSE}(\mathbf{dM}) \times 10^{-3}$, $\text{RE} \times 10^{-3}$).

	aSAM(M) (°)	GMSE(A)	GMSE(dM)	RE	time (s)	
$R = 3$	VCA/FCLS	14.0	1.23	/	3.20	1
	SISAL/FCLS	11.9	2.40	/	0.47	2
	rLMM	14.5	1.52	/	0.04	238
	OU	12.9	0.30	1.64	0.26	58
	Proposed (MCMC)	8.03	0.17	0.20	0.11	1590

V. CONCLUSION AND FUTURE WORK

A Bayesian model has been introduced to simultaneously represent smooth and abrupt variations possibly affecting multitemporal hyperspectral (MTHS) images. The proposed unmixing strategy is intended to address datasets in which the same materials are repeatedly observed, except in a few pixels of the time series. An MCMC algorithm was investigated to generate samples from the posterior distribution of the proposed Bayesian model, thus allowing estimators of its unknown parameters to be constructed. The unmixing performance of the proposed approach was assessed on synthetic data and is quite promising. Future work includes applications of the proposed method to real datasets, and the study of distributed unmixing algorithms to analyze larger sequences of HS images.

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 Observ. in Remote Sens.*, vol. 5, no. 2, pp. 354–379, April 2012.
- [2] A. Zare and K. C. Ho, "Endmember variability in hyperspectral imagery," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 95–104, Jan. 2014.
- [3] 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, June 2010.
- [4] X. Du, A. Zare, P. Gader, and D. Dranishnikov, "Spatial and spectral unmixing using the beta compositional model," *IEEE J. Sel. Topics Appl. Earth Observ. in Remote Sens.*, vol. 7, no. 6, pp. 1994–2003, June 2014.
- [5] A. Halimi, N. Dobigeon, and J.-Y. Tourneret, "Unsupervised unmixing of hyperspectral images accounting for endmember variability," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 4904–4917, Dec. 2015.
- [6] B. Somers, M. Zortea, A. Plaza, and G. Asner, "Automated extraction of image-based endmember bundles for improved spectral unmixing," *IEEE J. Sel. Topics Appl. Earth Observ. in Remote Sens.*, vol. 5, no. 2, pp. 396–408, April 2012.

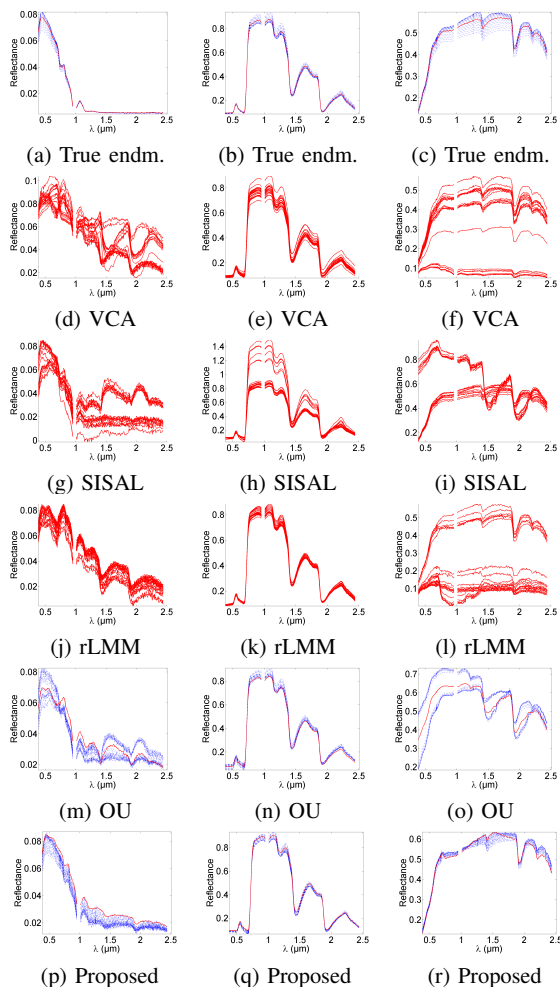


Fig. 5. Endmembers (red lines) and endmember + variability (blue dotted lines) extracted from the synthetic mixture by the compared methods (in row).

[7] J. Jin, B. Wang, and L. Zhang, "A novel approach based on Fisher discriminant null space for decomposition of mixed pixels in hyperspectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 4, pp. 699–703, Oct. 2010.

[8] T. Uezato, R. J. Murphy, A. Melkumyan, and A. Chlingaryan, "A novel spectral unmixing method incorporating spectral variability within endmember classes," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 5, pp. 2812–2831, May 2016.

[9] —, "A novel endmember bundle extraction and clustering approach for capturing spectral variability within endmember classes," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 11, pp. 6712–6731, Nov. 2016.

[10] P.-A. Thouvenin, N. Dobigeon, and J.-Y. Tournet, "Hyperspectral unmixing with spectral variability using a perturbed linear mixing model," *IEEE Trans. Signal Process.*, vol. 64, no. 2, pp. 525–538, Jan. 2016.

[11] L. Drumetz, M.-A. Veganzones, S. Henrot, R. Phlypo, J. Chanussot, and C. Jutten, "Blind hyperspectral unmixing using an extended linear mixing model to address spectral variability," *IEEE Trans. Image Process.*, vol. 25, no. 8, pp. 3890–3905, Aug. 2016.

[12] L. Drumetz, J. Chanussot, and C. Jutten, "Variability of the endmembers in spectral unmixing: recent advances," in *Proc. IEEE GRSS Workshop Hyperspectral Image Signal Process.: Evolution in Remote Sens. (WHISPERS)*, Los Angeles, United States, Aug. 2016. [Online]. Available: <https://hal.archives-ouvertes.fr/hal-01358368>

[13] S. Henrot, J. Chanussot, and C. Jutten, "Dynamical spectral unmixing of multitemporal hyperspectral images," *IEEE Trans. Image Process.*, vol. 25, no. 7, pp. 3219–3232, July 2016.

[14] A. Halimi, N. Dobigeon, J.-Y. Tournet, S. McLaughlin, and P. Honeine,

"Unmixing hyperspectral images accounting for temporal and spatial endmember variability," in *Proc. European Signal Process. Conf. (EU-SIPCO)*, Nice, France, Sept. 2015, pp. 1686–1690.

[15] P.-A. Thouvenin, N. Dobigeon, and J.-Y. Tournet, "Online unmixing of multitemporal hyperspectral images accounting for spectral variability," *IEEE Trans. Image Process.*, vol. 25, no. 9, pp. 3979–3990, Sept. 2016.

[16] C. Chenot, J. Bobin, and J. Rapin, "Robust sparse blind source separation," *IEEE Signal Process. Lett.*, vol. 22, no. 11, pp. 2172–2176, Nov. 2015.

[17] Y. Altmann, S. McLaughlin, and A. O. Hero, "Robust linear spectral unmixing using anomaly detection," *IEEE Trans. Comput. Imag.*, vol. 1, no. 2, pp. 74–85, June 2015.

[18] 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, April 2005.

[19] J. M. Bioucas-Dias and M. A. T. Figueiredo, "Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing," in *Proc. IEEE GRSS Workshop Hyperspectral Image Signal Process.: Evolution in Remote Sens. (WHISPERS)*, Reykjavik, Iceland, June 2010.

[20] J. M. Bioucas-Dias, "A variable splitting augmented Lagrangian approach to linear spectral unmixing," in *Proc. IEEE GRSS Workshop Hyperspectral Image Signal Process.: Evolution in Remote Sens. (WHISPERS)*, Grenoble, France, Aug. 2009.

[21] C. Févotte and N. Dobigeon, "Nonlinear hyperspectral unmixing with robust nonnegative matrix factorization," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 4904–4917, Dec. 2015.

[22] G. H. Golub and C. F. V. Loan, "An analysis of the total least squares problem," *SIAM J. Numer. Anal.*, vol. 17, no. 6, pp. 883 – 893, Dec. 1980.

[23] R. Tibshirani, "Regression shrinkage and selection via the LASSO," *J. Roy. Stat. Soc. Ser. B*, vol. 58, no. 1, pp. 267–288, 1996.

[24] J. P. Vila and P. Schniter, "Expectation-Maximization Gaussian-mixture approximate message passing," *IEEE Trans. Signal Process.*, vol. 61, no. 19, pp. 4658–4672, Oct. 2013.

[25] N. Dobigeon, A. O. Hero, and J.-Y. Tournet, "Hierarchical Bayesian sparse image reconstruction with application to MRFM," *IEEE Trans. Image Process.*, vol. 18, no. 9, pp. 2059–2070, Sept. 2009.

[26] J. J. Kormylo and J. M. Mendel, "Maximum likelihood detection and estimation of Bernoulli-Gaussian processes," *IEEE Trans. Inf. Theory*, vol. 28, no. 3, pp. 482–488, May 1982.

[27] M. Lavielle, "Bayesian deconvolution of Bernoulli-Gaussian processes," *Signal Process.*, vol. 33, no. 1, pp. 67–79, July 1993.

[28] S. Bourguignon and H. Carfantan, "Bernoulli-Gaussian spectral analysis of unevenly spaced astrophysical data," in *Proc. IEEE-SP Workshop Stat. and Signal Processing (SSP)*, Bordeaux, France, July 2005, pp. 811–816.

[29] C. Bazot, N. Dobigeon, and J.-Y. Tournet, "Bernoulli-Gaussian model for gene expression analysis," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*, Prague, Czech Republic, May 2011, pp. 5996–5999.

[30] Y. Altmann, M. Pereyra, and S. McLaughlin, "Bayesian nonlinear hyperspectral unmixing with spatial residual component analysis," *IEEE Trans. Comput. Imag.*, vol. 3, no. 1, pp. 174–185, Sept. 2015.

[31] 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 Process.*, vol. 57, no. 11, pp. 4355–4368, Nov. 2009.

[32] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *journal of ACM*, vol. 58, no. 1, pp. 1–37, 2009.

[33] P.-A. Thouvenin, N. Dobigeon, and J.-Y. Tournet, "A hierarchical bayesian model accounting for endmember variability and abrupt spectral changes to unmix multitemporal hyperspectral images – complementary results and supporting material," University of Toulouse, IRIT/INP-ENSEEIH, Tech. Rep., May 2017. [Online]. Available: <http://thouvenin.perso.enseiht.fr/papers/Thouvenin2017TR.pdf>