A Hierarchical Bayesian Model Accounting for Endmember Variability and Abrupt Spectral Changes to Unmix Multitemporal Hyperspectral Images – Complementary results and supporting material

Pierre-Antoine Thouvenin, *Member, IEEE*, Nicolas Dobigeon, *Senior Member, IEEE* and Jean-Yves Tourneret, *Senior Member, IEEE*

Abstract

Hyperspectral unmixing is a blind source separation problem which consists in estimating the reference spectral signatures contained in a hyperspectral image, as well as their relative contribution to each pixel according to a given mixture model. In practice, the process is further complexified by the inherent spectral variability of the observed scene and the possible presence of outliers. More specifically, multi-temporal hyperspectral images, i.e., sequences of hyperspectral images acquired over the same area at different time instants, are likely to simultaneously exhibit moderate endmember variability and abrupt spectral changes either due to outliers or to significant time intervals between consecutive acquisitions. Unless properly accounted for, these two perturbations can significantly affect the unmixing process. In this context, we propose a new unmixing model for multitemporal hyperspectral images accounting for smooth temporal variations, construed as spectral variability, and abrupt spectral changes interpreted as outliers. The proposed hierarchical Bayesian model is inferred using a Markov chain Monte-Carlo (MCMC) method allowing the posterior of interest to be sampled and Bayesian estimators to be approximated. A comparison with unmixing techniques

This report provides complementary results and supporting material to the paper [1].

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 Defence.

The authors are with the University of Toulouse, IRIT/INP-ENSEEIHT, 31071 Toulouse, France. (e-mail: {pierreantoine.thouvenin, Nicolas.Dobigeon, Jean-Yves.Tourneret}@enseeiht.fr

from the literature on synthetic and real data allows the interest of the proposed approach to be appreciated.

Index Terms

Hyperspectral imagery, multitemporal images, endmember variability, Markov chain Monte-Carlo (MCMC) methods.

I. Introduction

Acquired in hundreds of contiguous spectral bands, hyperspectral (HS) images have received an increasing interest due to the significant spectral information they convey, which is somewhat mitigated by their lower spatial resolution in remote sensing applications. This limitation, combined with possibly complex interactions between the incident light and the observed materials, implies that the observed spectra are mixtures of several signatures corresponding to distinct materials. Spectral unmixing then consists in identifying a limited number of reference spectral signatures composing the data – referred to as *endmembers* – and their abundance fractions in each pixel according to a predefined mixture model. The choice of a specific model generally reflects the practitioners' prior knowledge on the environmental factors possibly affecting the acquisitions, such as declivity or multiple reflections. Traditionally, a linear mixing model (LMM) is adopted since it is appropriate to describe hyperspectral data when the declivity of the scene and microscopic interactions between the observed materials are negligible [2]. Depending on the applications, various models have also been investigated to capture higher order interactions (i.e., nonlinearities) between the incident light and the observed materials (see [3], [4] for recent reviews on this topic). However, varying acquisition conditions, such as local illumination variations or the natural evolution of the scene, may significantly alter the shape and the amplitude of the acquired signatures [5], [6], thus affecting the extracted endmembers. Endmember variability has hitherto been extensively considered within a single HS image, either in a deterministic [7]–[10] or a statistical setting [11]–[13].

Recent works also considered temporal variability by exploiting the possibilities offered by multitemporal HS (MTHS) images [14], [15]. From a hyperspectral unmixing perspective, MTHS images, i.e., sequences of HS images acquired over the same area at different time instants, can be of interest to exploit information redundancy between consecutive images (e.g., through features exhibiting moderate or smooth temporal variations as in [16], [17]) while allowing the endmember temporal evolution to be characterized. For instance, MTHS have been recently exploited to improve endmember unmixing results [14], [15], [18] and used in a change detection problem involving two HS images [19], [20].

Even though the approaches proposed in [14], [15], [18] specifically allow smooth temporal variations of some of the mixture parameters to be considered, they do not account for abrupt spectral changes either due to outliers or to possibly significant time intervals between two consecutive images. In practice, such situations can be reasonably expected, depending on the acquisition dates and possible climatic hazards, e.g., when vegetation or water is present in the observed scene. Unless specifically accounted for, this situation frequently observed in real datasets has a significant impact on the recovered endmembers, which motivates the present work. Inspired by [14], [21], [22], and based on an original interpretation of the unmixing problem under study, our contribution consists in jointly accounting for smooth endmember variations – construed as temporal endmember variability – and abrupt changes interpreted as outliers (e.g., significant variability within a single image or presence of non-linearities) using a carefully designed hierarchical Bayesian model. More precisely, we focus our analysis on scenes in which mostly the same materials are expected to be observed from an image to another. In this context, using the endmembers extracted from the reference scene as a starting point to unmix the whole MTHS image constitutes a reasonable attempt to generalize the analyses previously conducted for a single image. On the one hand, the endmembers identified in each single image can *in fine* be considered as time-varying instances of reference signatures shared by the different images, thus justifying the use of a modified version of the perturbed linear mixing model (PLMM) proposed in [15]. This formulation will notably allow smooth spectral variations occurring over time to be captured, leading to competitive results when compared to methods analyzing the images individually. On the other hand, the signatures corresponding to materials appearing in only a few images, which induce abrupt spectral changes, can be regarded as outliers with respect to the commonly shared endmembers. This paper studies a new Bayesian model allowing both spectral variability and presence of outliers to be considered in the unmixing of MTHS images. The resulting unmixing task is solved using a Markov chain Monte-Carlo (MCMC) allowing the posterior of interest to be sampled and Bayesian estimators to be approximated.

The paper is organized as follows. The mixing model considered in this paper is introduced in Section II, and the associated hierarchical Bayesian model is developed in Section III. Section IV investigates a Gibbs sampler to solve the resulting mixed integer non-linear problem. The performance of the proposed approach on synthetic and real data is studied in Sections V and VI. In particular, the results obtained with the proposed method are compared to those of the VCA/FCLS algorithm [23], [24], the SISAL/FCLS algorithm [25], the algorithm associated with the robust LMM (RLMM) proposed in [26] and the MTHS optimization method [15]. Finally, Section VII concludes this work and outlines further research perspectives.

II. Problem statement

We consider a sequence of HS images acquired at T different time instants over the same area, where mostly the same materials are expected to be observed over time. In the following, at most R endmembers are assumed to be shared between the T images composing the sequence, where R is *a priori* known. Since the observed instances of a given endmember can be reasonably expected to vary from an image to another, we propose to account for smooth endmember spectral variations via a modified version of the perturbed linear mixing model (PLMM) proposed in [10], [15]. Inspired by the total least squares problem [27], the PLMM consists in representing each pixel $\mathbf{y}_{nt}$ by a linear combination of the R endmembers – denoted by $\mathbf{m}_r$ – affected by an additive error term $\mathbf{dm}_{r,t}$ accounting for temporal endmember variability. However, this model shows notable limitations when the vector $\mathbf{y}_{n,t}$ is affected by abrupt changes. Consequently, this paper investigates a new unmixing model jointly accounting for endmember variability and abrupt changes possibly affecting MTHS images. To this end, the proposed model is a generalized PLMM, which includes an additional term $\mathbf{x}_{n,t}$ to capture significant deviations from the LMM, i.e., significant spatial variability or non-linearities within each image [21], [26]. The resulting observation model can thus be written

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

for $n = 1, \ldots, N$ and $t = 1, \ldots, T$, where $\mathbf{y}_{n,t}$ denotes the nth image pixel at time t, $\mathbf{m}_r$ is the rth endmember, $a_{r,n,t}$ is the proportion of the rth endmember in the nth pixel at time t, $\mathbf{dm}_{r,t}$ denotes the perturbation of the rth endmember at time t, and $\mathbf{x}_{n,t}$ denotes the contribution of outliers in the nth pixel at time t. Finally, $\mathbf{b}_{n,t}$ represents an additive noise resulting from the data acquisition and the modeling errors. The so-called robust PLMM can be written

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

where $\mathbf{Y}_t = [\mathbf{y}_{1,t}, \ldots, \mathbf{y}_{N,t}]$ is an $L \times N$ matrix containing the pixels of the tth image, $\mathbf{M}$ denotes an $L \times R$ matrix containing the endmembers that are common to all the images of the sequence, $\mathbf{A}_t$ is an $R \times N$ matrix composed of the abundance vectors $\mathbf{a}_{n,t}$, $\mathbf{dM}_t$ is an $L \times R$ matrix whose columns contain the variability inherent to the tth image, $\mathbf{X}_t$ is an $L \times N$ matrix whose columns are the outliers present in the image t, and $\mathbf{B}_t$ is an $L \times N$ matrix accounting for the noise at time t. The constraints considered to reflect physical considerations are

$$\begin{aligned}
&\mathbf{A}_t \succeq \mathbf{0}_{R,N}, \quad \mathbf{A}_t^T \mathbf{1}_R = \mathbf{1}_N, \, \forall t \in \{1, \ldots, T\} \\
&\mathbf{M} \succeq \mathbf{0}_{L,R}, \quad \mathbf{M} + \mathbf{dM}_t \succeq \mathbf{0}_{L,R}, \, \forall t \in \{1, \ldots, T\} \\
&\mathbf{X}_t \succeq \mathbf{0}_{L,N}, \quad \forall t \in \{1, \ldots, T\}
\end{aligned} \tag{3}$$

where $\succeq$ denotes a term-wise inequality. Note that the outlier term $\mathbf{X}_t$ is intended to describe abrupt changes due for instance to the appearance of one or several new endmembers that were not present

in the reference image. This justifies the corresponding non-negativity constraint, similar to the one imposed on the other endmembers. Note however that different phenomena not considered in this work, possibly represented by the terms $\mathbf{X}_t$, can induce a decrease in the total reflectance, e.g., shadowing effects or some nonlinearities as detailed in [28]. To address this case, the non-negativity constraint on the outlier terms $\mathbf{X}_t$ should be removed.

Given the mixture model (2), the unmixing problem considered in this work consists in inferring the abundances $\mathbf{A}_t$, the endmembers $\mathbf{M}$, the variability $\mathbf{dM}_t$ and the outliers $\mathbf{X}_t$ from the observations $\mathbf{Y}_t$, $t = 1, \ldots, T$. In the next section, this problem is tackled in a Bayesian framework to easily incorporate all the prior knowledge available on the mixture parameters.

III. Bayesian model

This section details the specific structure imposed on the parameters to be inferred *via* appropriate prior distributions. Note that dependencies with respect to constant parameters are omitted in the following paragraphs to simplify the notations.

A. Likelihood

Assuming the additive noise $\mathbf{b}_{n,t}$ is distributed according to a Gaussian distribution $\mathbf{b}_{n,t} \sim \mathcal{N}(\mathbf{0}_L, \sigma_t^2 \mathbf{I}_L)$, the observation model (2) leads to

$$\mathbf{y}_{n,t} \mid \mathbf{M}, \mathbf{dM}_t, \mathbf{A}_t, \mathbf{X}_t, \sigma_t^2 \sim \mathcal{N}\Big((\mathbf{M} + \mathbf{dM}_t)\mathbf{a}_{n,t} + \mathbf{x}_{n,t}, \sigma_t^2 \mathbf{I}_L\Big).$$

In addition, assuming prior independence between the pixels within each image and between the images $\mathbf{Y}_t$ themselves, the likelihood function of all images $\underline{\mathbf{Y}} = [\mathbf{Y}_1, \ldots, \mathbf{Y}_T]$ is

$$\begin{aligned} p(\underline{\mathbf{Y}} \mid \Theta) \propto & \prod_{t=1}^{T} (\sigma_t^2)^{-NL/2} \times \\ & \exp\Big(-\frac{1}{2\sigma_t^2} \|\mathbf{Y}_t - (\mathbf{M} + \mathbf{dM}_t)\mathbf{A}_t - \mathbf{X}_t\|_{\mathrm{F}}^2\Big) \end{aligned} \tag{4}$$

where the underline notation stands for the overall set of the corresponding parameters, $\|\cdot\|_{\mathrm{F}}$ is the Frobenius norm, $\Theta = \{\Theta_p, \Theta_h\}$ and

$$\Theta_p = \{\mathbf{M}, \underline{\mathbf{dM}}, \underline{\mathbf{A}}, \underline{\mathbf{X}}, \boldsymbol{\sigma}^2, \mathbf{Z}\}, \ \Theta_h = \{\boldsymbol{\Psi}^2, \mathbf{s}^2, \boldsymbol{\beta}\} \tag{5}$$

denote the parameters and hyperparameters whose priors are defined in the following paragraphs. Note that the independence assumption between the observed images conditionally on the unknown parameters is justified by the fact that the sequence of images has been acquired by possibly different sensors at different time instants.

Remark. The proposed method can easily accommodate different structures for the noise covariance matrix (e.g., diagonal or full covariance matrix) in case the correlation between the spectral bands is

significant (see, e.g., [29]). However, this modification would increase the computational and memory cost of the estimation algorithm introduced in Section IV.

B. Parameter priors

1) Abundances: We propose to promote smooth temporal variations of the abundances between successive time instants for pixels that are not classified as outliers. To this end, we first introduce the binary latent variables $\mathbf{z}_t \in \{0,1\}^N$ to describe the support of the outliers (i.e., $z_{n,t} = 0$ in the absence of outliers in the pixel (n,t), 1 otherwise). With this notation, we introduce a new abundance prior defined for $n = 1, \dots N$ as

$$\mathbf{a}_{n,1} \mid z_{n,1} = 0 \sim \mathcal{U}_{\mathcal{S}_R} \tag{6}$$

$$\mathbf{a}_{n,t} \mid z_{n,t} = 1 \sim \mathcal{U}_{\widetilde{\mathcal{S}_R}}, \text{ for } t = 1, \dots, T \tag{7}$$

$$\begin{aligned} &p\left(\mathbf{a}_{n,t} \mid z_{n,t} = 0, \underline{\mathbf{A}}_{\setminus\{\mathbf{a}_{n,t}\}}\right) \propto \exp\Bigg\{-\frac{1}{2\varepsilon_n^2} \times \\ &\left([\mathscr{T}_{n,t}^1 \neq \emptyset] \| a_{n,t} - a_{n,\tau_{n,t}^1} \|_2^2\right)\Bigg\} \mathbb{1}_{\mathcal{S}_R}(\mathbf{a}_{n,t}), \text{ for } t \geq 2 \end{aligned} \tag{8}$$

where $\mathcal{U}_{\mathcal{S}_R}$ denote the uniform distribution on the set $\mathcal{S}_R$, $\mathbb{1}_{\mathcal{S}_R}$ is the indicator function of the set $\mathcal{S}_R$, $[\mathscr{P}]$ denotes the Iverson bracket applied to the logical proposition $\mathscr{P}$, i.e.,

$$[\mathscr{P}] = \begin{cases} 1, & \text{if } \mathscr{P} \text{ is true;} \\ 0, & \text{otherwise} \end{cases}$$

and

$$\mathcal{S}_R = \{\mathbf{x} \in \mathbb{R}^R \mid \forall i, x_i \geq 0 \text{ and } \mathbf{x}^{\mathrm{T}} \mathbf{1}_R = 1\} \tag{9}$$

$$\widetilde{\mathcal{S}_R} = \{\mathbf{x} \in \mathbb{R}^R \mid \forall i, x_i \geq 0 \text{ and } \mathbf{x}^{\mathrm{T}} \mathbf{1}_R \leq 1\} \tag{10}$$

$$\mathscr{T}_{n,t}^1 = \{\tau < t \mid z_{n,\tau} = 0\}, \quad \tau_{n,t}^1 = \max_{\tau \in \mathscr{T}_{n,t}^1} \tau. \tag{11}$$

By convention, we set $\mathscr{T}_{n,t}^1 = \emptyset$ when $t = 1$. To be more explicit, consider an image at time t and a pixel n within this image which is not corrupted by outliers (i.e., $z_{n,t} = 0$). For $t = 1$, a uniform distribution defined in the unit simplex is selected to reflect the absence of specific prior knowledge while accounting for the related constraints in (3). For $t > 1$, smooth variations of $\mathbf{a}_{n,t}$ are promoted *via* a one-dimensional Gaussian Markov field [13], [30] penalizing the Euclidean distance between $\mathbf{a}_{n,t}$ and the abundance of the last corresponding outlier-free pixel in the preceding images of the sequence, i.e., at time instant $\tau_{n,t}^1$. On the contrary, when outliers are present in the pixel (n,t) ($\mathbf{x}_{n,t} = 1$), the usual abundance sum-to-one constraint is relaxed ($\mathbf{a}_{n,t}^T \mathbf{1}_R \leq 1$) so that the prior allows cases in which the linear model does not exhaustively describe the data to be addressed. Note that the *a priori* independence assumptions between the abundance vectors $\mathbf{a}_{n,t}$ (conditionally to the labels

$z_{n,t}$) is reasonable from a physical point of view, since they can evolve independently from a pixel to another. In the following, the joint abundance prior is denoted by

$$p(\underline{\mathbf{A}} \mid \mathbf{Z}) = \prod_{n=1}^{N} \Bigg[\prod_{\substack{j=1 \\ t_j : z_{n,t_j}=1}}^{J_n} p(\mathbf{a}_{n,t_j} \mid z_{n,t_j}=1) \Bigg] \times \Bigg[\prod_{\substack{i=1 \\ t_i : z_{n,t_i}=0}}^{I_n} p(\mathbf{a}_{n,t_i} \mid \mathbf{a}_{n,t_{i-1}}, z_{n,t_i}=0) \Bigg] \tag{12}$$

with $I_n = \sharp\{t : z_{n,t} = 0\}$, $J_n = T - I_n$ and $\sharp$ denotes the cardinal operator. Note that the events $[z_{n,t} = 0]$ and $[\mathbf{x}_{n,t} = \mathbf{0}_L]$ (respectively $[z_{n,t} = 1]$ and $[\mathbf{x}_{n,t} \neq \mathbf{0}_L]$) are equivalent, which allows $p(\underline{\mathbf{A}} \mid \underline{\mathbf{X}})$ to be defined.

In the following paragraph, the latent variables $z_{n,t}$ are assigned a specific prior to reflect the fact that outliers are *a priori* assumed to represent a limited number of pixels within the sequence of image.

2) Outliers $\underline{\mathbf{X}}$ *and label maps* $\mathbf{Z}$*:* Similarly to [21], outliers are a priori assumed to be spatially sparse. Different approaches have been proposed in the literature to include this prior knowledge, either relying on the ℓ_1 penalty (such as the LASSO [31]) or on mixtures of probability distributions involving a Dirac mass at zero and a continuous probability distribution [32] (such as the Bernoulli-Laplace [33] or Bernoulli-Gaussian distributions [34], [35], extensively used in the literature [36]–[38]). In this work, we propose to assign the following prior to the outliers $\mathbf{x}_{n,t}$ to promote spatial sparsity

$$p(\mathbf{x}_{n,t} \mid z_{n,t}, s_t^2) = (1 - z_{n,t})\delta(\mathbf{x}_{n,t}) + z_{n,t}\,\mathcal{N}_{\mathbb{R}_+^L}(\mathbf{0}_L, s_t^2) \tag{13}$$

where $\mathcal{N}_{\mathbb{R}_+^L}$ denotes a Gaussian distribution truncated to the set $\mathbb{R}_+^L$. Note that $z_{n,t} = 1$ if an outlier is present in the corresponding pixel, and 0 otherwise. The proposed prior notably allows outliers to be *a priori* described by a truncated Gaussian distribution when $z_{n,t} = 1$, since the outliers $\mathbf{x}_{n,t}$ are mainly due to the appearance of new endmembers (i.e., that were not present in the reference image). With this context in mind, we further propose to promote spatial correlations between the outliers' support, since new materials are likely to appear in multiple contiguous pixels. The binary label maps $\mathbf{z}_t \in \mathbb{R}^N$ $(t = 1, \ldots, T)$ are consequently modeled as Ising-Markov random fields [13], [39], [40], for which the Hammersley-Clifford theorem yields

$$p(\mathbf{z}_t \mid \beta_t) = \frac{1}{C(\beta_t)} \exp\left(\beta_t \sum_{n=1}^{N} \sum_{k \in \mathcal{V}(n)} \delta(z_{n,t} - z_{k,t}) \right) \tag{14}$$

where $\mathcal{V}(n)$ denotes the 4-neighbourhood of the pixel n, and $C(\beta_t)$ is the partition function [41]. In practice, the outlier terms $\mathbf{x}_{n,t}$ can be assumed to be *a priori* independent conditionally on $\mathbf{z}_{n,t}$ (since the values of outliers are not *a priori* correlated over space and time from a physical point of

view). A similar assumption can be made on the labels $\mathbf{z}_t$, which leads to

$$p(\underline{\mathbf{X}} \mid \mathbf{Z}, \mathbf{s}^2) = \prod_{n,t} p(\mathbf{x}_{n,t} \mid z_{n,t}, s_t^2) \tag{15}$$

$$p(\mathbf{Z} \mid \boldsymbol{\beta}) = \prod_{t} p(\mathbf{z}_t \mid \beta_t) \tag{16}$$

with $\mathbf{Z} \in \mathbb{R}^{N \times T}$ and $\boldsymbol{\beta} \in \mathbb{R}^T$. Note that the prior (13) leads to the following result, which will be useful to sample the label maps in Section IV-E

$$p(\mathbf{x}_{n,t} \mid \mathbf{z}_{\backslash n,t}, s_t^2, \beta_t) = (1-\omega_{n,t})\delta(\mathbf{x}_{n,t}) + \omega_{n,t}\mathcal{N}_{\mathbb{R}_+^L}(\mathbf{0}_L, s_t^2\mathbf{I}_L)$$

where $\mathbf{z}_{\backslash n,t}$ denotes the label map $\mathbf{z}_t$ whose nth entry has been removed, and

$$\omega_{n,t} = \frac{1}{C}\exp\left(\beta_t \sum_{k \in \mathcal{V}(n)} \delta(1 - z_{k,t})\right). \tag{17}$$

with $C = \sum_{i=0}^{1} \exp\left(\beta_t \sum_{k \in \mathcal{V}(n)} \delta(i - z_{k,t})\right)$.

3) Endmembers: A non-informative prior is adopted for the endmember matrix $\mathbf{M}$ to reflect the absence of specific prior knowledge about the spectral signatures contained in the image. More precisely, as in previous studies related to hyperspectral unmixing [13], [21], we consider the following truncated multivariate Gaussian distribution

$$\mathbf{m}_r \sim \mathcal{N}_{\mathbb{R}_+^L}(\mathbf{0}_L, \xi\mathbf{I}_L), \text{ for } r = 1, \ldots, R \tag{18}$$

where ξ is set to a sufficiently large value to ensure an uninformative prior (e.g., $\xi = 1$). Assuming the endmembers $\mathbf{m}_r$ are independent (which is physically reasonable since the endmembers characterize different materials), the joint prior for the endmembers can be written as

$$p(\mathbf{M}) = \prod_{r=1}^{R} p(\mathbf{m}_r). \tag{19}$$

In addition, the endmembers can be *a priori* assumed to live in a subspace of dimension $K = R - 1$ [42] whose practical determination can be performed by a principal component analysis (PCA) or a robust PCA (rPCA) [43]. This dimensionality reduction step is essentially aimed at reducing the computational complexity of the proposed approach. More explicitly, the PCA applied to the original data $\underline{\mathbf{Y}}$ leads to a decomposition which can be expressed as [42]

$$\mathbf{m}_r = \mathbf{U}\mathbf{e}_r + \check{\mathbf{y}}, \quad \check{\mathbf{y}} = (\mathbf{I}_L - \mathbf{U}\mathbf{U}^{\mathrm{T}})\bar{\mathbf{y}}, \quad \mathbf{U}^{\mathrm{T}}\mathbf{U} = \mathbf{I}_K \tag{20}$$

where $\mathbf{U}$ denotes a basis of the subspace of dimension K and $\bar{\mathbf{y}}$ denotes the average spectral signature obtained from $\underline{\mathbf{Y}}$. Note that using an rPCA would result in similar expressions (modulo a simple change of notations). The projected endmembers $\mathbf{e}_r$ are then assigned the following truncated Gaussian prior, which ensures the non-negativity of the endmembers

$$\mathbf{e}_r \sim \mathcal{N}_{\mathcal{E}_r}(\mathbf{0}_K, \xi\mathbf{I}_K), \text{ for } r = 1, \ldots, R \tag{21}$$

with

$$\mathcal{E}_r = [e_{1,r}^-, e_{1,r}^+] \times \ldots \times [e_{K,r}^-, e_{K,r}^+] \tag{22}$$

$$e_{k,r}^- = \max_{\ell \in \mathcal{U}_k^+} \left(-\frac{\check{y}_\ell + \sum_{j \neq k} u_{\ell,j} e_{j,r}}{u_{\ell,k}} \right) \tag{23}$$

$$e_{k,r}^+ = \min_{\ell \in \mathcal{U}_k^-} \left(-\frac{\check{y}_\ell + \sum_{j \neq k} u_{\ell,j} e_{j,r}}{u_{\ell,k}} \right) \tag{24}$$

$$\mathcal{U}_k^- = \{r : u_{k,r} < 0\}, \quad \mathcal{U}_k^+ = \{r : u_{k,r} > 0\}. \tag{25}$$

4) Endmember variability: We consider a prior for the vectors $\mathbf{dm}_{r,t}$ (associated with the endmember variability) promoting smooth temporal variations while accounting for the term-wise non-negativity of the observed endmembers (i.e., $\mathbf{m}_r + \mathbf{dm}_{r,t} \succeq \mathbf{0}_{L,R}$), expressed as

$$\mathbf{dm}_{r,1} \mid \mathbf{m}_r \sim \mathcal{N}_{\mathcal{I}_r}(\mathbf{0}_L, \nu \mathbf{I}_L) \tag{26}$$

$$dm_{\ell,r,t} \mid m_{\ell,r}, dm_{\ell,r,(t-1)}, \psi_{\ell,r}^2 \sim \mathcal{N}_{\mathcal{I}_{\ell,r}}\left(dm_{\ell,r,(t-1)}, \psi_{\ell,r}^2\right) \tag{27}$$

for $\ell = 1, \ldots, L$, $r = 1, \ldots, R$, $t = 1, \ldots, T$, where $\mathcal{I}_r = \mathcal{I}_{1,r} \times \ldots \times \mathcal{I}_{L,r}$ and $\mathcal{I}_{\ell,r} = [-m_{\ell,r}, +\infty)$. Assuming *a priori* independence between the different endmember variabilities (since the variability can be independent from a material to another), the joint variability prior can finally be expressed as

$$\begin{aligned} p(\underline{\mathbf{dM}} \mid \mathbf{M}, \mathbf{\Psi}^2) = \prod_{r=1}^{R} \Big[& p(\mathbf{dm}_{r,1} \mid \mathbf{m}_r) \\ & \times \prod_{t=2}^{T} p(\mathbf{dm}_{r,t} \mid \mathbf{m}_r, \mathbf{dm}_{r,(t-1)}, \boldsymbol{\psi}_r^2) \Big]. \end{aligned} \tag{28}$$

5) Noise variance: A non-informative inverse-gamma conjugate prior is selected for the noise variance

$$\sigma_t^2 \sim \mathcal{IG}(a_\sigma, b_\sigma) \tag{29}$$

for $t = 1, \ldots, T$, with $a_\sigma = b_\sigma = 10^{-3}$ in order to ensure a weakly informative prior. The noise variances σ_t^2 can be assumed to be *a priori* independent (given the absence of *a priori* correlation between the noise in different images), thus leading to

$$p(\boldsymbol{\sigma}^2) = \prod_t p(\sigma_t^2). \tag{30}$$

C. Hyperparameters

In order to complete the description of the proposed hierarchical Bayesian model, we consider the following generic priors for the different hyperparameters. Note that the *a priori* independence assumptions made in this section are more of a computational nature, i.e., aimed at simplifying the estimation procedure detailed in the next section.

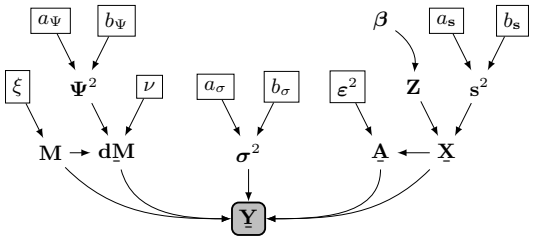


Fig. 1. Directed acyclic graph associated with the proposed Bayesian model (fixed parameters appear in boxes).

(i) Non-informative conjugate inverse-gamma priors for the variability variances $\boldsymbol{\Psi}^2$ and the outlier variances $\mathbf{s}^2$, i.e., for $\ell = 1, \ldots, L$, $r = 1, \ldots, R$ and $t = 1, \ldots, T$

$$\psi^2_{\ell,r} \sim \mathcal{IG}(a_\Psi, b_\Psi), \quad s^2_t \sim \mathcal{IG}(a_{\mathbf{s}}, b_{\mathbf{s}}) \tag{31}$$

where $\mathcal{IG}(a_\Psi, b_\Psi)$ denotes the inverse gamma distribution and $a_\Psi = b_\Psi = a_{\mathbf{s}} = b_{\mathbf{s}} = 10^{-3}$. Classical independence assumptions for the different hyperparameters lead to

$$p(\boldsymbol{\Psi}^2) = \prod_{\ell,r} p(\psi^2_{\ell,r}), \quad p(\mathbf{s}^2) = \prod_t p(s^2_t). \tag{32}$$

(ii) A uniform prior for the granularity parameter of a Potts-Markov random field (*a fortiori* of an Ising-Markov random field). Previous studies have shown that it is reasonable to constrain the granularity parameter to belong to the interval $[0, 2]$ [44], leading to

$$\beta_t \sim \mathcal{U}_{[0,2]}, \text{ for } t = 1, \ldots, T. \tag{33}$$

Assuming the granularity parameters are *a priori* independent for different time instants finally yields

$$p(\boldsymbol{\beta}) = \prod_t p(\beta_t). \tag{34}$$

D. Joint posterior distribution

Applying Bayes' theorem, the joint posterior distribution of the parameters of interest is given by

$$\begin{aligned} p(\Theta \mid \underline{\mathbf{Y}}) \propto\ & p(\underline{\mathbf{Y}} \mid \Theta) p(\underline{\mathbf{A}} \mid \underline{\mathbf{X}}) p(\underline{\mathbf{X}} \mid \mathbf{Z}, \mathbf{s}^2) p(\mathbf{s}^2) \\ & \times p(\mathbf{Z} \mid \boldsymbol{\beta}) p(\boldsymbol{\beta}) p(\underline{\mathbf{dM}} \mid \mathbf{M}, \boldsymbol{\Psi}^2) p(\mathbf{M}) p(\boldsymbol{\Psi}^2) p(\boldsymbol{\sigma}^2). \end{aligned} \tag{35}$$

The complexity of the proposed Bayesian model summarized in the directed acyclic graph of Fig. 1 and its resulting posterior (35) prevent a simple computation of the maximum *a posteriori* (MAP) or minimum mean square (MMSE) estimators. For instance, the optimization problem associated with the determination of the MAP estimator of Θ is clearly complex, since the negative log-posterior is non-convex and parameterized by mixed continuous and discrete variables. In this context, classical matrix factorization techniques such as [26] cannot be used efficiently. An MCMC method is consequently adopted to sample the posterior (35) and to build estimators of the parameters involved in the proposed Bayesian model using the generated samples.

IV. Gibbs sampler

This section studies a Gibbs sampler, which is guaranteed to produce samples asymptotically distributed according to the target distribution (35). This sampler described in Algo. 1 consists in generating samples distributed according to the conditional distribution of each parameter of interest. Section IV-A introduces the proposed sampling method, and the conditional distributions of all the parameters of interest (see Fig. 1) are detailed in the following paragraphs.

A. Bayesian inference and parameter estimation

The main steps of the proposed Gibbs sampler are summarized in Algo. 1. Similarly to [21], the sequence $\{\Theta^{(q)}\}_{q=N_{\text{bi}}+1}^{N_{\text{MC}}}$ generated by the proposed sampler (i.e., after N_{bi} burn-in iterations) is used to approximate the MMSE estimators of the different unknown parameters $\mathbf{M}$, $\mathbf{A}_t$, $\mathbf{dM}_t$ and $\mathbf{X}_t$ by replacing the expectations by empirical averages.

$$\widehat{\mathbf{M}}^{\text{MMSE}} \simeq \frac{1}{N_{\text{MC}} - N_{\text{bi}}} \sum_{q=N_{\text{bi}}+1}^{N_{\text{MC}}} \mathbf{M}^{(q)} \tag{36}$$

$$\widehat{\mathbf{A}}_t^{\text{MMSE}} \simeq \frac{1}{N_{\text{MC}} - N_{\text{bi}}} \sum_{q=N_{\text{bi}}+1}^{N_{\text{MC}}} \mathbf{A}_t^{(q)} \tag{37}$$

$$\widehat{\mathbf{dM}}_t^{\text{MMSE}} \simeq \frac{1}{N_{\text{MC}} - N_{\text{bi}}} \sum_{q=N_{\text{bi}}+1}^{N_{\text{MC}}} \mathbf{dM}_t^{(q)} \tag{38}$$

$$\widehat{\mathbf{X}}_t^{\text{MMSE}} \simeq \frac{1}{N_{\text{MC}} - N_{\text{bi}}} \sum_{q=N_{\text{bi}}+1}^{N_{\text{MC}}} \mathbf{X}_t^{(q)}. \tag{39}$$

This choice is justified by the fact that MAP estimators computed using MCMC algorithms are often less accurate when the number of unknown parameters is relatively large [45]. Finally, the following *marginal maximum a posterior* (mMAP) estimator is considered for the label maps

$$\widehat{z}_{n,t}^{\text{mMAP}} = \underset{z_{n,t}\in\{0,1\}}{\arg\max}\, p\left(z_{n,t} \mid \mathbf{y}_{n,t}, \Theta_{\backslash\{z_{n,t}\}}\right). \tag{40}$$

It is approximated by

$$\widehat{z}_{n,t}^{\text{mMAP}} \simeq \begin{cases} 0, \text{ if } \sharp\{q > N_{\text{bi}} : z_{n,t}^{(q)} = 0\} \leq \frac{N_{\text{MC}} - N_{\text{bi}}}{2} \\ 1, \text{ otherwise.} \end{cases} \tag{41}$$

B. Sampling the abundances $\underline{\mathbf{A}}$

The likelihood function (III-A) combined with the prior given in Section III-B1 leads to the following conditional distribution for the abundances

$$\mathbf{a}_{n,t} \mid \mathbf{y}_{n,t}, \Theta_{\backslash\{\mathbf{a}_{n,t}\}} \sim \mathcal{N}_{\mathcal{S}_R}(\boldsymbol{\mu}_{n,t}^{(\mathbf{A})}, \boldsymbol{\Lambda}_{n,t}) \tag{42}$$

$$\boldsymbol{\Lambda}_{n,t}^{-1} = \frac{1}{\sigma_t^2}\mathbf{M}_t^{\mathrm{T}}\mathbf{M}_t + \frac{1}{\varepsilon_n^2}\Big([\mathscr{T}_{n,t}^1 \neq \emptyset] + [\mathscr{T}_{n,t}^2 \neq \emptyset]\Big)\mathbf{I}_R \tag{43}$$

Algorithm 1: Proposed hybrid Gibbs sampler.

Data: N_{bi}, N_{MC}, $\mathbf{M}^{(0)}$, $\underline{\mathbf{A}}^{(0)}$, $\mathbf{d}\underline{\mathbf{M}}^{(0)}$, $\underline{\mathbf{X}}^{(0)}$, $\boldsymbol{\sigma}^{2(0)}$, $\mathbf{Z}^{(0)}$, $\boldsymbol{\beta}^{(0)}$, $\mathbf{s}^{2(0)}$, $\boldsymbol{\Psi}^{2(0)}$.

begin

  for $q = 1$ **to** N_{MC} **do**

    Sample the endmembers $\mathbf{M}^{(q)}$, cf. §IV-C ;

    Sample the variability terms $\mathbf{d}\underline{\mathbf{M}}^{(q)}$, cf. §IV-D;

    Sample the abundances $\underline{\mathbf{A}}^{(q)}$, cf. §IV-B ;

    Sample the labels and the outliers $(\mathbf{Z}^{(q)}, \underline{\mathbf{X}}^{(q)})$, cf. §IV-E;

    Sample the outlier variances $\mathbf{s}^{2(q)}$, cf. §IV-F;

    Sample the noise variances $\boldsymbol{\sigma}^{2(q)}$, cf. §IV-G;

    Sample the variability variances $\boldsymbol{\Psi}^{2(q)}$, cf. §IV-H;

    Sample the granularity parameters $\boldsymbol{\beta}^{(q)}$, cf. §IV-I;

Result: $\Big\{\mathbf{M}^{(q)}, \mathbf{d}\underline{\mathbf{M}}^{(q)}, \underline{\mathbf{A}}^{(q)}, \mathbf{Z}^{(q)}, \underline{\mathbf{X}}^{(q)}, \boldsymbol{\sigma}^{2(q)}, \mathbf{Z}^{(q)}, \boldsymbol{\beta}^{(q)}, \mathbf{s}^{2(q)}, \boldsymbol{\Psi}^{2(q)}\Big\}_{q=1}^{N_{\text{MC}}}$.

$$\mathbf{M}_t \triangleq \mathbf{M} + \mathbf{d}\mathbf{M}_t \tag{44}$$

$$\begin{aligned}\boldsymbol{\mu}_{n,t}^{(\mathbf{A})} = \boldsymbol{\Lambda}_{n,t}\Bigg[&\frac{1}{\sigma_t^2}\mathbf{M}_t^{\mathrm{T}}(\mathbf{y}_{n,t} - \mathbf{x}_{n,t}) \\ &+ \frac{1}{\varepsilon_n^2}\Big([\mathscr{T}_{n,t}^1 \neq \emptyset]\mathbf{a}_{n,\tau_{n,t}^1} + [\mathscr{T}_{n,t}^2 \neq \emptyset]\mathbf{a}_{n,\tau_{n,t}^2}\Big)\Bigg]\end{aligned} \tag{45}$$

where $\mathcal{N}_{\mathcal{S}_R}(\boldsymbol{\mu}, \boldsymbol{\Lambda})$ denotes a Gaussian distribution truncated to the set $\mathcal{S}_R$ and

$$\mathscr{T}_{n,t}^2 = \{\tau > t \mid z_{n,\tau} = 0\}, \quad \tau_{n,t}^2 = \min_{\tau \in \mathscr{T}_{n,t}^2} \tau \tag{46}$$

with the convention $\mathscr{T}_{n,t}^2 = \emptyset$ if $t = T$.

Samples distributed according to the above truncated multivariate Gaussian distributions can be generated by a Gibbs sampler described in [46, Section IV.B.] [47], by an Hamiltonian Monte-Carlo procedure [48], [49] or by the general method recently proposed in [50]. In this work, the Gibbs sampler [46, Section IV.B.] has been adopted to sample the parameters of interest. Note that the abundance vectors $\mathbf{a}_{n,t}$ can be sampled in parallel to accelerate the algorithm.

C. Sampling the endmembers $\mathbf{M}$

Combining (III-A) and the endmember prior given in Section III-B3 leads to

$$m_{\ell,r} \mid \underline{\mathbf{Y}}, \Theta_{\setminus\{m_{\ell,r}\}} \sim \mathcal{N}_{[b_{\ell,r},+\infty)}(\mu_{\ell,r}^{(\mathbf{M})}, \kappa_{\ell,r}^2) \tag{47}$$

$$b_{\ell,r} = \max\Big\{0, \max_t(-dm_{\ell,r,t})\Big\} \tag{48}$$

$$\begin{aligned}\mu_{\ell,r}^{(\mathbf{M})} = \kappa_{\ell,r}^2 \sum_t \frac{1}{\sigma_t^2}\big[&\widetilde{\mathbf{y}}_{\ell,t} - \widetilde{\mathbf{x}}_{\ell,t} - \widetilde{\mathbf{m}}_{l,\setminus r}\mathbf{A}_{\setminus r,t} \\ &- \widetilde{\mathbf{dm}}_{\ell,t}\mathbf{A}_t\big]\widetilde{\mathbf{a}}_{r,t}^{\mathrm{T}}\end{aligned} \tag{49}$$

$$\kappa^2_{\ell,r} = \left[\sum_{n,t} \frac{a^2_{r,n,t}}{\sigma^2_t} + \frac{1}{\xi}\right]^{-1} \tag{50}$$

where $\widetilde{\mathbf{dm}}_{\ell,t}$ is the ℓth row of $\mathbf{dM}_t$, $\widetilde{\mathbf{m}}_{\ell,\backslash r}$ is the ℓth row of $\mathbf{M}$ whose rth entry has been removed, $\mathbf{A}_{\backslash r,t}$ denotes the matrix $\mathbf{A}_t$ without its rth row and $\widetilde{\mathbf{a}}_{r,t}$ is the rth row of $\mathbf{A}_t$. Samples distributed according to the above truncated Gaussian distributions can be efficiently generated using the algorithm described in [51]. When using a PCA as a preprocessing step (20), the projected endmembers $\mathbf{e}_r$, for $r = 1, \ldots, R$ have a truncated multivariate Gaussian distribution [42]

$$\mathbf{e}_r \mid \underline{\mathbf{Y}}, \Theta_{\backslash\{\mathbf{e}_r\}} \sim \mathcal{N}_{\mathscr{E}_r}(\boldsymbol{\mu}_r^{(\mathbf{E})}, \boldsymbol{\Lambda}_r) \tag{51}$$

where $\mathscr{E}_r$, $\boldsymbol{\mu}_r^{(\mathbf{E})}$ and $\boldsymbol{\Lambda}_r$ have been reported in Appendix A. Note that the rows of $\mathbf{M}$ (resp. of the projected endmember matrix $\mathbf{E}$) can be sampled in parallel to decrease the computational time required by the algorithm.

D. Sampling the variability terms $\underline{\mathbf{dM}}$

Similarly, the likelihood function (III-A) and the prior given in Section III-B4 lead to

$$dm_{\ell,r,t} \sim \mathcal{N}_{[-m_{\ell,r},+\infty)}(\mu^{(\mathbf{dM})}_{\ell,r,t}, \eta^2_{\ell,r,t}) \tag{52}$$

with

$$\begin{aligned}\frac{1}{\eta^2_{\ell,r,t}} = \frac{1}{\sigma^2_t}\sum_n a^2_{r,n,t} + \frac{1}{\nu}\big[t = 1\big] \\ + \frac{1}{\psi^2_{\ell,r}}\Big(1 + \big[1 < t < T\big]\Big)\end{aligned} \tag{53}$$

$$\begin{aligned}\mu^{(\mathbf{dM})}_{\ell,r,t} = \Bigg[\frac{1}{\sigma^2_t}\big(\widetilde{\mathbf{y}}_{\ell,t} - \widetilde{\mathbf{dm}}_{\ell,\backslash r,t}\mathbf{A}_{\backslash r,t} - \widetilde{\mathbf{m}}_\ell \mathbf{a}_{n,t} - \mathbf{x}_{\ell,n,t}\big)\widetilde{\mathbf{a}}^{\mathrm{T}}_{r,t} \\ + \frac{1}{\psi^2_{\ell,r}}\Big([t < T]dm_{\ell,r,(t+1)} + [t > 1]dm_{\ell,r,(t-1)}\Big)\Bigg]\eta^2_{\ell,r,t}\end{aligned} \tag{54}$$

where $\widetilde{\mathbf{dm}}_{\ell,\backslash r,t}$ denotes the ℓth row of $\mathbf{dM}_t$ whose rth element has been removed, $\widetilde{\mathbf{m}}_\ell$ is the ℓth row of $\mathbf{M}$ and $\mathbf{A}_{\backslash r,t}$ is the matrix $\mathbf{A}_t$ without its rth row. The rows of each variability matrix $\mathbf{dM}_t$ can be sampled in parallel to reduce the computational time of the sampler.

E. Sampling the label maps $\mathbf{Z}$ *and the outliers* $\underline{\mathbf{X}}$

According to (III-A) and Section III-B2, the outliers admit the following group-sparsity promoting conditional distributions

$$\begin{aligned}p(\mathbf{x}_{n,t} \mid \mathbf{y}_{n,t}, \Theta_{\backslash\{z_{n,t},\mathbf{x}_{n,t}\}}) = \\ (1 - w_{n,t})\delta(\mathbf{x}_{n,t}) + w_{n,t}\,\mathcal{N}_{\mathbb{R}^L_+}(\boldsymbol{\mu}^{(\mathbf{X})}_{n,t}, \vartheta^2_t \mathbf{I}_L)\end{aligned} \tag{55}$$

which are mixtures of a Dirac mass at $\mathbf{0}$ and of truncated multivariate Gaussian distributions, where

$$w_{n,t} = \frac{\tilde{w}_{n,t}}{\tilde{w}_{n,t} + (1 - \omega_{n,t})}, \quad \vartheta_t^2 = \frac{\sigma_t^2 s_t^2}{\sigma_t^2 + s_t^2} \tag{56}$$

$$\tilde{w}_{n,t} = \frac{\omega_{n,t}}{(s_t^2)^{L/2}} (\vartheta_t^2)^{L/2} \exp\left(\frac{1}{2\vartheta_t^2} \|\boldsymbol{\mu}_{n,t}^{(\mathbf{X})}\|_2^2 \right) \tag{57}$$

$$\boldsymbol{\mu}_{n,t}^{(\mathbf{X})} = \frac{s_t^2}{\sigma_t^2 + s_t^2} \left[\mathbf{y}_{n,t} - (\mathbf{M} + \mathbf{dM}_t)\mathbf{a}_{n,t} \right]. \tag{58}$$

In practice, the labels $z_{n,t}$ are first sampled according to a Bernoulli distribution to select one of the two models for $\mathbf{x}_{n,t}$, with probability $\mathbb{P}[z_{n,t} = 1 \mid \mathbf{y}_{n,t}, \Theta_{\setminus\{z_{n,t}, \mathbf{x}_{n,t}\}}] = w_{n,t}$. Note that the labels $\mathbf{z}_{n,t}$ can be sampled in parallel by using a checkerboard scheme. In addition, the outliers $\mathbf{x}_{n,nt}$ can be sampled in parallel to decrease the computational time.

F. Sampling the outlier variances $\mathbf{s}^2$

According to Sections III-B2 and III-C, we can easily identify the conditional law of s_t^2 for $t = 1, \ldots, T$ as the following inverse gamma distribution

$$\begin{aligned} s_t^2 \mid & \Theta_{\setminus\{s_t^2\}} \sim \\ & \mathcal{IG}\Big(a_{\mathbf{s}} + \frac{\sharp\{n : z_{n,t} = 1\} L}{2}, b_{\mathbf{s}} + \frac{1}{2} \|\mathbf{X}_t\|_{\mathrm{F}}^2 \Big). \end{aligned} \tag{59}$$

G. Sampling the noise variances $\boldsymbol{\sigma}^2$

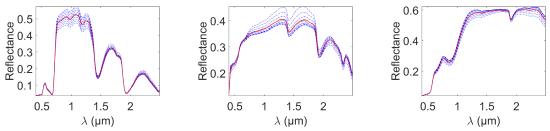
Using Sections III-B5 and III-C, we obtain for $t = 1, \ldots, T$

$$\begin{aligned} \sigma_t^2 \mid \mathbf{Y}_t, \Theta_{\setminus\{\sigma_t^2\}} \sim & \mathcal{IG}\Big(a_\sigma + \frac{LN}{2}, b_\sigma + \\ & \frac{1}{2} \|\mathbf{Y}_t - (\mathbf{M} + \mathbf{dM}_t)\mathbf{A}_t - \mathbf{X}_t\|_{\mathrm{F}}^2 \Big) \end{aligned}. \tag{60}$$

H. Sampling the variability variances $\boldsymbol{\Psi}^2$

Similarly, Section III-B4 and III-C lead to

$$\begin{aligned} \psi_{\ell,r}^2 \mid \Theta_{\setminus\{\psi_{\ell,r}^2\}} \sim & \mathcal{IG}\Big(a_\Psi + \frac{T-1}{2}, b_\Psi + \\ & \frac{1}{2} \sum_{t=2}^{T} (dm_{\ell,r,t} - dm_{\ell,r,t-1})^2 \Big). \end{aligned} \tag{61}$$



(a) Endmember 1 (b) Endmember 2 (c) Endmember 3

Fig. 2. Endmembers ($\mathbf{m}_r$, red lines) and their variants affected by variability ($\mathbf{m}_r + \mathbf{dm}_{r,t}$, blue dotted lines) used to generate the synthetic mixtures with $R = 3$. Signatures corresponding to different time instants are represented in a single figure to better appreciate the variability introduced in the data.

I. Sampling the granularity parameters β_t

Provided square images are considered, the partition functions $C(\beta_t)$ have the closed-form expressions [41], [52]

$$\widetilde{C}(\beta_t) = \frac{1}{2}\log(2\sinh\beta_t) + \frac{1}{2N}\sum_{n=1}^{N}\operatorname{acosh}\Delta_n(\beta_t) + \beta_t$$

$$\Delta_n(\beta_t) = v(\beta_t) - C_n,\ v(\beta_t) = \frac{\cosh^2\beta_t}{\sinh\beta_t}$$

$$\widetilde{C}(\beta_t) = \frac{1}{N}\log C(\beta_t), \quad C_n = \cos\left(\frac{2n-1}{2N}\pi\right).$$

The exact partition function can then be used to sample the parameters β_t using Metropolis-Hastings steps. In this work, new values of the granularity parameters have been proposed by the following Gaussian random walk

$$\beta_t^* = \beta_t^{(q)} + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}\big(0, \sigma_\beta^2(t)\big) \tag{62}$$

where the parameters $\sigma_\beta^2(t)$ are adjusted during the burn-in iterations to yield acceptance rates in the interval $[0.4, 0.6]$.

J. Computational complexity

Assuming elementary arithmetic operations and scalar pseudo-random number generations are $O(1)$ operations, the overall computational complexity is dominated by matrix products needed to compute the parameters related to the conditional distribution of the variability vectors. Since $R \ll L \ll N$ and $T \ll L$, the per-iteration computational cost of the proposed algorithm is $O(LR^2NT)$ per iteration. As detailed in the preceding paragraphs, many parameters can be sampled in parallel to reduce the computational time of the proposed algorithm. In comparison, the computational complexity of VCA is $O(R^2N)$ [23] per image, and the per iteration complexity of the others algorithms for a single image are respectively: $O(N^2)$ for FCLS [24], $O(RN)$ for SISAL [25], $O(LRN)$ for rLMM [26] and $O(R^2(L+N))$ for OU [15].

V. Experiments with synthetic data

The proposed method has been applied to an MTHS image composed of 10 acquisitions of size 50×50 with $L = 413$ bands. The first scenario deals with the appearance of a new material in specific regions of a few images. To this end, 4 images out of the 10 have been corrupted by spatially sparse outliers, corresponding to a new endmember extracted from a spectral library. Each image of the sequence corresponds to a linear mixture of 3 endmembers affected by smooth time-varying variability, and the synthetic abundances vary smoothly from one image to another.

First, so-called reference abundance maps corresponding to the first time instant have been generated (e.g., for Fig. 4, we have taken the abundance maps obtained by VCA/FCLS on the widely studied Moffett dataset [42]). Then, the abundance maps corresponding to the remaining time instants have been generated by multiplying the reference maps with trigonometric functions to ensure a sufficiently smooth temporal evolution. For the first dataset composed of $R = 3$ endmembers, the reference maps associated with the first two endmembers have been respectively multiplied by $\cos\left(\frac{\pi}{100} + t\frac{48\pi}{100}\right)$ and $\sin\left(\frac{\pi}{100} + t\frac{48\pi}{100}\right)$, with $t \in \{1, \ldots, T\}$. The temporal evolution of the last abundance map has finally been obtained by leveraging the sum-to-one condition. With these abundance maps, the contribution of a given endmember (assumed to punctually disappear) has been replaced at specific time instants by a new endmember signature in pixels originally corresponding to its highest abundance coefficients (e.g., above 0.8).

The mixtures have finally been corrupted by an additive white Gaussian noise to ensure a resulting signal-to-noise ratio (SNR) between 25 and 30 dB. Similarly, two complementary scenarii involving 5 HS images, of size 100×100, composed of 6 and 9 endmembers, have been considered to analyze the performance of the method in the presence of a larger number of endmembers. Note that a larger image size has been considered for these two datasets to reflect the fact that a larger number of endmembers is expected to be observed in larger scenes. In practice, the two other datasets (i.e., composed of 6 and 9 endmembers) have been generated from the 100×100 abundance maps used in [53]. In addition, the images of this experiment do not satisfy the pure pixel assumption to assess the proposed method in challenging situations.

Controlled spectral variability has been introduced by using the product of reference endmembers with randomly generated piecewise-affine functions as in [10], where different affine functions have been generated for each endmember at each time instant. Typical instances of the signatures used in this experiment are depicted in Fig. 2. The robustness of the proposed method to moderate spatial variability, i.e., endmember variability occurring within single images, has also been evaluated. The corresponding results can be found in Appendix E.

TABLE I

FIXED PARAMETERS, AND INITIAL VALUES ASSOCIATED IN THE EXPERIMENTS TO PARAMETERS LATER INFERRED FROM THE MODEL.

		Synthetic data	Real data
Fixed parameters	ε_n^2	10^{-3}	10^{-2}
	ξ	1	1
	ν	10^{-3}	10^{-5}
	$a_{\mathbf{s}}, a_\Psi, a_\sigma$	10^{-3}	10^{-3}
	$b_{\mathbf{s}}, b_\Psi, b_\sigma$	10^{-3}	10^{-3}
	N_{bi}	350	450
	N_{MC}	400	500
Initial values	σ_t^2	10^{-4}	10^{-4}
	s_t^2	5×10^{-3}	5×10^{-3}
	$\psi_{\ell,r}^2$	10^{-3}	10^{-2}
	β_t	1.7	1.7

A. *Compared methods*

The results of the proposed algorithm have been compared to those of several unmixing methods from the literature, some of which are specifically designed to unmix a single HS image. In the following lines, the most relevant implementation details specific to each method are briefly recalled.

1) VCA/FCLS (no variability, single image): the endmembers are first extracted on each image using the vertex component analysis (VCA) [23], which requires pure pixels to be present. The abundances are then estimated for each pixel by solving a fully constrained least squares problem (FCLS) using the alternating direction method of multipliers (ADMM) [24]. Note that the estimates provided by the VCA algorithm vary from one run to another, given its stochastic nature;
2) SISAL/FCLS (no variability, single image): the endmembers are extracted on each image by the simplex identification via split augmented Lagrangian (SISAL) [25], and the abundances are estimated for each pixel by FCLS. The tolerance for the stopping rule has been set to 10^{-3};
3) RLMM (no variability, single image): the unmixing method associated with the robust linear mixing model (RLMM) proposed in [26] has been applied to each image of the series independently. The algorithm has been initialized with SISAL/FCLS, and the regularization parameter specific to this method is set as in [26];
4) OU: the endmembers are estimated using the online unmixing (OU) algorithm introduced in [15] with endmembers initialized by the output of VCA applied to the first image of the sequence.

The abundances are initialized by FCLS, and the variability matrices are initialized with all their entries equal to 0. The other parameters are set to the same values as those given in [15, Table I];

5) Proposed approach: the endmembers are initialized with VCA applied to the first image of the sequence, within which the observed materials are well represented (i.e., with sufficiently high abundance coefficients for each material). In this context the VCA algorithm, which requires pure pixels to be present in the data, has been observed to yield relevant results for the initialization. However, other endmember extraction techniques might be used to initialize the proposed algorithm if needed. The abundances are initialized by FCLS, and the variability matrices and label maps are initialized with all their entries equal to 0 (i.e., the images are a priori assumed to contain no outlier). The values chosen for the other parameters are summarized in Table I. Further details on these values can be found in Appendix B.

Performance assessment has been conducted in terms of

(i) endmember estimation through the average spectral angle mapper (aSAM)

$$\mathrm{aSAM}(\mathbf{M}) = \frac{1}{R}\sum_{r=1}^{R} \arccos\left(\frac{\mathbf{m}_r^{\mathrm{T}}\widehat{\mathbf{m}}_r}{\|\mathbf{m}_r\|_2\|\widehat{\mathbf{m}}_r\|_2}\right); \tag{63}$$

(ii) abundance and variability estimation through the global mean square errors (GMSEs)

$$\mathrm{GMSE}(\mathbf{A}) = \frac{1}{TRN}\sum_{t=1}^{T}\|\mathbf{A}_t - \widehat{\mathbf{A}}_t\|_{\mathrm{F}}^2 \tag{64}$$

$$\mathrm{GMSE}(\mathbf{dM}) = \frac{1}{TLR}\sum_{t=1}^{T}\|\mathbf{dM}_t - \widehat{\mathbf{dM}}_t\|_{\mathrm{F}}^2; \tag{65}$$

(iii) quadratic reconstruction error (RE)

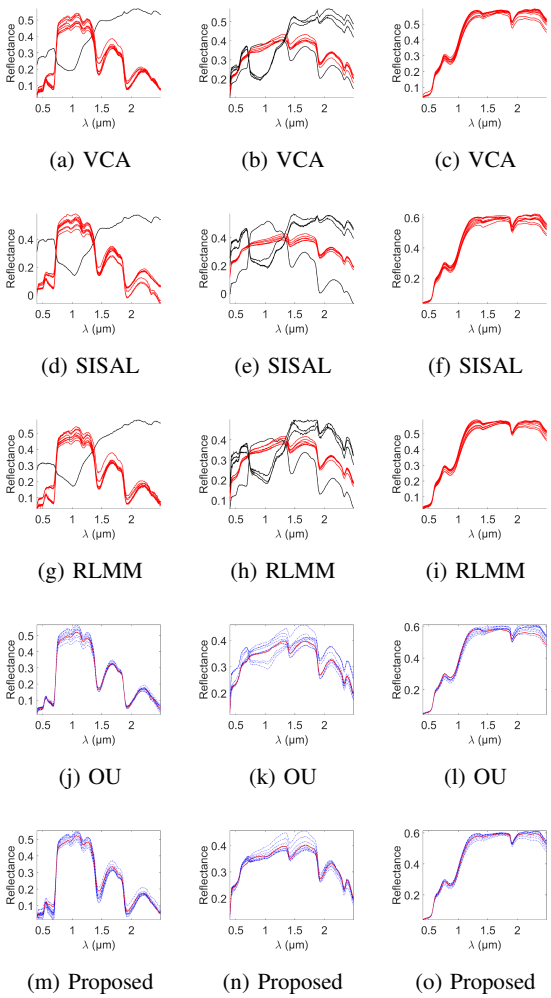
$$\mathrm{RE} = \frac{1}{TLN}\sum_{t=1}^{T}\|\mathbf{Y}_t - \widehat{\mathbf{Y}}_t\|_{\mathrm{F}}^2 \tag{66}$$

where $\widehat{\mathbf{Y}}_t$ is the matrix composed of the pixels reconstructed with the estimated parameters.

B. Results

The endmembers estimated by the proposed algorithm are compared to those of VCA/FCLS, SISAL/FCLS, RLMM and OU in Fig. 3, whereas the corresponding abundance maps are displayed in Fig. 4. Note that the abundance maps and the endmembers obtained for the mixtures of 6 and 9 endmembers are reported in Appendix D to ease the reading of this paper. The unmixing performance of each method, reported in Table II, leads to the following conclusions.

- **Endmember estimation:** the proposed method shows an interesting robustness with respect to spatially sparse outliers in the sense that the estimated signatures (Figs. 3m to 3o) are very close to the corresponding ground truth (Fig. 2). In comparison, the shape of the endmembers recovered by VCA, SISAL and RLMM and the variability extracted by OU are significantly affected by



(a) VCA (b) VCA (c) VCA

(d) SISAL (e) SISAL (f) SISAL

(g) RLMM (h) RLMM (i) RLMM

(j) OU (k) OU (l) OU

(m) Proposed (n) Proposed (o) Proposed

Fig. 3. Endmembers ($\mathbf{m}_r$, red lines) and their variants affected by variability ($\mathbf{m}_r + \mathbf{dm}_{r,t}$, blue dotted lines) recovered by the different methods from the synthetic mixtures with $R = 3$. Each row of the figure corresponds to the endmembers extracted for the 3 materials by each algorithm. Signatures corresponding to different time instants are represented on a single figure to better appreciate the variability recovered from the data. The spectra represented in black in Figs. 3a, 3b, 3d, 3e and 3h correspond to signatures corrupted by outliers.

outliers, as exemplified in Figs. 3b, 3e, 3h and 3k respectively. These qualitative results are confirmed by the quantitative performance measures of each method provided in Table II. Note that the endmembers recovered by the SISAL and RLMM methods are very sensitive to the VCA initialization, as illustrated by the similarity between the signatures estimated by these methods (Figs. 3a to 3i).

- **Abundance estimation:** the abundance maps estimated by FCLS, RLMM and SISAL reflect the high sensitivity of VCA (used to initialize SISAL and RLMM) to the presence of outliers (see the figures delineated in red in Fig. 4). On the contrary, the abundances recovered by OU and the proposed approach are much closer to the ground truth. These observations are confirmed by the abundance estimation performance reported in Table II. The proposed abundance smoothness prior appears to mitigate the errors induced by the presence of outliers as can be seen in Figs. 13

Fig. 4. Abundance map of the first endmember recovered by the different methods (in each row) at each time instant (given in column) for the experiment with $R = 3$ [the different rows correspond to the true abundances, VCA/FCLS, SISAL/FCLS, RLMM, OU and the proposed method]. The images delineated in red show that several methods are highly sensitive to the presence of outliers, and the time instants represented with $*$ denote images containing outliers.

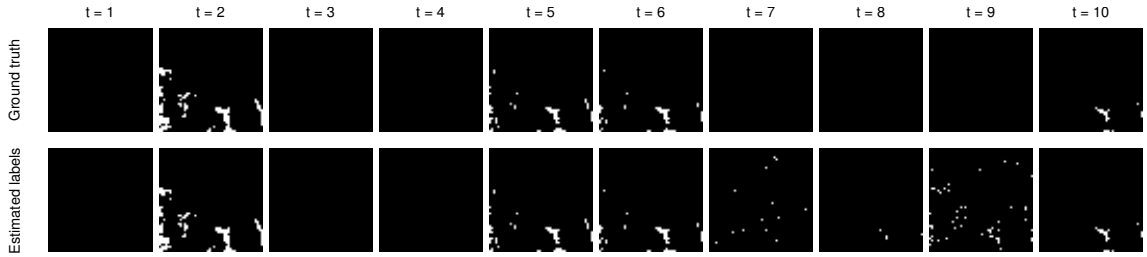


Fig. 5. Outlier labels $\mathbf{z}_t$ estimated for each image of the synthetic dataset with 3 endmembers (the different rows correspond to the true labels, and the estimated labels) [0 in black, 1 in white].

and 14 in Appendix C. A more detailed version of Table II (Table V), along with complementary figures illustrating the interest of the proposed abundance prior, can be found in Appendix C.

- **Overall performance:** the performance measures reported in Table II are globally favorable to the proposed approach. It is important to mention that the price to pay with the good performance of the proposed method is its computational complexity, which is common with MCMC methods.

As a complementary output, the proposed algorithm is able to recover the location of the outliers within each image, as illustrated in Fig. 5. Up to a few false detections, the estimated labels are very close to the ground truth. The label errors observed for $t = 7, 8$ and 9 partly result from the different abundance constraints considered when an outlier is detected or not.

TABLE II

SIMULATION RESULTS ON SYNTHETIC DATA (aSAM($\mathbf{M}$) IN (°), GMSE($\mathbf{A}$)$\times 10^{-2}$, GMSE(d$\mathbf{M}$)$\times 10^{-4}$, RE $\times 10^{-4}$, TIME IN (S)).

		aSAM(M)	GMSE(A)	GMSE(dM)	RE	time
$R=3$	VCA/FCLS	6.07	2.32	/	3.91	**1**
	SISAL/FCLS	5.07	1.71	/	2.28	2
	RLMM	5.13	2.04	/	**0.31**	463
	OU	**1.90**	0.42	3.22	2.61	98
	Proposed	2.03	**0.15**	**1.85**	2.00	2530
$R=6$	VCA/FCLS	3.81	1.57	/	3.09	**2**
	SISAL/FCLS	5.76	0.91	/	4.49	3
	RLMM	2.73	1.26	/	**0.29**	1453
	OU	2.74	0.38	3.70	1.13	420
	Proposed	**1.48**	**0.16**	**2.84**	0.51	8691
$R=9$	VCA/FCLS	3.74	0.65	/	6.83	**4**
	SISAL/FCLS	5.91	0.36	/	5.56	5
	RLMM	2.48	0.54	/	**0.31**	1447
	OU	6.08	0.47	**2.19**	0.89	1024
	Proposed	**2.23**	**0.15**	8.38	0.82	17151

VI. EXPERIMENT WITH REAL DATA

A. *Description of the dataset*

We consider a real sequence of AVIRIS HS images acquired over the Lake Tahoe region (California, United States of America) between 2014 and 2015[1]. The scene of interest (100×100), composed of a lake and a nearby field, has been unmixed with $R = 3$ endmembers based on the results of the noise-whitened eigengap algorithm (NWEGA) [54] applied to each image of the series (see Table III). This choice is further supported by results obtained from a previous analysis conducted on the same dataset [55, Appendix E]. For $R = 4$ and 5, the signatures of water, soil and vegetation were split into two or more components by the different algorithms, suggesting $R = 3$ is more appropriate for this study. Note that prior studies led in [15] revealed that this dataset contains outliers (area delineated in red in Fig. 6e). After removing the seemingly corrupted bands and the water absorption bands, 173 out of the 224 spectral bands were finally exploited. The initial parameters used for the proposed algorithm are given in Table I. The other methods have been run with the same parameters as in Section V. Note that the VCA results reported in this section are representative of those obtained over multiple runs (no significant differences have been observed from one run to another).

[1]The images from which the scene under study is extracted are freely available from the online AVIRIS flight locator tool at http://aviris.jpl.nasa.gov/alt_locator/.

(a) 04/10/2014 (b) 06/02/2014 (c) 09/19/2014 (d) 11/17/2014 (e) 04/29/2015 (f) 10/13/2015

Fig. 6. Scenes used in the experiment, given with their respective acquisition date. The area delineated in red in Fig. 6e highlights a region known to contain outliers (this observation results from a previous analysis led on this dataset in [15]).

B. Results

In the absence of any ground truth, the performance of the unmixing methods is assessed in terms of RE (Table IV) while taking into account the consistency of the estimated abundance maps reported in Figs. 8, 9 and 10. More precisely, the abundances associated with the vegetation area are expected to be very high for $t = 1, 3, 5$ (corresponding to Figs. 6a, 6c and 6e) where the vegetation visually appears to be sufficiently irrigated (hence well represented). On the contrary, the abundance coefficients are supposed to be much lower for $t = 2, 4, 6$ (corresponding to Figs. 6b, 6d and 6f), where the vegetation is visually drier or almost absent. Concerning the presence of water in the bottom left-hand corner of the images, the latent variables introduced in Section III-B2 are expected to reflect the abrupt variations in the presence of water observed at $t = 3, 4$ and 5. These observations, combined with the extracted signatures (Fig. 7) and the estimated abundances (Figs. 8 to 10) lead to the following comments.

- **Endmember estimation:** the signature recovered for the soil by VCA, SISAL and RLMM at time $t = 5$ shows an amplitude which is significantly greater than the amplitude of the signatures extracted at the other time instants, and a shape incompatible with what can be expected based on physical considerations (see the black signatures in Figs. 7a, 7d and 7g). This is a clear indication that outliers are present in the corresponding image. A similar observation can be made for the vegetation signature obtained by VCA, SISAL and RLMM at time $t = 5$. On the contrary, the endmembers recovered by OU and the proposed approach are much more consistent from this point of view.
- **Abundance estimation:** the estimated abundances globally reflect the previous comments made on the extracted endmembers. Notably, the abundance coefficients estimated at $t = 5$ by VCA, SISAL and RLMM (delineated in red in Figs. 8 to 10) are visually inconsistent with the temporal evolution of the materials observed in the true color composition given in Fig. 6. More explicitly, the soil is not supposed to be concentrated on a few pixels as suggested by the corresponding abundance maps in Fig. 8. Similarly, the water is not supposed to be present in high proportions in all the pixels of the image as indicated in Fig. 9. These results, in contradiction with Fig. 6,

TABLE III

ENDMEMBER NUMBER R ESTIMATED BY NWEGA [54] ON EACH IMAGE OF THE REAL DATASET.

	04/10/2014	06/02/2014	09/19/201	11/17/2014	04/29/2015	10/13/2015
NWEGA	3	3	3	4	3	4

TABLE IV

SIMULATION RESULTS ON REAL DATA (RE $\times 10^{-4}$).

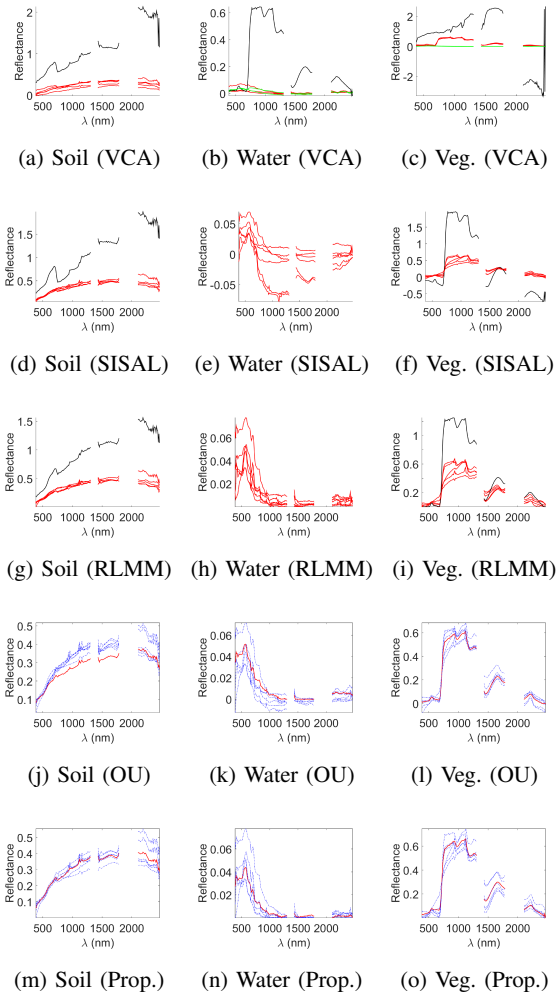
		RE	time (s)
$R=3$	VCA/FCLS	45.05	**1**
	SISAL/FCLS	1.65	2
	rLMM	2.51	390
	OU	2.50	508
	Proposed	**0.34**	23608

suggest that outliers are present at $t = 5$. In addition, the abundance maps estimated at $t = 4$ and 6 by FCLS for the water and the vegetation (delineated in green in Figs. 8 and 9) suggest that the water contribution has been split into two spectra. The corresponding signatures are represented in green in Figs. 7a and 7c. On the contrary, the results reported for OU and the proposed method are consistent with the expected evolution of water and vegetation over time (abundance values close to 1 at time $t = 1$, 3, 5, lower values at time $t = 2$, 4, 6). Finally, the vegetation abundance maps estimated by the proposed method globally presents a better contrast than those obtained with OU (Fig. 10).

The previous comments, along with the lower reconstruction error reported in Table IV, suggest that the proposed approach is robust to spatially sparse outliers while allowing smooth temporal variations to be exploited. Indeed, the pixels corresponding to abrupt variations of the water signature have been properly detected. Furthermore, the outliers previously detected in this dataset [15] for $t = 5$ (highlighted in red in Fig. 6e) are well captured by the latent variables $\mathbf{Z}$ (see Fig. 11). In addition, the spatial distribution of the estimated outlier labels (Fig. 11) is in agreement with the results of the RLMM (in terms of the spatial distribution of the outlier energy) and with the non-linearity detector [56] applied to each image of the sequence with the SISAL-estimated endmembers (see Fig. 12). Concentrated on regions where non-linear effects can be reasonably expected, the active latent variables $\mathbf{Z}$ tend to capture the spatial distribution of the non-linearities possibly occurring in the observed scene.

VII. CONCLUSION AND FUTURE WORK

This paper introduced a Bayesian model accounting for both smooth and abrupt variations possibly occurring in multitemporal hyperspectral images. The adopted model was specifically designed to



(a) Soil (VCA) (b) Water (VCA) (c) Veg. (VCA)

(d) Soil (SISAL) (e) Water (SISAL) (f) Veg. (SISAL)

(g) Soil (RLMM) (h) Water (RLMM) (i) Veg. (RLMM)

(j) Soil (OU) (k) Water (OU) (l) Veg. (OU)

(m) Soil (Prop.) (n) Water (Prop.) (o) Veg. (Prop.)

Fig. 7. Endmembers ($\mathbf{m}_r$, red lines) and their variants affected by variability ($\mathbf{m}_r + \mathbf{dm}_{r,t}$, blue dotted lines) recovered by the different methods from the real dataset depicted in Fig. 6. The spectral gaps in the recovered signatures correspond to the low SNR bands which have been removed prior to the unmixing procedure. Signatures corresponding to different time instants are represented in a single figure to better appreciate the variability recovered from the data. The spectra represented in black correspond to signatures corrupted by outliers, while those given in green represent endmembers which have been split into several components by the associated estimation procedure.

handle datasets in which mostly the same materials were expected to be observed at different time instants, thus allowing information redundancy to be exploited. An MCMC algorithm was derived to solve the resulting unmixing problem in order to precisely assess the performance of the proposed approach on multitemporal HS images of moderate size (i.e., moderate spatial and temporal dimensions). This algorithm was used to sample the posterior of the proposed hierarchical Bayesian model and to use the generated samples to build estimators of the unknown model parameters. Given its computational cost, the proposed approach is not intended to be applied to large datasets, for which different unmixing methods can provide a rougher analysis at a smaller computational cost. The proposed approach is rather meant to be used as a complementary tool to carry out an in-depth analysis of scenes of moderate size. Future research perspectives include the use of

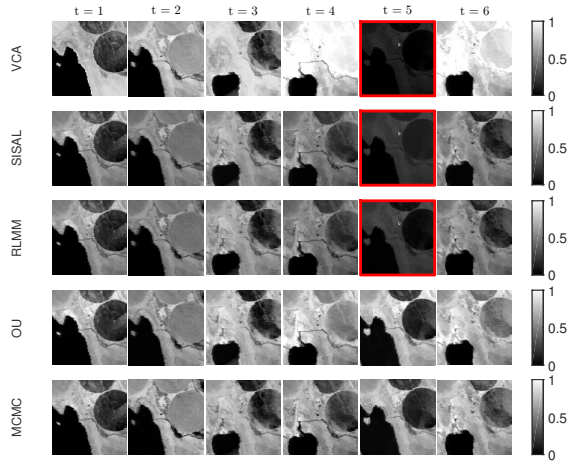


Fig. 8. Soil abundance map recovered by the different methods (in each row) at each time instant (given in column) for the experiment on the real dataset [the different rows correspond to VCA/FCLS, SISAL/FCLS, RLMM, OU, and the proposed method]. The images delineated in red suggest that some of the methods are particularly sensitive to the presence of outliers.

relaxation methods to the Ising field to tackle similar problems with online optimization techniques, and the development of distributed unmixing procedures to efficiently unmix larger datasets. Designing unmixing methods scaling with the problem dimension while simultaneously accounting for temporal and spatial endmember variability is another interesting prospect.

APPENDIX A
SAMPLING THE PROJECTED ENDMEMBERS $\mathbf{E}$

When using a PCA as a preprocessing step, the projected endmembers $\mathbf{e}_r$, for $r = 1, \ldots, R$, are distributed according to the following truncated Gaussian distributions

$$\mathbf{e}_r \mid \underline{\mathbf{Y}}, \Theta_{\setminus\{\mathbf{e}_r\}} \sim \mathcal{N}_{\mathscr{E}_r}(\boldsymbol{\mu}_r^{(\mathbf{E})}, \boldsymbol{\Lambda}_r) \tag{67}$$

with $\mathscr{E}_r = [c_{1,r}, d_{1,r}] \times \ldots \times [c_{K,r}, d_{K,r}]$, and for $k = 1, \ldots, K$

$$\begin{aligned}
c_{k,r} &= \max_{\ell \in \mathcal{U}_k^+} \left(-\frac{\check{y}_\ell + \sum_{j \neq k} u_{\ell,j} e_{j,r} + b_{\ell,r}}{u_{\ell,k}} \right) \\
d_{k,r} &= \min_{\ell \in \mathcal{U}_k^-} \left(-\frac{\check{y}_\ell + \sum_{j \neq k} u_{\ell,j} e_{j,r} + b_{\ell,r}}{u_{\ell,k}} \right) \\
b_{\ell,r} &= \min\Big\{0, \min_t (dm_{\ell,r,t})\Big\} \\
\boldsymbol{\Lambda}_r^{-1} &= \left[\frac{1}{\xi} + \sum_{n,t} \frac{a_{r,n,t}^2}{\sigma_t^2} \right] \mathbf{I}_{R-1}
\end{aligned} \tag{68}$$

$$\begin{aligned}
\boldsymbol{\mu}_r^{(\mathbf{E})} = \boldsymbol{\Lambda}_r \mathbf{U}^{\mathrm{T}} \Bigg[\sum_{t,n} \frac{1}{\sigma_t^2} \Big(\mathbf{y}_{n,t} - \mathbf{x}_{n,t} - \mathbf{d}\mathbf{M}_t \mathbf{a}_{n,t} \\
- \check{\mathbf{y}} a_{r,n,t} - \sum_{j \neq r} a_{j,n,t} \mathbf{m}_j \Big) a_{r,n,t} \Bigg].
\end{aligned}$$

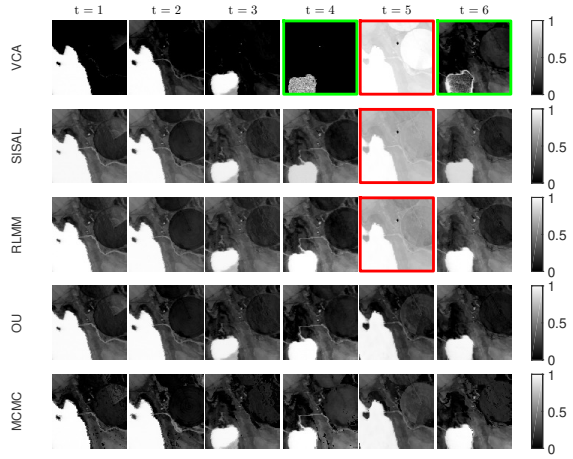


Fig. 9. Water abundance map recovered by the different methods (in each row) at each time instant (given in column) for the experiment on the real dataset [the different rows correspond to VCA/FCLS, SISAL/FCLS, RLMM, OU, and the proposed method]. On the one hand, the images delineated in red suggest that some of the methods are particularly sensitive to the presence of outliers. On the other hand, the images delineated in green represent the abundance maps associated with signatures which have been split into two components by the corresponding unmixing procedures.

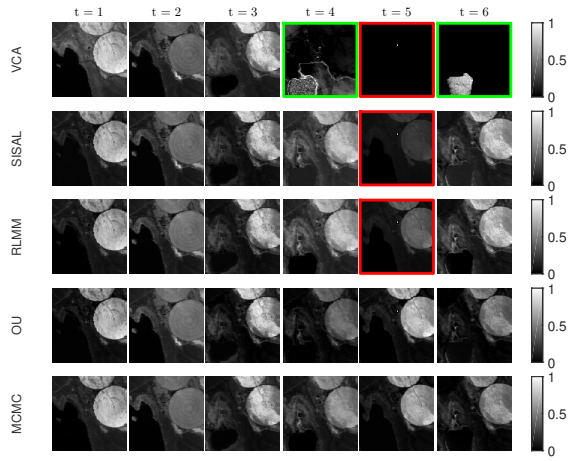


Fig. 10. Vegetation abundance map recovered by the different methods (in each row) at each time instant (given in column) for the experiment on the real dataset [the different rows correspond to VCA/FCLS, SISAL/FCLS, RLMM, OU, and the proposed method]. The images delineated in red suggest that some of the methods are particularly sensitive to the presence of outliers.

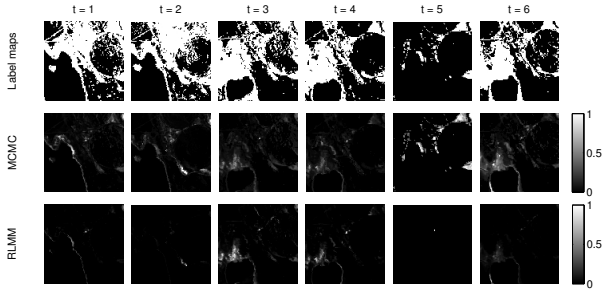


Fig. 11. mMAP estimates of the label maps recovered by the proposed approach, displayed at each time instant (the different rows correspond to: the estimated label map (pixels detected as outliers appear in white), the outlier energy map re-scaled in the interval $[0, 1]$ obtained by the proposed method, and by RLMM).

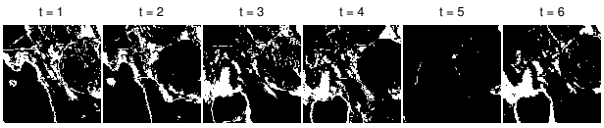


Fig. 12. Non-linearity maps estimated by the detector [56] applied to each image with the SISAL-extracted endmembers, with a probability of false alarm of 10^{-3} (pixels detected as non-linearities appear in white).

APPENDIX B
DETAILS ON THE VALUES CHOSEN IN TABLE I

The initial values chosen for the parameters in Table I are based on the following considerations.

(a) The initial noise variance σ_t^2 reflects a typical order of magnitude of the SNR (HS images are typically known to have an SNR around $30\,\mathrm{dB}$).

(b) The initial outlier variance s_t^2 has been taken an order of magnitude greater than σ_t^2 to ensure outlier contributions that can be captured by the algorithm.

(c) The initial variability variance $\psi_{\ell,r}^2$, which controls the temporal smoothness of the variability term, can be *a priori* chosen of the same order of magnitude as s_t^2.

(d) The granularity parameters β_t were initially selected to reflect the practitioner's prior knowledge on the smoothness of the outlier spatial support. A value between 1 and 2 (i.e., above the phase-transition temperature of the Ising MRF [57]) is particularly appropriate for natural scenes, in which the observed materials exhibit a relatively smooth spatial distribution.

Similarly, the values of the fixed parameters given in Table I are selected as follows.

(a) ε_n^2, which controls the confidence given to the abundance smoothness prior, has been fixed by cross-validation (i.e., based on the estimation results obtained by multiple runs for different values);

(b) Since no specific prior knowledge is available on the endmembers, the endmember variance ξ is chosen sufficiently large (typically equal to 1) to ensure the endmember prior is weakly informative;

(c) ν, which controls the energy of the variability captured by the algorithm, has been set by cross validation;

(d) Given the absence of specific prior knowledge on the outlier variances s_t^2, the variability variances $\psi_{\ell,r}^2$ and the noise variances σ_t^2, the hyperparameters $a_{\mathbf{s}}, a_{\Psi}, a_{\sigma}, b_{\mathbf{s}}, b_{\Psi}, b_{\sigma}$ are set to a small value (typically 10^{-3}) to ensure the chosen conjugate inverse-gamma priors are uninformative [21];

(e) The number of Monte-Carlo and burn-in iterations, respectively denoted by N_{MC} and N_{bi}, are set according to a classical convergence diagnosis, namely the potential scale reduction factor (PSRF) [58].

APPENDIX C
COMPLEMENTARY ILLUSTRATION OF THE PROPOSED ABUNDANCE PRIOR

In this section, two complementary figures (Figs. 13 and 14) illustrate the interest of considering the proposed abundance prior whenever outliers are observed in the MTHS image for synthetic dataset composed of $R = 3$ endmembers. These illustrations are further substantiated by the estimation metrics detailed for each time instant in Table V.

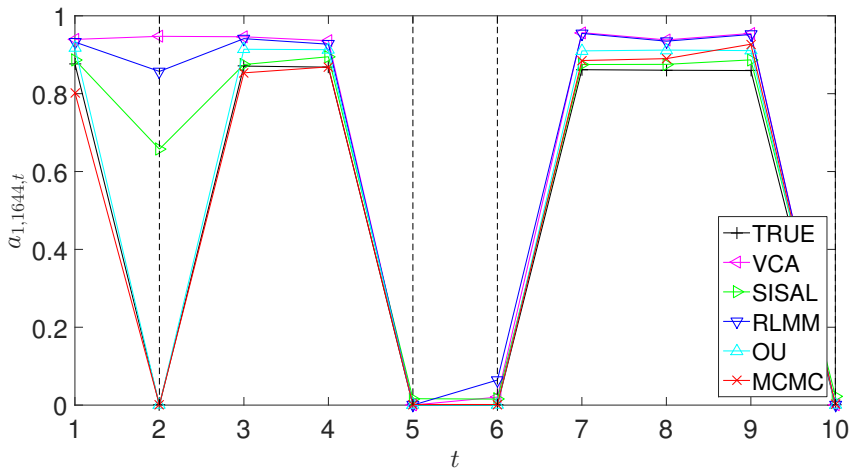


Fig. 13. Evolution of the abundance corresponding to a pixel in which outliers are observed at several time instants (indicated by vertical dotted lines). The abundance coefficient estimated at each time instant by the different unmixing strategies are reported in different colors. The proposed method proves to be particularly appropriate to mitigate the errors induced by the presence of outliers.

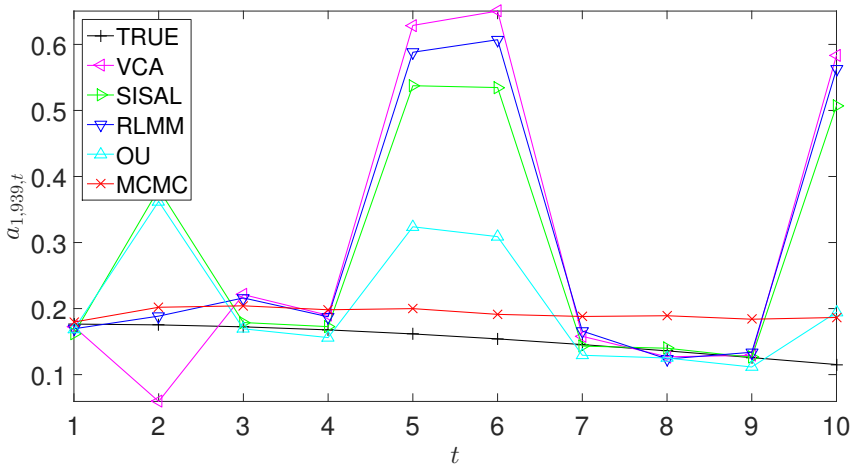


Fig. 14. Evolution of the abundance corresponding to a pixel for which no outliers are observed over time. The abundance coefficient estimated at each time instant by the different unmixing strategies are reported in different colors.

TABLE V

DETAILED TABLE OF RESULTS OBTAINED ON THE SYNTHETIC DATA COMPOSED OF $R = 3$ ENDMEMBERS. THE CRITERION $\text{aSAM}(\mathbf{M}_t)$ DENOTES THE AVERAGE SPECTRAL ANGLE MAPPER BETWEEN THE GROUND TRUTH PERTURBED ENDMEMBERS $\mathbf{M}_t = \mathbf{M} + \mathbf{dM}_t$ AND THE ESTIMATED PERTURBED ENDMEMBERS AT TIME INSTANT t. THE TIME INSTANTS REPRESENTED WITH $*$ DENOTE IMAGES CONTAINING OUTLIERS.

	t	1	2$*$	3	4	5$*$	6$*$	7	8	9	10$*$
	VCA	1.835e-03	3.764e-02	2.379e-03	1.644e-03	5.942e-02	6.598e-02	1.823e-03	1.963e-03	1.228e-03	5.849e-02
	SISAL	1.088e-04	3.631e-02	4.532e-05	3.667e-04	3.973e-02	4.002e-02	1.304e-04	1.136e-04	1.708e-04	4.242e-02
GMSE($\mathbf{A}_t$)	RLMM	1.689e-03	3.788e-02	2.092e-03	1.526e-03	4.924e-02	5.442e-02	1.716e-03	1.859e-03	1.188e-03	5.242e-02
	OU	9.990e-04	1.438e-02	9.873e-04	9.541e-04	9.571e-03	8.797e-03	8.247e-04	8.437e-04	9.699e-04	3.633e-03
	MCMC	4.826e-04	2.715e-03	4.899e-04	5.583e-04	1.880e-03	1.780e-03	9.692e-04	1.201e-03	1.716e-03	3.104e-03
	VCA	2.095e+00	1.840e+01	3.085e+00	2.560e+00	1.055e+01	1.100e+01	2.622e+00	2.706e+00	2.252e+00	1.035e+01
	SISAL	4.842e-01	2.672e+01	2.669e-01	7.102e-01	1.055e+01	1.004e+01	5.483e-01	5.191e-01	5.956e-01	9.231e+00
aSAM($\mathbf{M}_t$)	RLMM	1.502e+00	2.134e+01	1.647e+00	1.401e+00	8.894e+00	9.427e+00	1.782e+00	1.665e+00	1.520e+00	8.874e+00
	OU	1.231e+00	4.402e+00	1.317e+00	1.346e+00	3.197e+00	2.984e+00	1.312e+00	1.292e+00	1.353e+00	1.785e+00
	MCMC	1.522e+00	1.461e+00	8.965e-01	6.649e-01	7.913e-01	8.633e-01	1.055e+00	1.226e+00	1.679e+00	2.797e+00
	VCA	1.308e-04	5.450e-04	2.119e-04	1.942e-04	6.402e-04	5.543e-04	3.400e-04	2.817e-04	3.416e-04	6.675e-04
	SISAL	1.252e-04	2.438e-04	1.621e-04	1.635e-04	2.714e-04	1.824e-04	3.323e-04	2.110e-04	3.343e-04	2.500e-04
RE	RLMM	1.844e-04	7.596e-04	2.593e-04	2.608e-04	6.193e-04	4.322e-04	6.139e-04	3.579e-04	6.193e-04	4.894e-04
	OU	1.254e-04	4.108e-04	1.624e-04	1.637e-04	3.367e-04	2.458e-04	3.325e-04	2.113e-04	3.346e-04	2.843e-04
	MCMC	1.252e-04	1.238e-04	1.621e-04	1.635e-04	2.060e-04	1.307e-04	3.314e-04	2.109e-04	3.319e-04	2.102e-04
	VCA	1.904e+00	3.025e+00	2.387e+00	2.284e+00	3.844e+00	3.480e+00	2.954e+00	2.605e+00	2.930e+00	3.523e+00
	SISAL	1.872e+00	2.351e+00	2.099e+00	2.093e+00	2.557e+00	2.059e+00	2.929e+00	2.321e+00	2.904e+00	2.392e+00
aSAM($\mathbf{Y}_t$)	RLMM	1.964e+00	3.422e+00	2.231e+00	2.220e+00	3.040e+00	2.469e+00	3.170e+00	2.480e+00	3.150e+00	2.650e+00
	OU	1.873e+00	2.527e+00	2.100e+00	2.095e+00	2.603e+00	2.084e+00	2.930e+00	2.322e+00	2.905e+00	2.414e+00
	MCMC	1.872e+00	1.785e+00	2.099e+00	2.093e+00	2.303e+00	1.828e+00	2.925e+00	2.320e+00	2.892e+00	2.282e+00

APPENDIX D
COMPLEMENTARY ILLUSTRATIONS ASSOCIATED WITH THE EXPERIMENTS

A. *Experiments on synthetic data ($R = 3$)*

The abundances associated with the two remaining endmembers are given in Figs. 15 and 16.

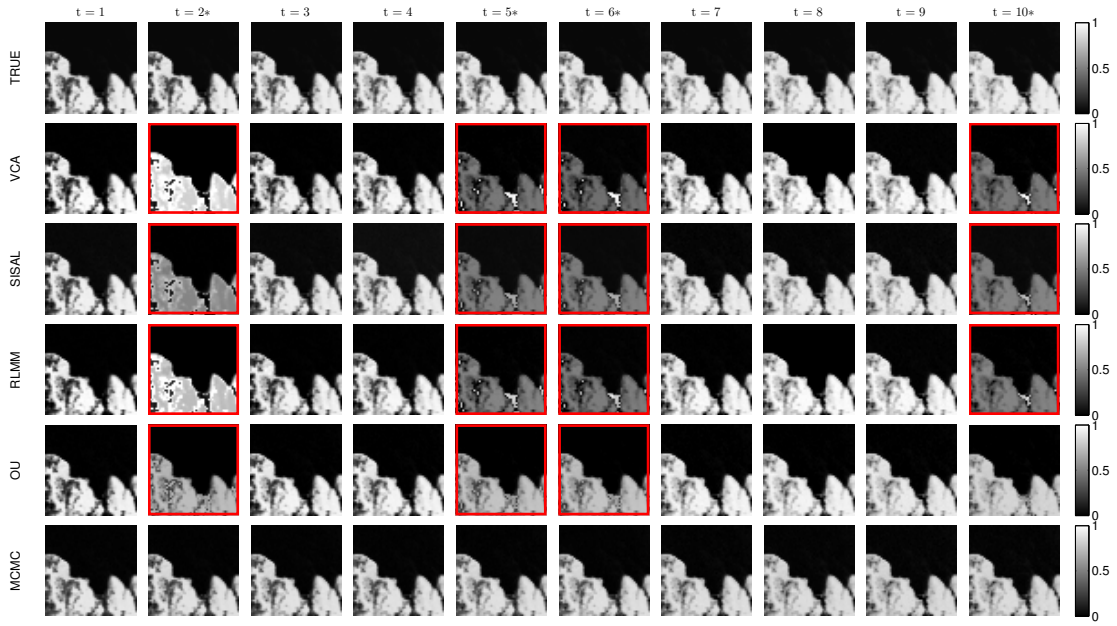


Fig. 15. Abundance map of the second endmember recovered by the different methods (in each row) at each time instant (given in column) for the experiment with $R = 3$ [the different rows correspond to the true abundances, VCA/FCLS, SISAL/FCLS, RLMM, OU and the proposed method]. The images delineated in red show that several methods are highly sensitive to the presence of outliers, and the time instants represented with $*$ denote images containing outliers.

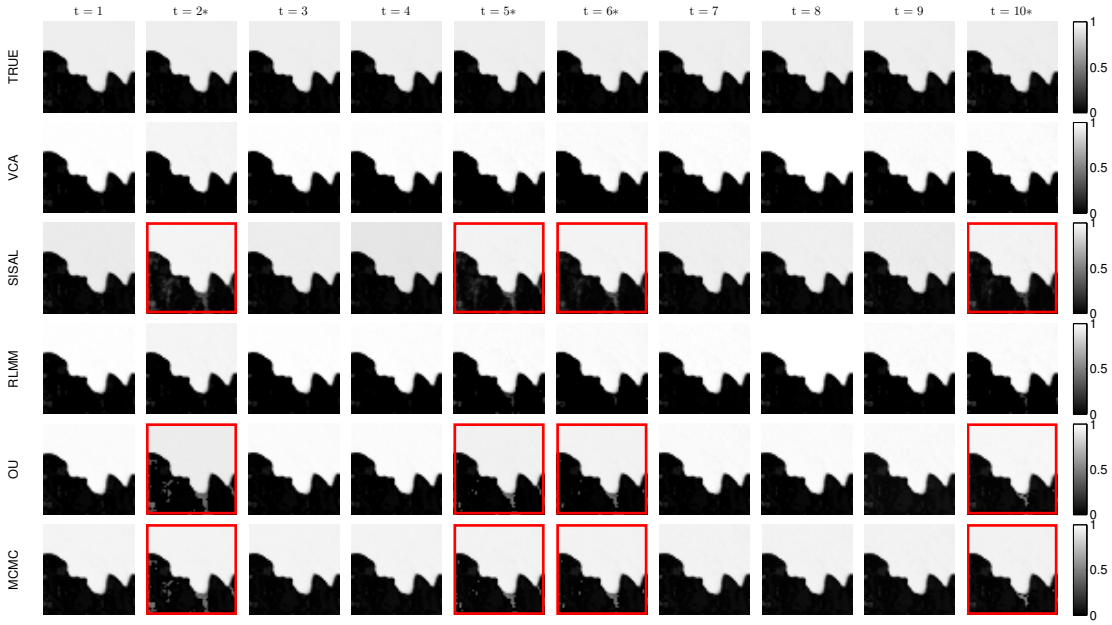


Fig. 16. Abundance map of the third endmember recovered by the different methods (in each row) at each time instant (given in column) for the experiment with $R = 3$ [the different rows correspond to the true abundances, VCA/FCLS, SISAL/FCLS, RLMM, OU and the proposed method]. The images delineated in red show that several methods are highly sensitive to the presence of outliers, and the time instants represented with $*$ denote images containing outliers.

B. *Experiments on synthetic data ($R = 6$)*

The abundances, endmembers and the estimated labels are depicted in Figs. 17 to 25 for the datasets composed of $R = 6$. Note that the synthetic abundance maps involved in these experiments are those used in [53].

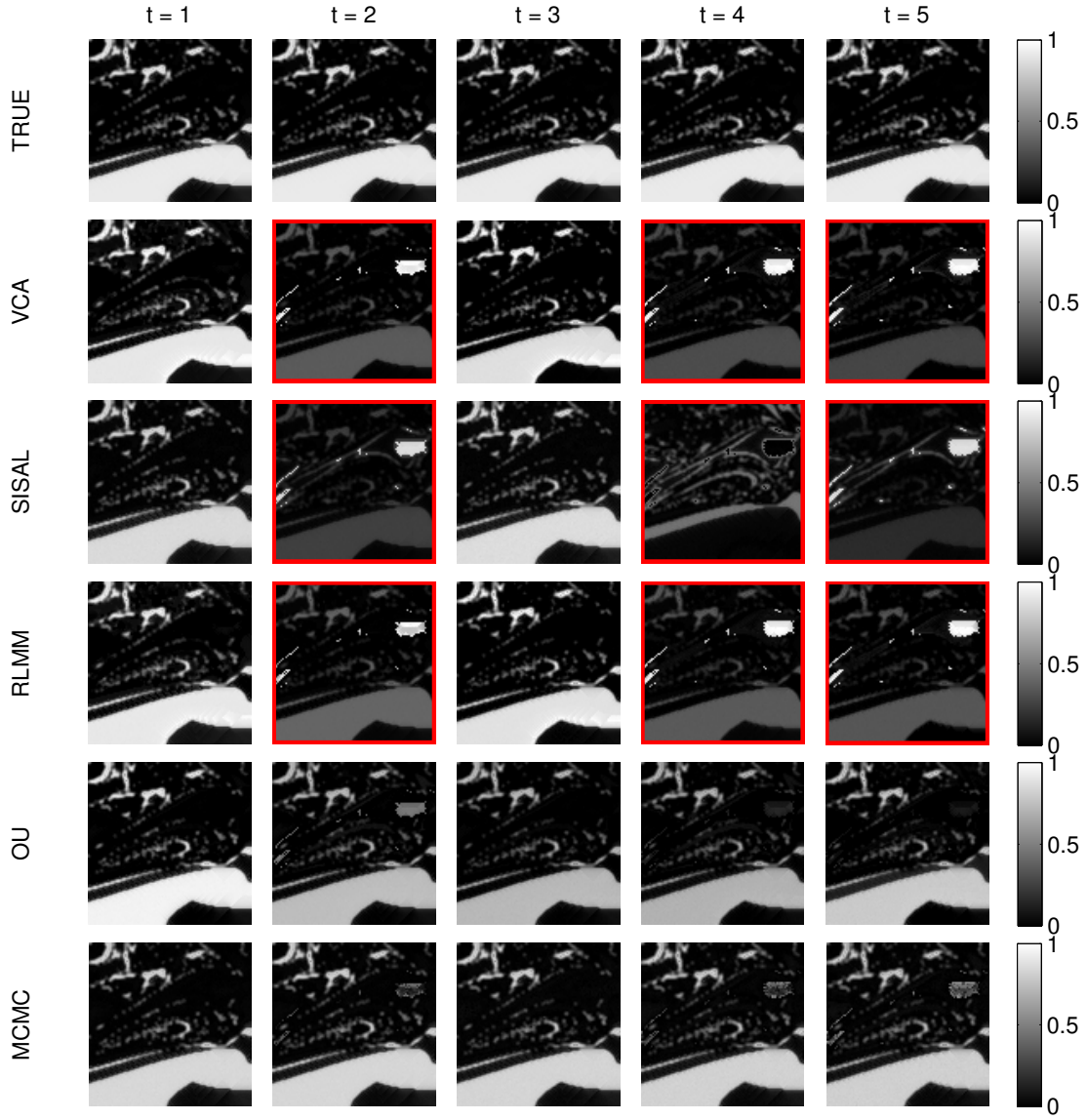


Fig. 17. Abundance map of the endmember 1 recovered by the different methods (in each row) at each time instant (given in column) for the experiment with $R = 6$ [the different rows correspond to the true abundances, VCA/FCLS, SISAL/FCLS, RLMM, OU and the proposed method]. The images delineated in red show that several methods are highly sensitive to the presence of outliers, and the time instants represented with $*$ denote images containing outliers.

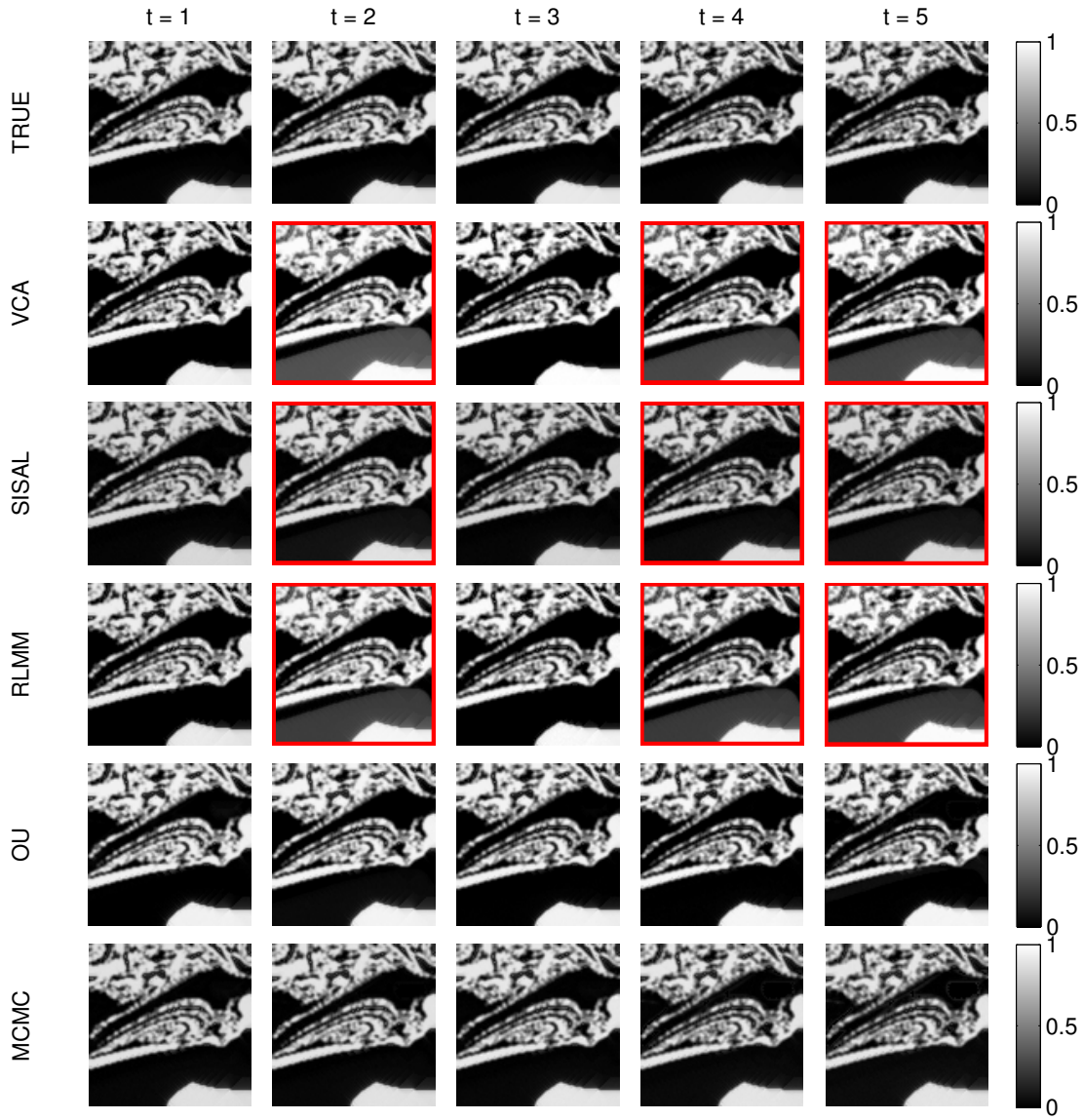


Fig. 18. Abundance map of the endmember 2 recovered by the different methods (in each row) at each time instant (given in column) for the experiment with $R = 6$ [the different rows correspond to the true abundances, VCA/FCLS, SISAL/FCLS, RLMM, OU and the proposed method]. The images delineated in red show that several methods are highly sensitive to the presence of outliers, and the time instants represented with $*$ denote images containing outliers.

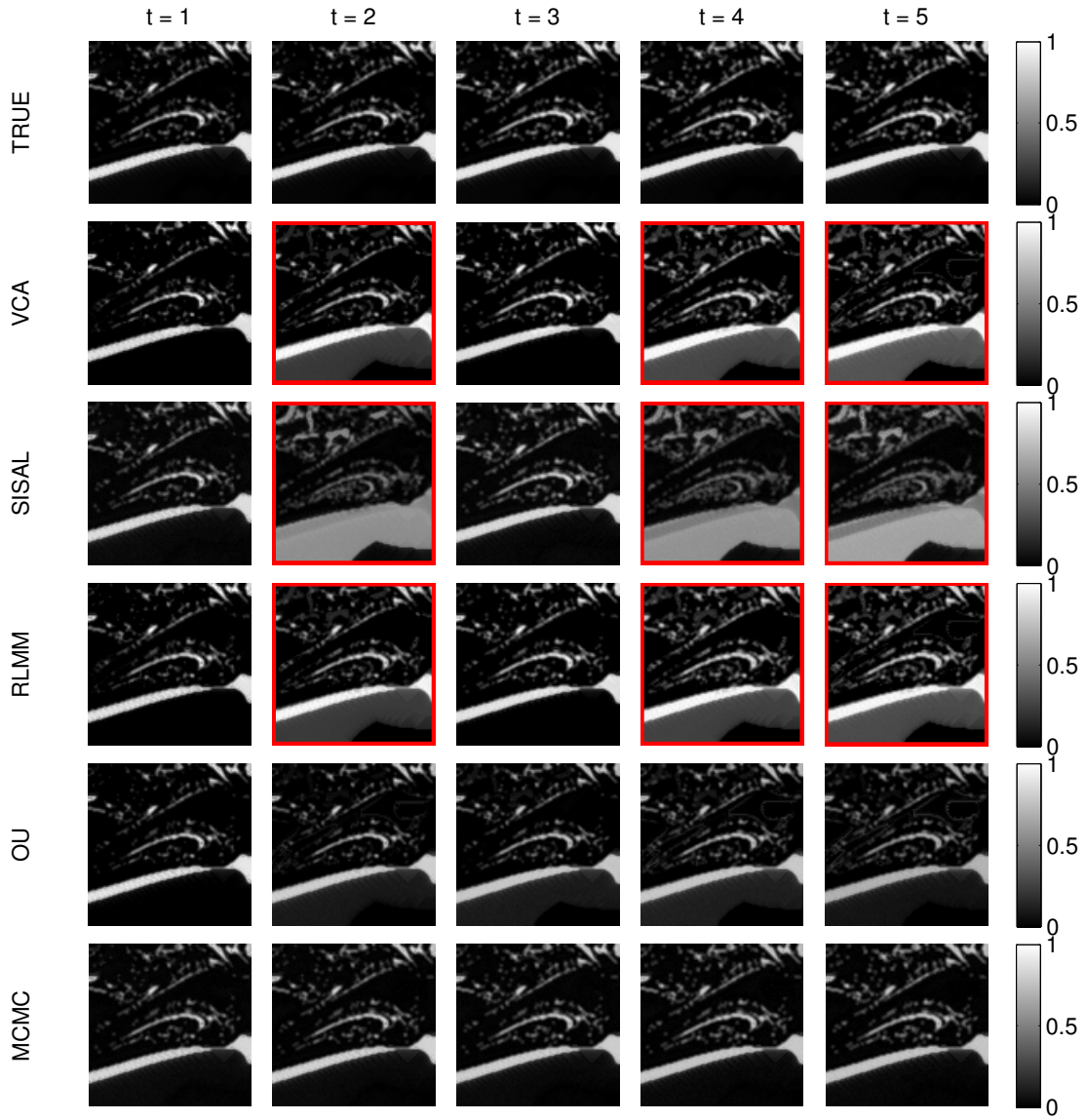


Fig. 19. Abundance map of the endmember 3 recovered by the different methods (in each row) at each time instant (given in column) for the experiment with $R = 6$ [the different rows correspond to the true abundances, VCA/FCLS, SISAL/FCLS, RLMM, OU and the proposed method]. The images delineated in red show that several methods are highly sensitive to the presence of outliers, and the time instants represented with $*$ denote images containing outliers.

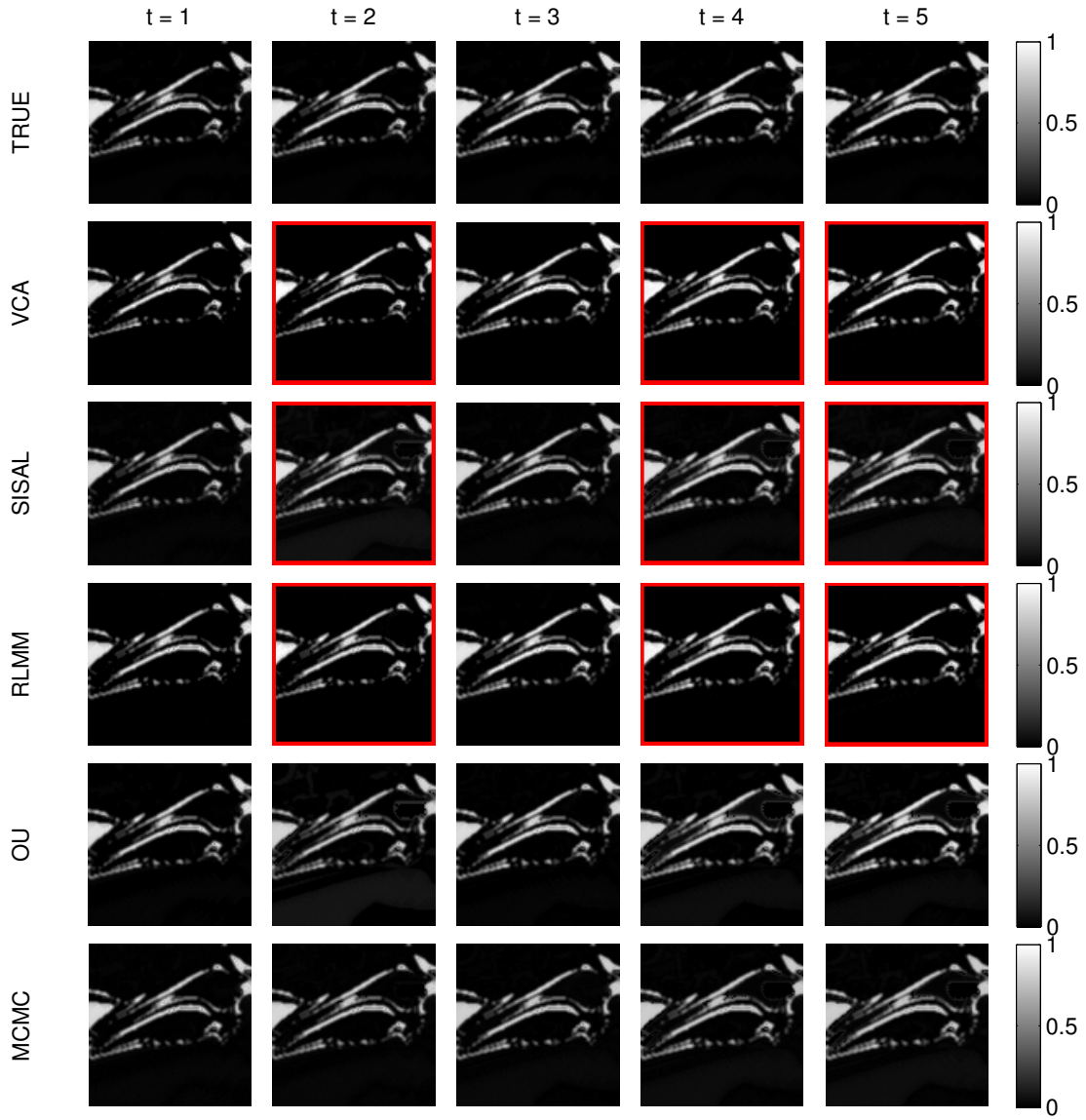


Fig. 20. Abundance map of the endmember 4 recovered by the different methods (in each row) at each time instant (given in column) for the experiment with $R = 6$ [the different rows correspond to the true abundances, VCA/FCLS, SISAL/FCLS, RLMM, OU and the proposed method]. The images delineated in red show that several methods are highly sensitive to the presence of outliers, and the time instants represented with $*$ denote images containing outliers.

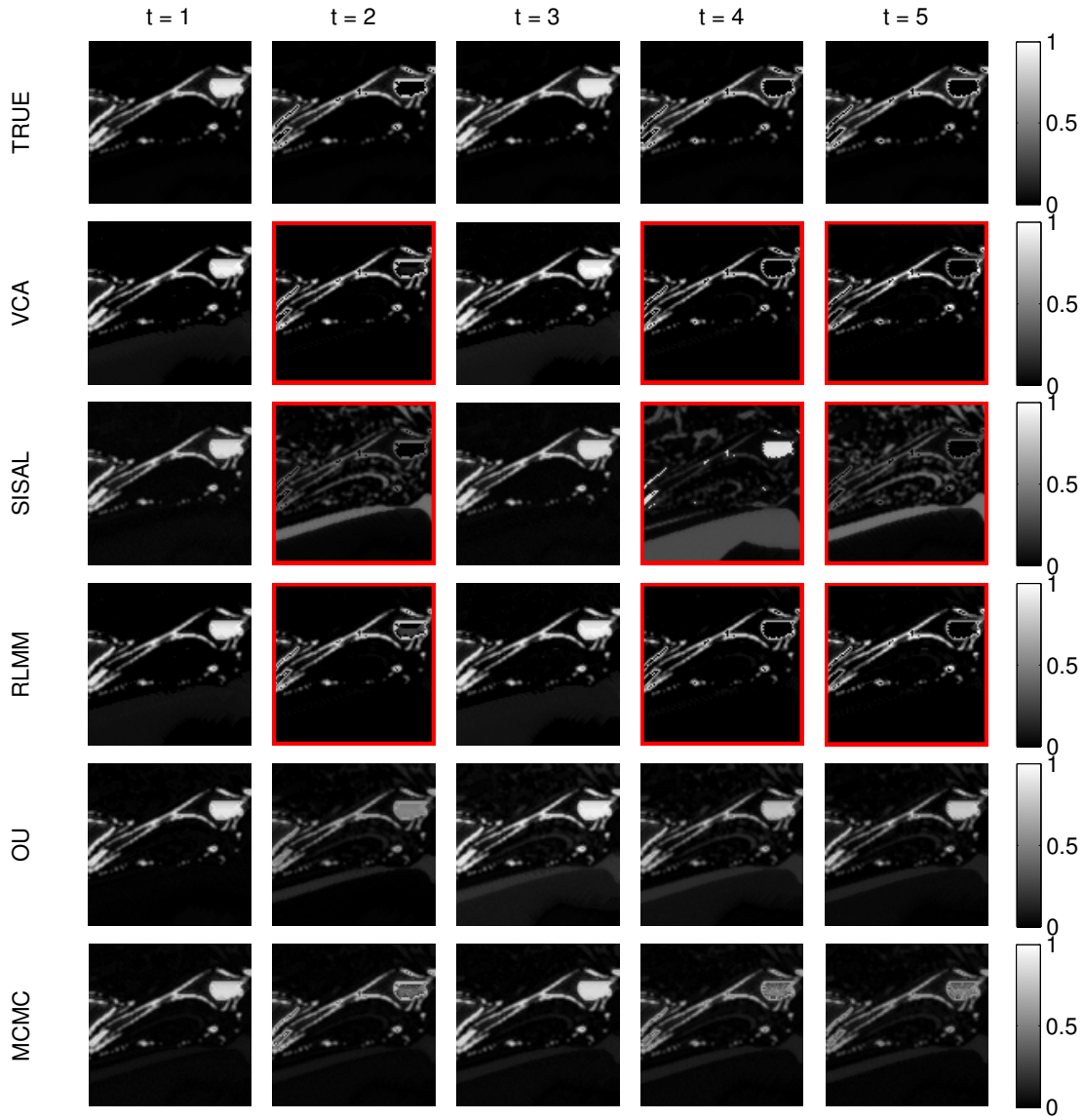


Fig. 21. Abundance map of the endmember 5 recovered by the different methods (in each row) at each time instant (given in column) for the experiment with $R = 6$ [the different rows correspond to the true abundances, VCA/FCLS, SISAL/FCLS, RLMM, OU and the proposed method]. The images delineated in red show that several methods are highly sensitive to the presence of outliers, and the time instants represented with $*$ denote images containing outliers.

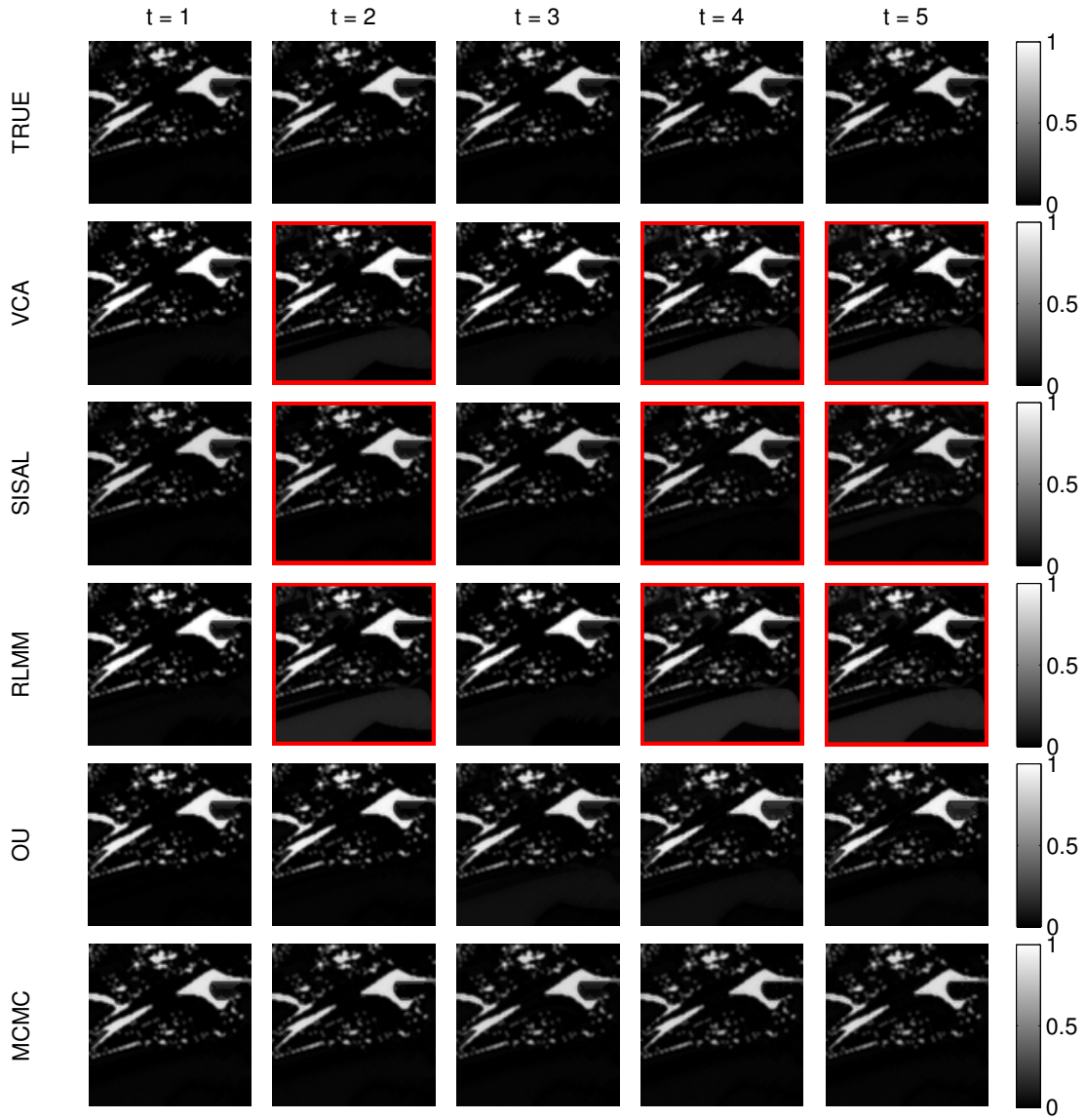
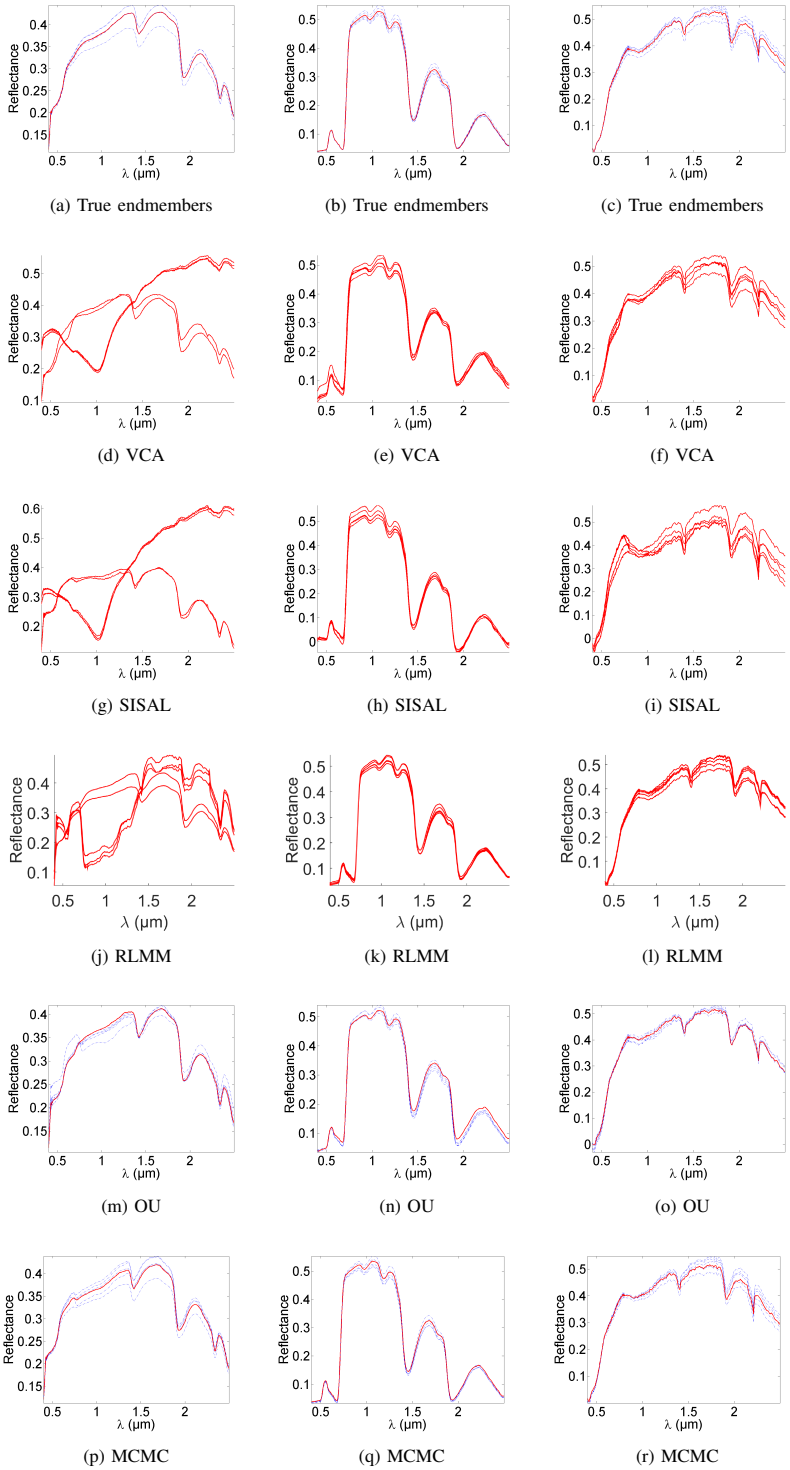


Fig. 22. Abundance map of the endmember 6 recovered by the different methods (in each row) at each time instant (given in column) for the experiment with $R = 6$ [the different rows correspond to the true abundances, VCA/FCLS, SISAL/FCLS, RLMM, OU and the proposed method]. The images delineated in red show that several methods are highly sensitive to the presence of outliers, and the time instants represented with $*$ denote images containing outliers.



(a) True endmembers (b) True endmembers (c) True endmembers

(d) VCA (e) VCA (f) VCA

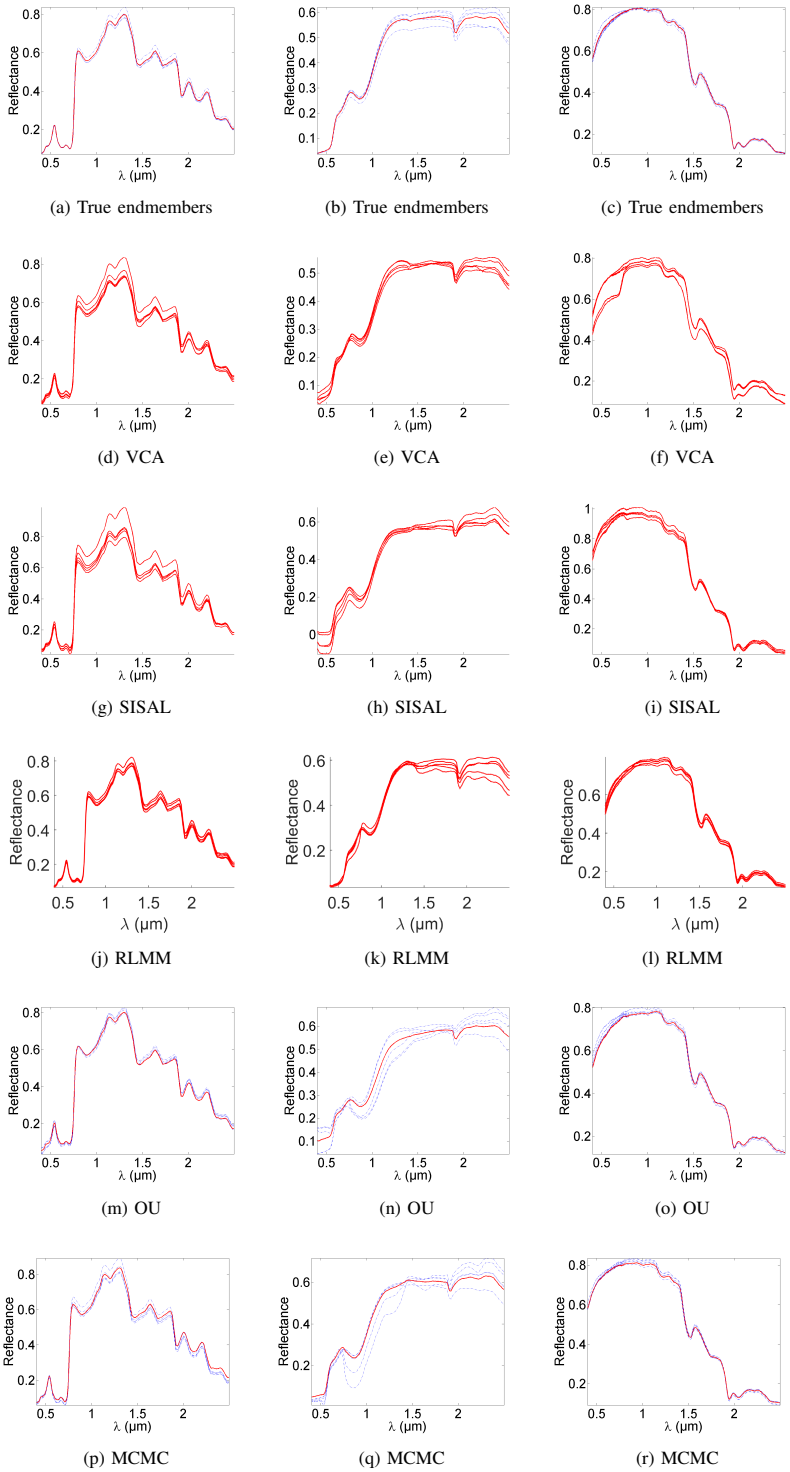
(g) SISAL (h) SISAL (i) SISAL

(j) RLMM (k) RLMM (l) RLMM

(m) OU (n) OU (o) OU

(p) MCMC (q) MCMC (r) MCMC

Fig. 23. Endmembers ($\mathbf{m}_r$, red lines) and their variants affected by variability ($\mathbf{m}_r + \mathbf{dm}_{r,t}$, blue dotted lines) recovered by the different methods from the synthetic mixtures with $R = 6$. Signatures corresponding to different time instants are represented on a single figure to better appreciate the variability recovered from the data. The spectra represented in black correspond to signatures corrupted by outliers.



(a) True endmembers (b) True endmembers (c) True endmembers

(d) VCA (e) VCA (f) VCA

(g) SISAL (h) SISAL (i) SISAL

(j) RLMM (k) RLMM (l) RLMM

(m) OU (n) OU (o) OU

(p) MCMC (q) MCMC (r) MCMC

Fig. 24. Endmembers ($\mathbf{m}_r$, red lines) and their variants affected by variability ($\mathbf{m}_r + \mathbf{dm}_{r,t}$, blue dotted lines) recovered by the different methods from the synthetic mixtures with $R = 6$. Signatures corresponding to different time instants are represented on a single figure to better appreciate the variability recovered from the data. The spectra represented in black correspond to signatures corrupted by outliers.

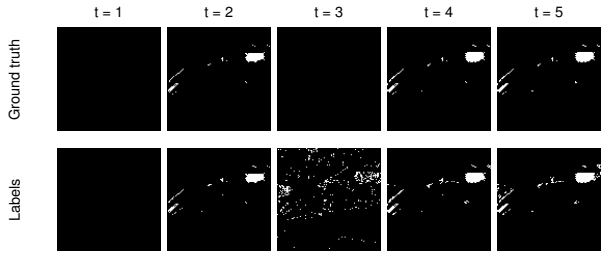


Fig. 25. Outlier labels $\mathbf{z}_t$ estimated for each image of the synthetic dataset with $R = 6$ (the different rows correspond to the true labels, and the estimated labels) [0 in black, 1 in white].

C. Experiments on synthetic data ($R = 9$)

The abundances, endmembers and the estimated labels are depicted in Figs. 26 to 38 for the datasets composed of $R = 9$. Note that the synthetic abundance maps involved in these experiments are those used in [53].

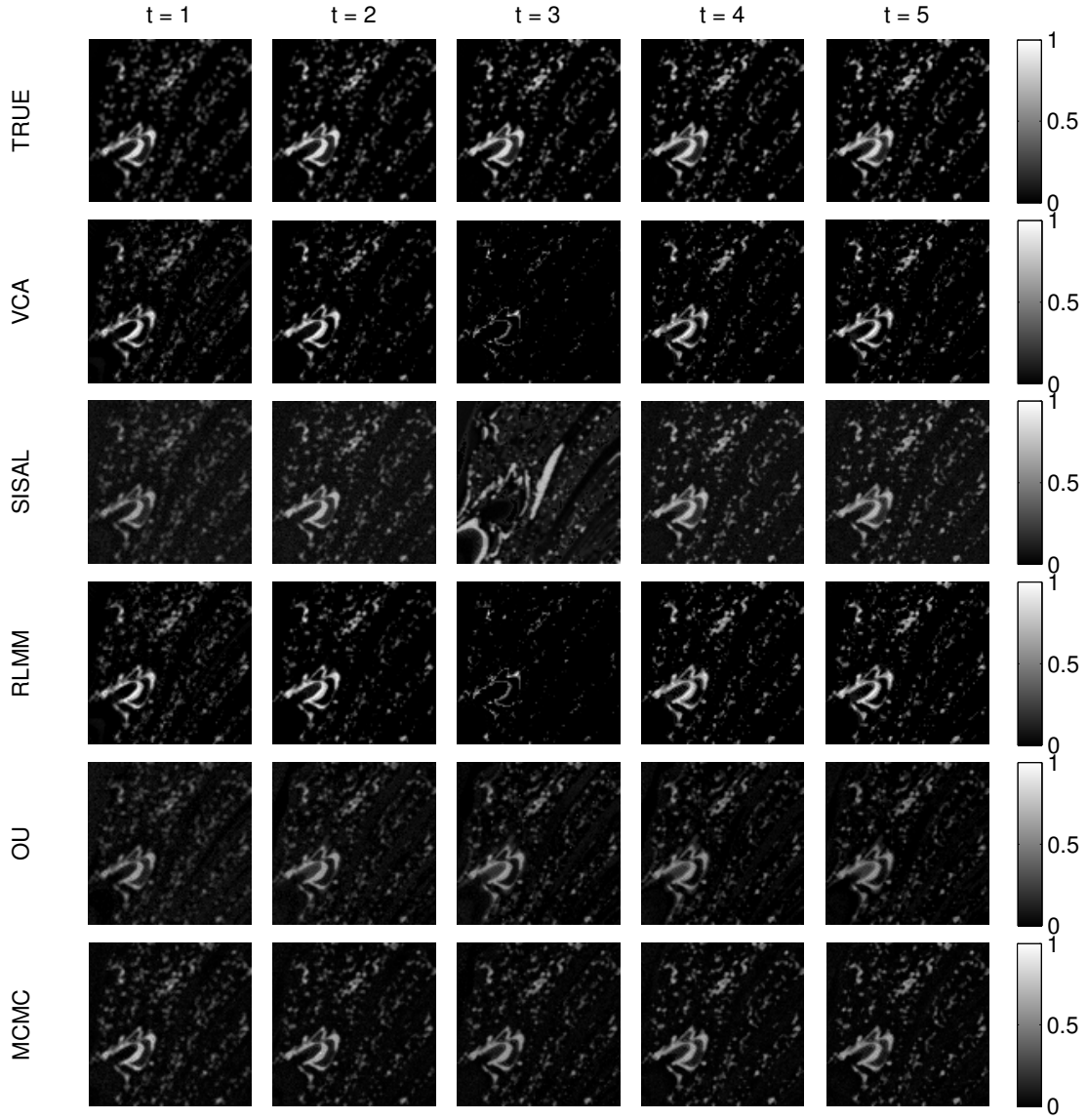


Fig. 26. Abundance map of the endmember 1 recovered by the different methods (in each row) at each time instant (given in column) for the experiment with $R = 6$ [the different rows correspond to the true abundances, VCA/FCLS, SISAL/FCLS, RLMM, OU and the proposed method]. The images delineated in red show that several methods are highly sensitive to the presence of outliers, and the time instants represented with $*$ denote images containing outliers.

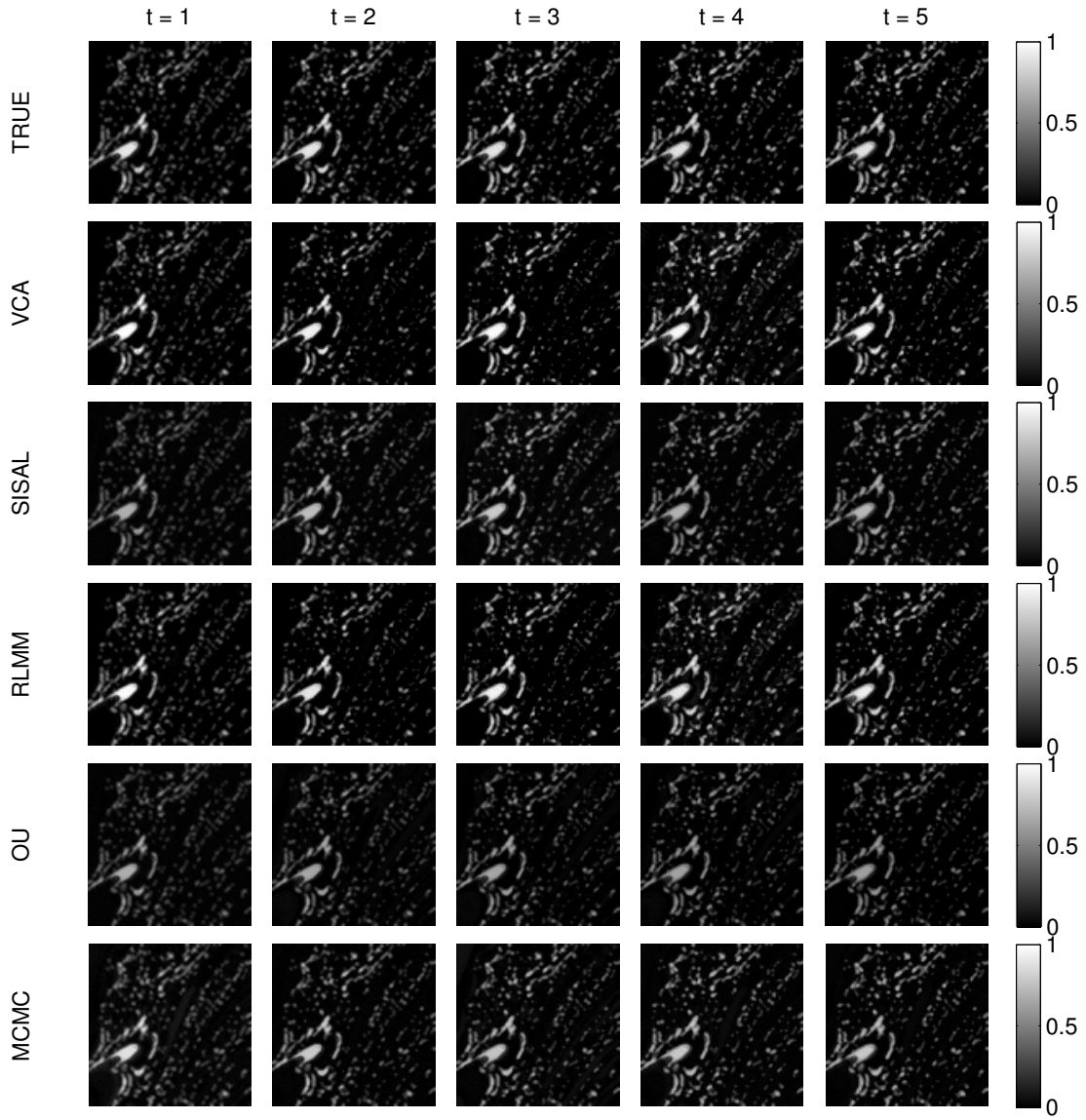


Fig. 27. Abundance map of the endmember 2 recovered by the different methods (in each row) at each time instant (given in column) for the experiment with $R = 6$ [the different rows correspond to the true abundances, VCA/FCLS, SISAL/FCLS, RLMM, OU and the proposed method]. The images delineated in red show that several methods are highly sensitive to the presence of outliers, and the time instants represented with $*$ denote images containing outliers.

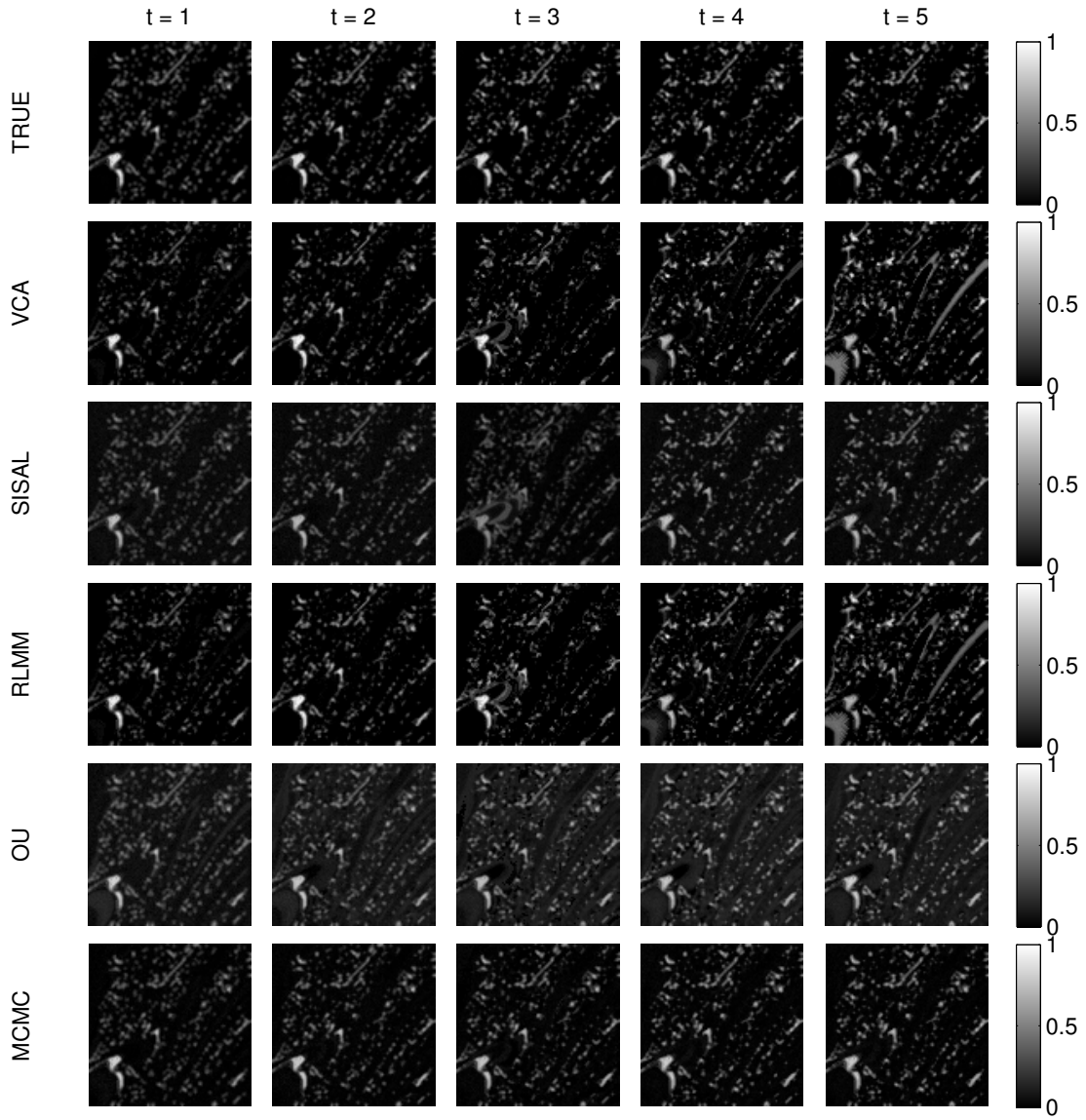


Fig. 28. Abundance map of the endmember 3 recovered by the different methods (in each row) at each time instant (given in column) for the experiment with $R = 6$ [the different rows correspond to the true abundances, VCA/FCLS, SISAL/FCLS, RLMM, OU and the proposed method]. The images delineated in red show that several methods are highly sensitive to the presence of outliers, and the time instants represented with $*$ denote images containing outliers.

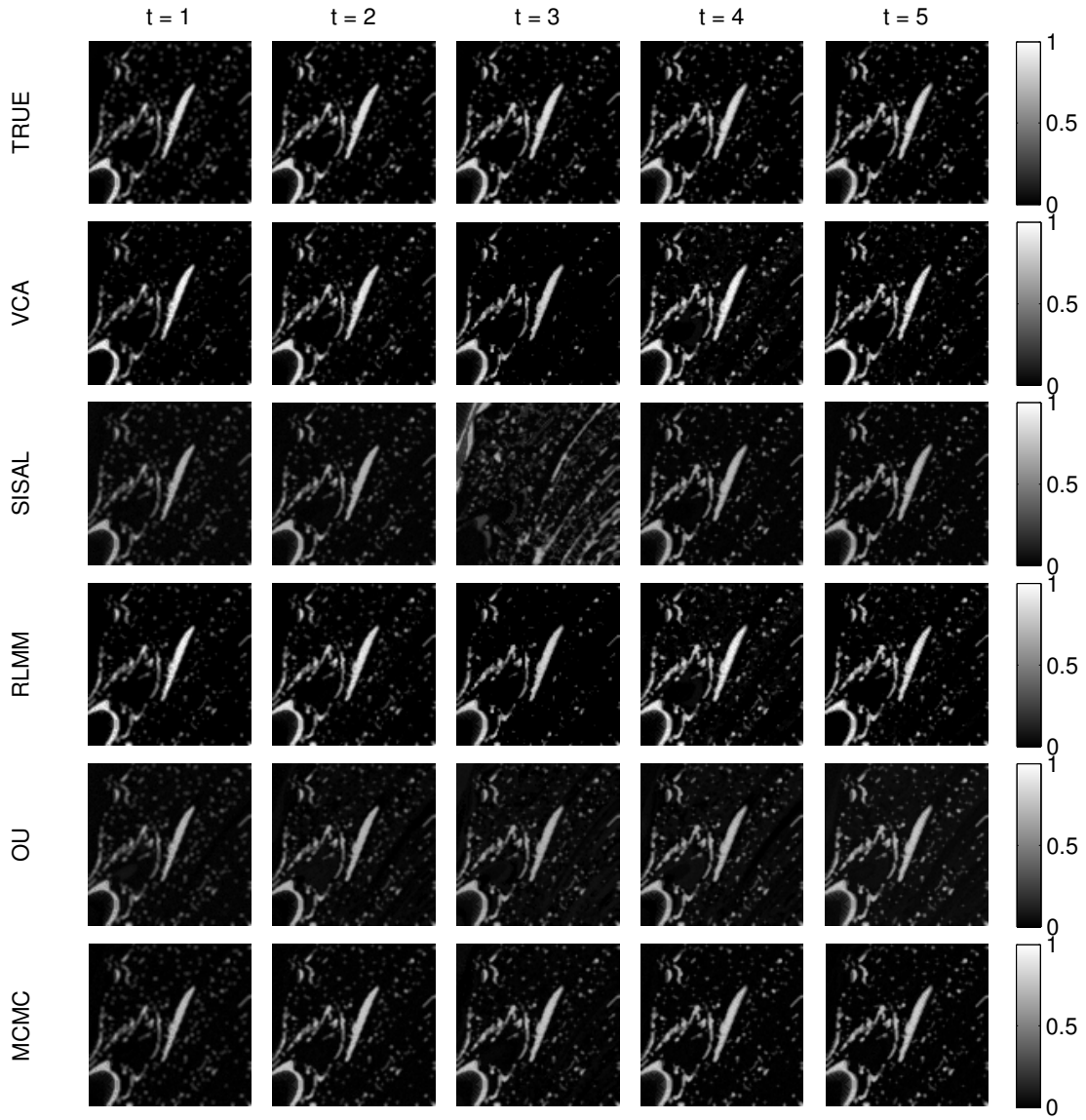


Fig. 29. Abundance map of the endmember 4 recovered by the different methods (in each row) at each time instant (given in column) for the experiment with $R = 6$ [the different rows correspond to the true abundances, VCA/FCLS, SISAL/FCLS, RLMM, OU and the proposed method]. The images delineated in red show that several methods are highly sensitive to the presence of outliers, and the time instants represented with $*$ denote images containing outliers.

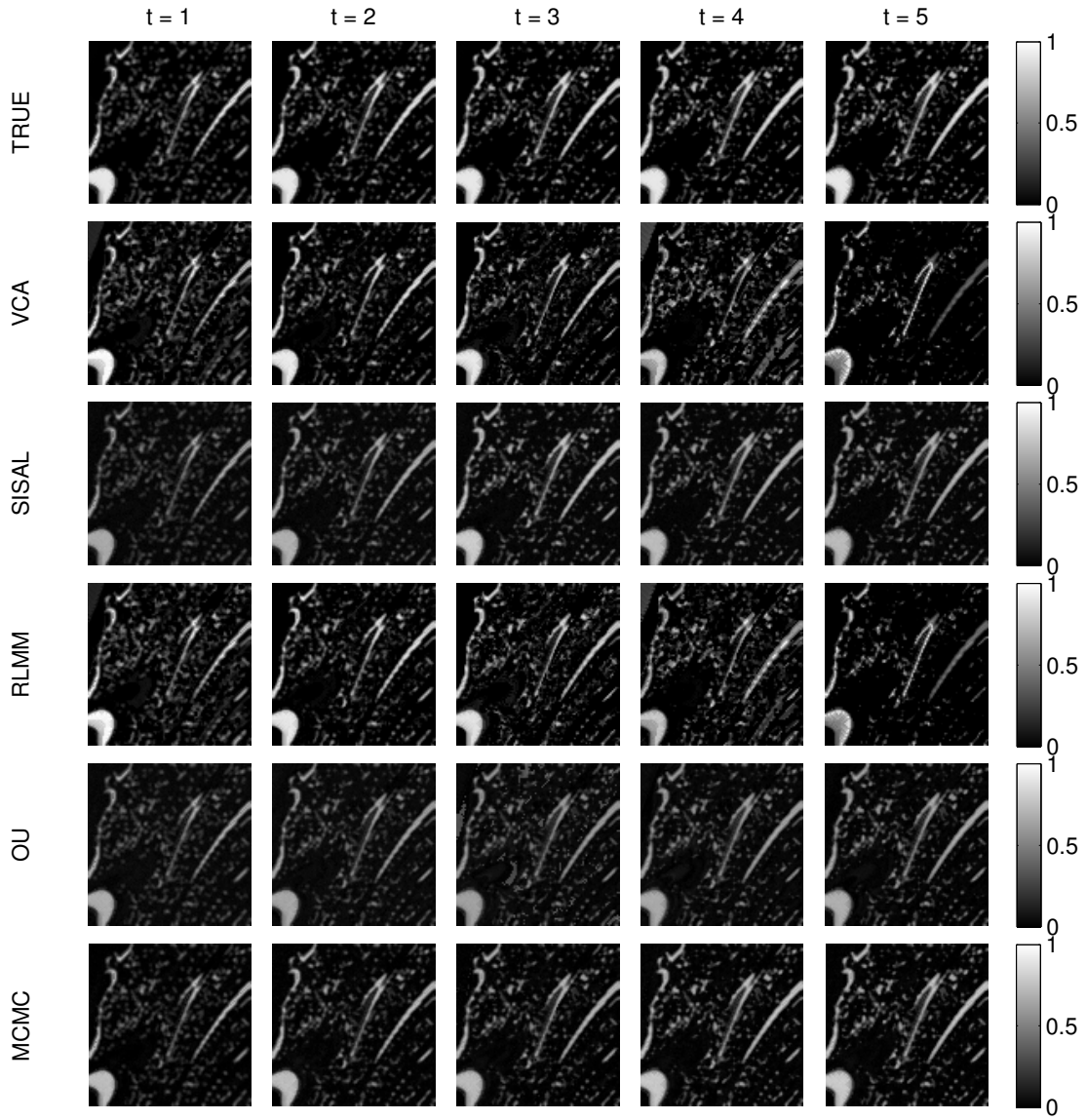


Fig. 30. Abundance map of the endmember 5 recovered by the different methods (in each row) at each time instant (given in column) for the experiment with $R = 6$ [the different rows correspond to the true abundances, VCA/FCLS, SISAL/FCLS, RLMM, OU and the proposed method]. The images delineated in red show that several methods are highly sensitive to the presence of outliers, and the time instants represented with $*$ denote images containing outliers.

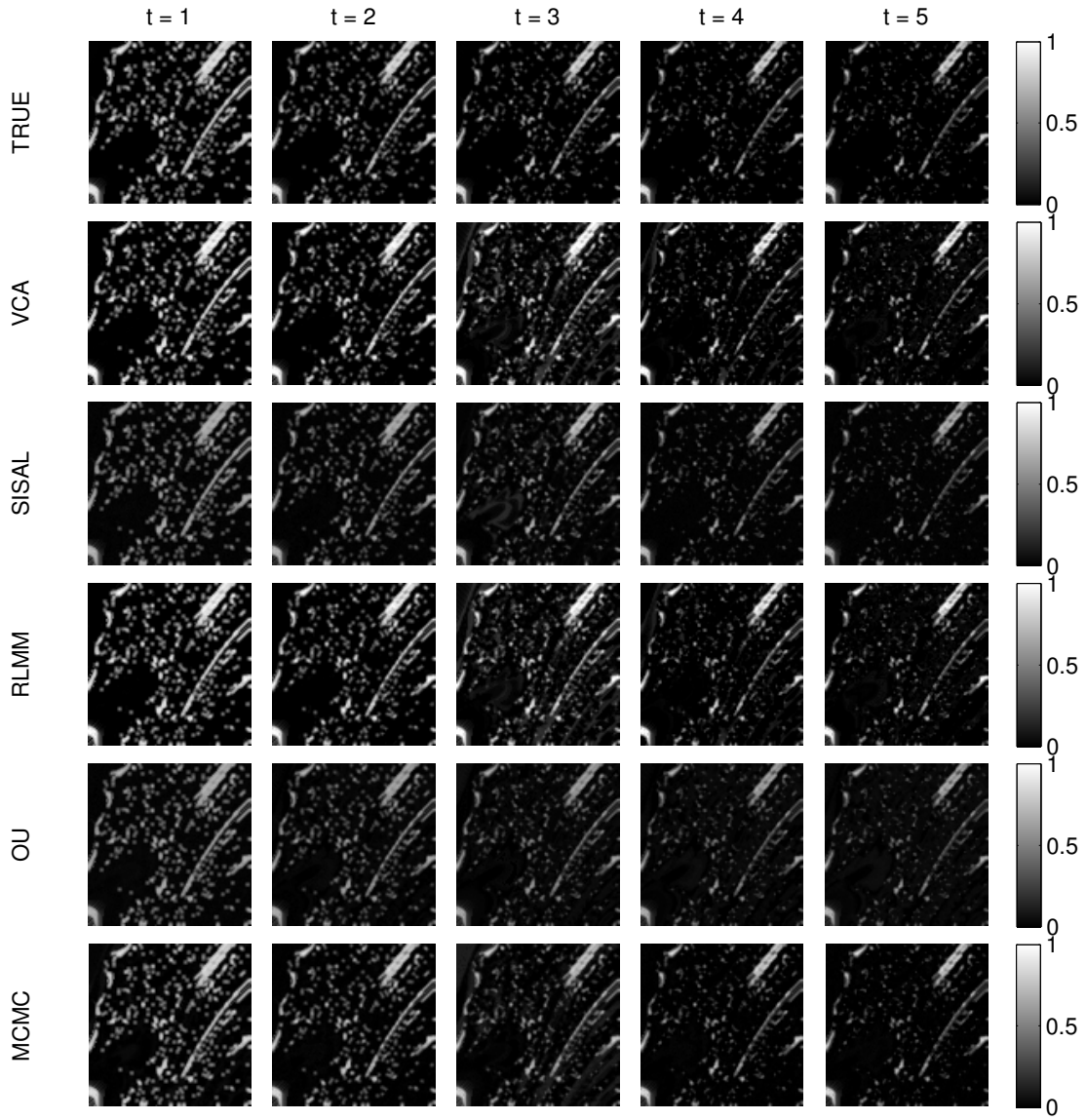


Fig. 31. Abundance map of the endmember 6 recovered by the different methods (in each row) at each time instant (given in column) for the experiment with $R = 6$ [the different rows correspond to the true abundances, VCA/FCLS, SISAL/FCLS, RLMM, OU and the proposed method]. The images delineated in red show that several methods are highly sensitive to the presence of outliers, and the time instants represented with $*$ denote images containing outliers.

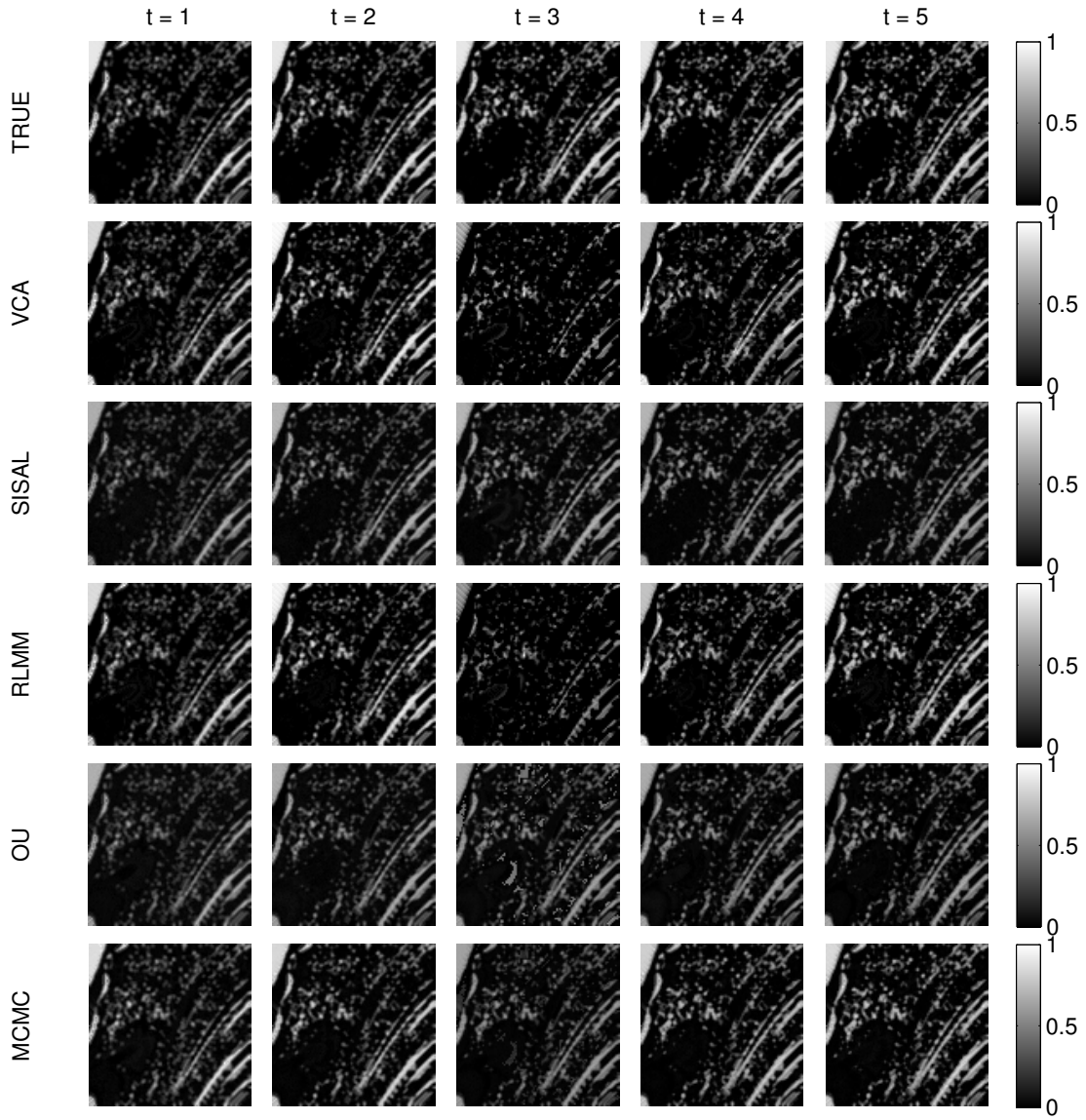


Fig. 32. Abundance map of the endmember 7 recovered by the different methods (in each row) at each time instant (given in column) for the experiment with $R = 6$ [the different rows correspond to the true abundances, VCA/FCLS, SISAL/FCLS, RLMM, OU and the proposed method]. The images delineated in red show that several methods are highly sensitive to the presence of outliers, and the time instants represented with $*$ denote images containing outliers.

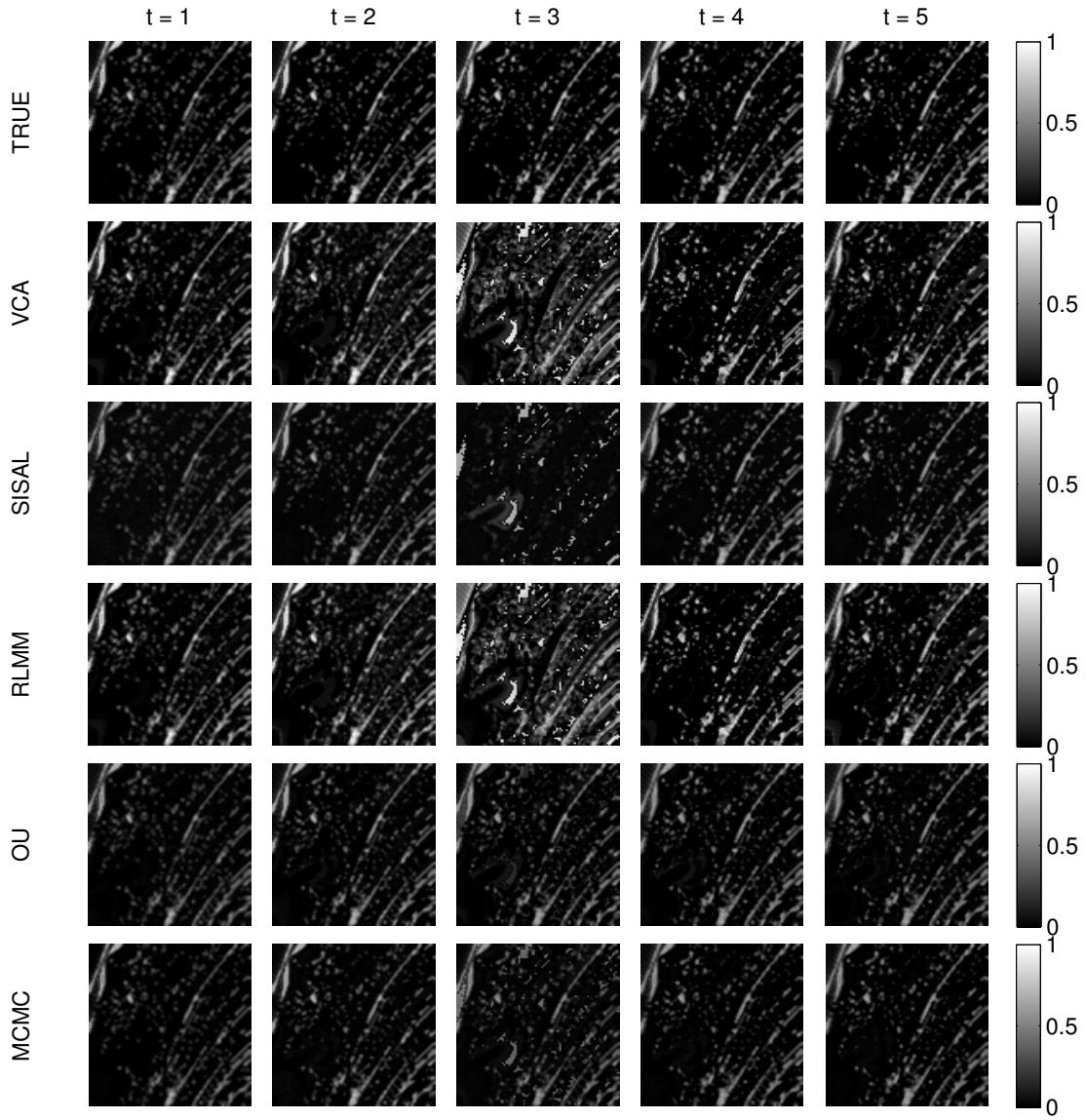


Fig. 33. Abundance map of the endmember 8 recovered by the different methods (in each row) at each time instant (given in column) for the experiment with $R = 6$ [the different rows correspond to the true abundances, VCA/FCLS, SISAL/FCLS, RLMM, OU and the proposed method]. The images delineated in red show that several methods are highly sensitive to the presence of outliers, and the time instants represented with $*$ denote images containing outliers.

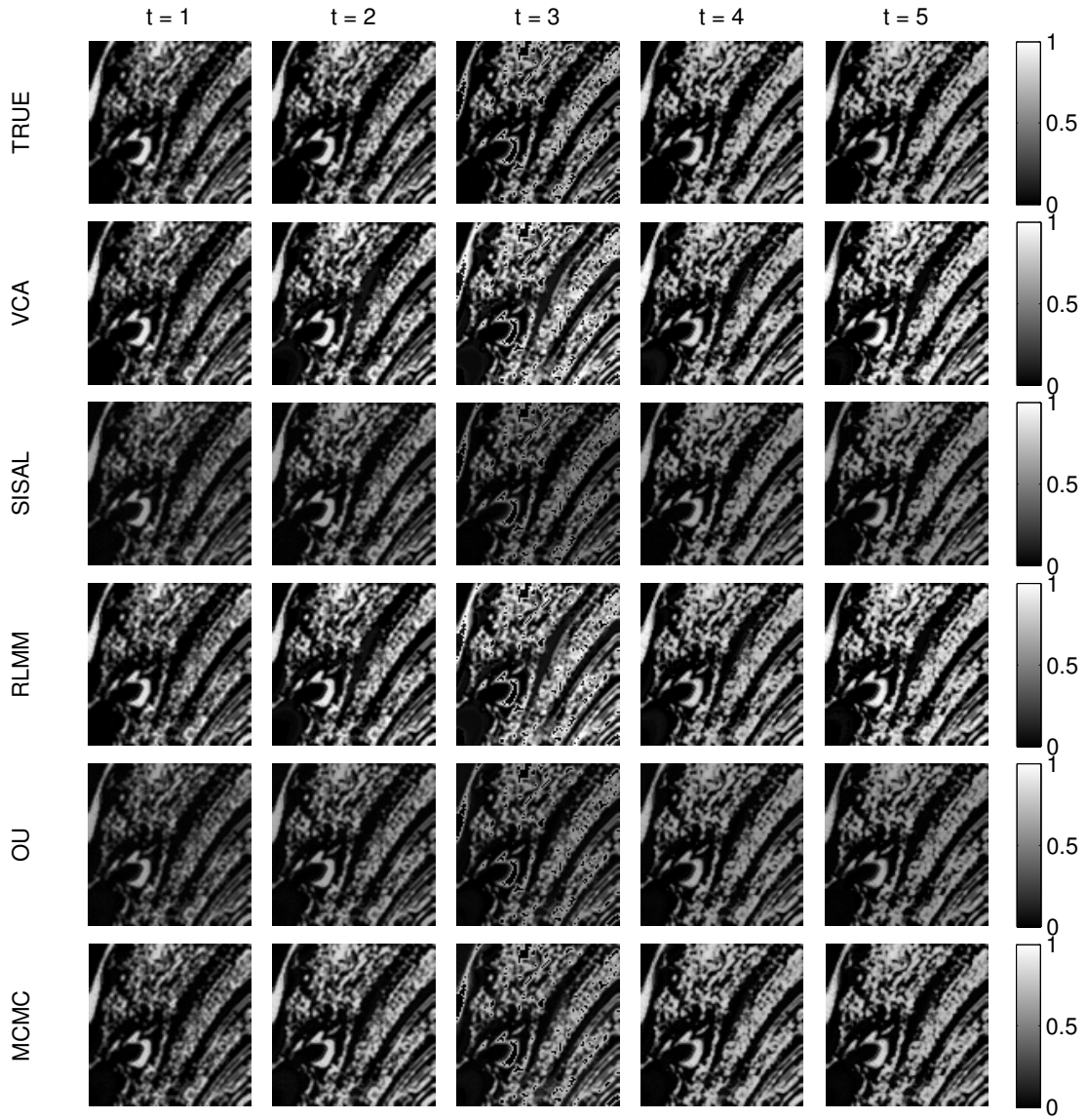
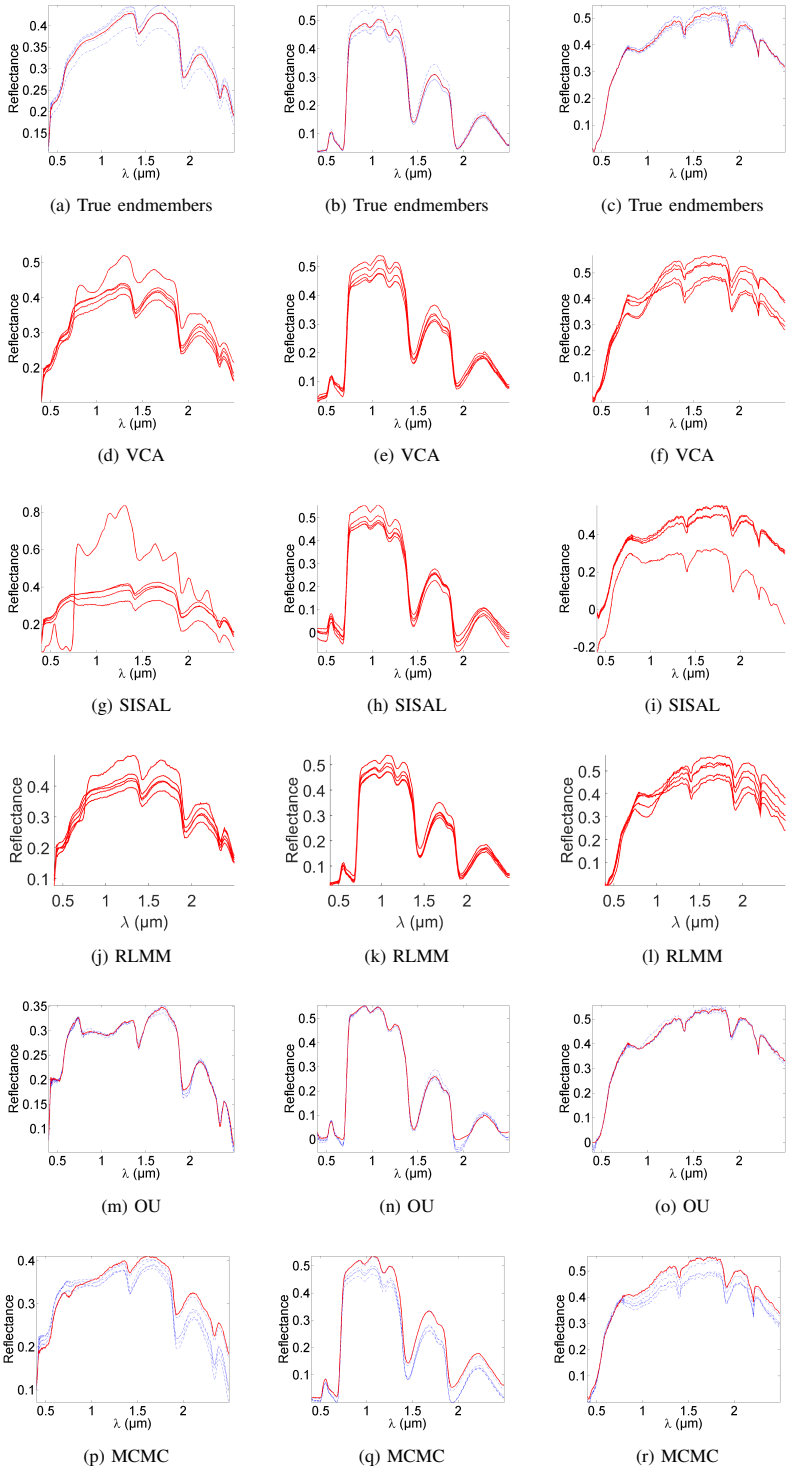


Fig. 34. Abundance map of the endmember 9 recovered by the different methods (in each row) at each time instant (given in column) for the experiment with $R = 6$ [the different rows correspond to the true abundances, VCA/FCLS, SISAL/FCLS, RLMM, OU and the proposed method]. The images delineated in red show that several methods are highly sensitive to the presence of outliers, and the time instants represented with $*$ denote images containing outliers.



(a) True endmembers (b) True endmembers (c) True endmembers

(d) VCA (e) VCA (f) VCA

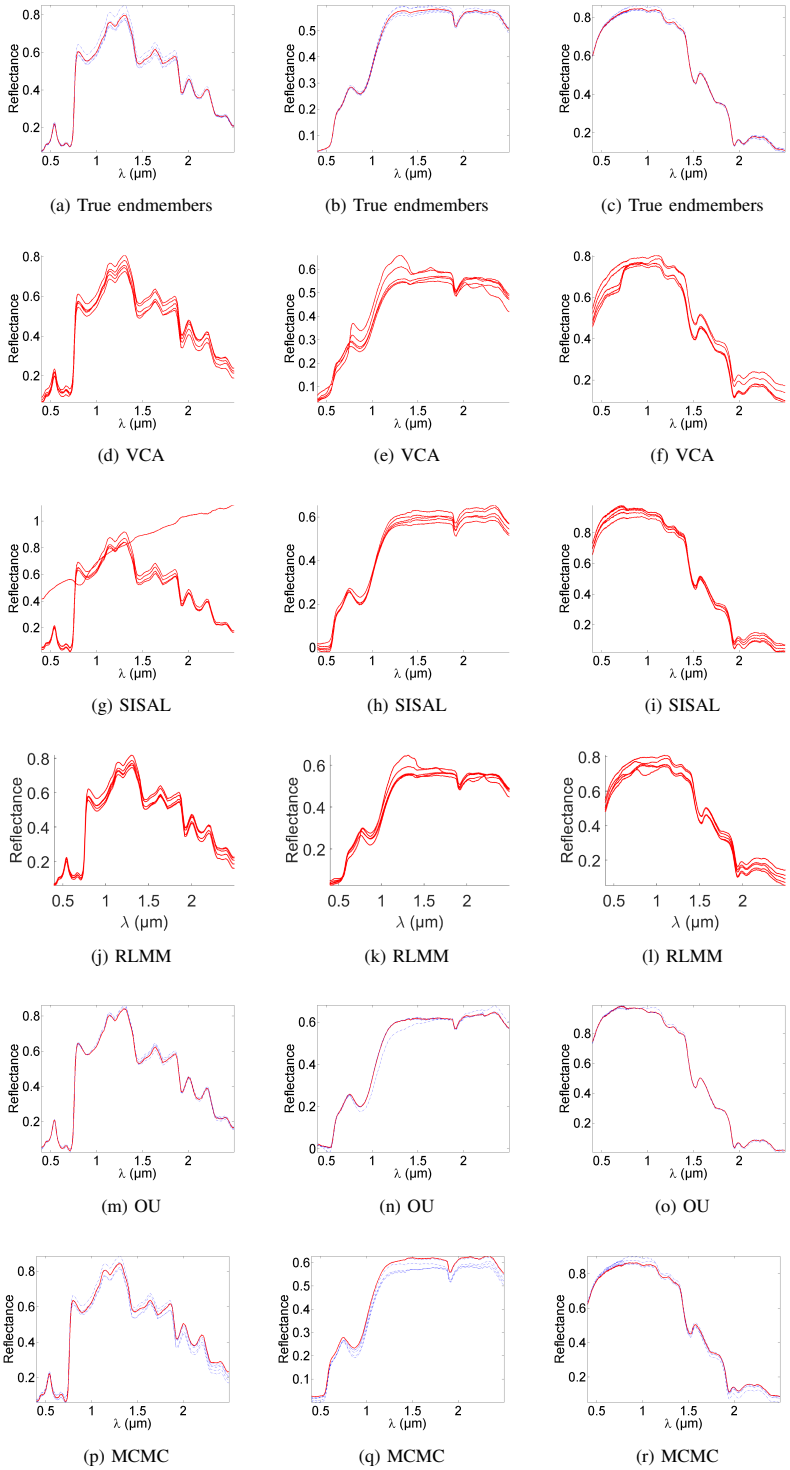
(g) SISAL (h) SISAL (i) SISAL

(j) RLMM (k) RLMM (l) RLMM

(m) OU (n) OU (o) OU

(p) MCMC (q) MCMC (r) MCMC

Fig. 35. Endmembers ($\mathbf{m}_r$, red lines) and their variants affected by variability ($\mathbf{m}_r + \mathbf{dm}_{r,t}$, blue dotted lines) recovered by the different methods from the synthetic mixtures with $R = 9$. Signatures corresponding to different time instants are represented on a single figure to better appreciate the variability recovered from the data.



(a) True endmembers (b) True endmembers (c) True endmembers

(d) VCA (e) VCA (f) VCA

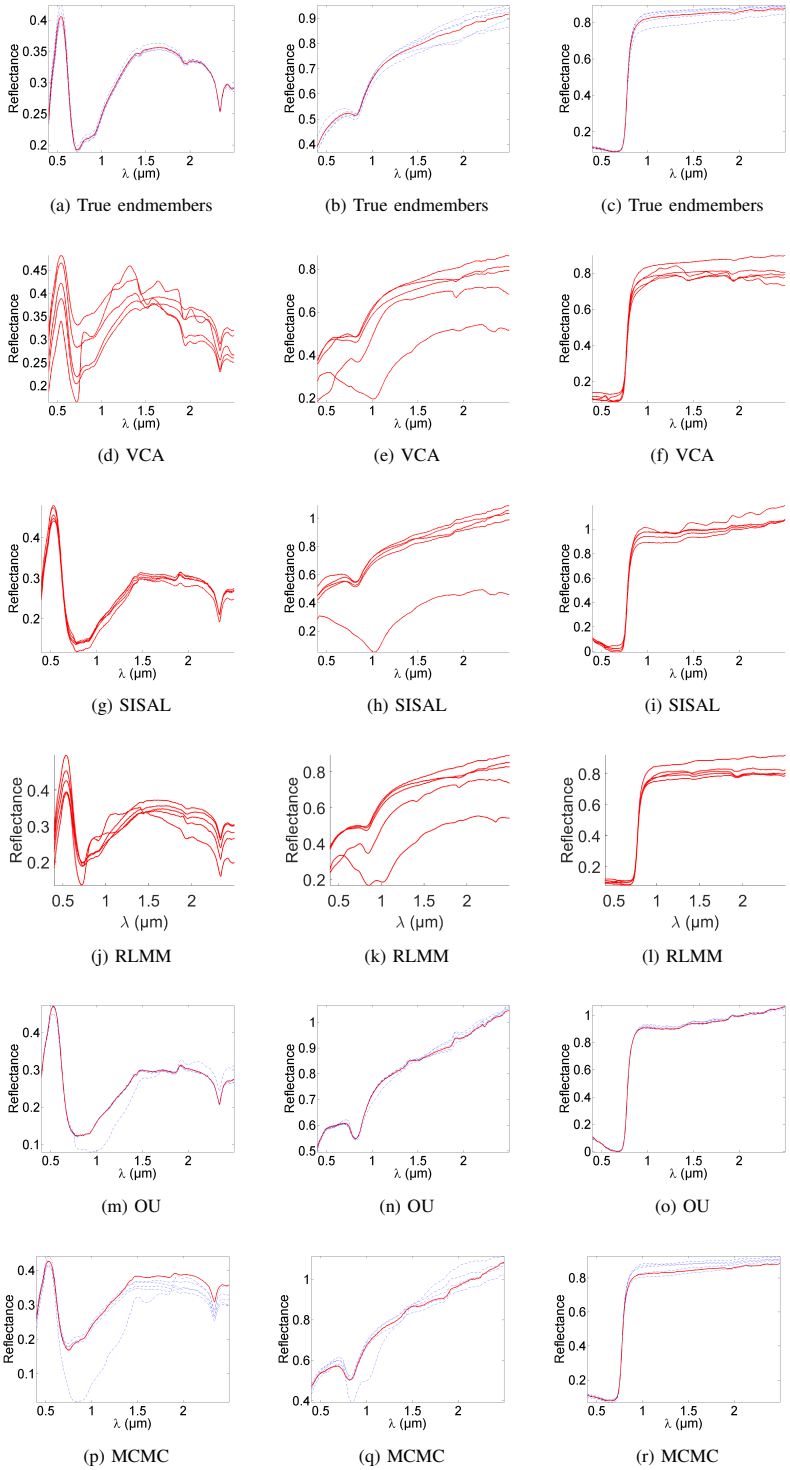
(g) SISAL (h) SISAL (i) SISAL

(j) RLMM (k) RLMM (l) RLMM

(m) OU (n) OU (o) OU

(p) MCMC (q) MCMC (r) MCMC

Fig. 36. Endmembers ($\mathbf{m}_r$, red lines) and their variants affected by variability ($\mathbf{m}_r + \mathbf{dm}_{r,t}$, blue dotted lines) recovered by the different methods from the synthetic mixtures with $R = 9$. Signatures corresponding to different time instants are represented on a single figure to better appreciate the variability recovered from the data.



(a) True endmembers (b) True endmembers (c) True endmembers

(d) VCA (e) VCA (f) VCA

(g) SISAL (h) SISAL (i) SISAL

(j) RLMM (k) RLMM (l) RLMM

(m) OU (n) OU (o) OU

(p) MCMC (q) MCMC (r) MCMC

Fig. 37. Endmembers ($\mathbf{m}_r$, red lines) and their variants affected by variability ($\mathbf{m}_r + \mathbf{dm}_{r,t}$, blue dotted lines) recovered by the different methods from the synthetic mixtures with $R = 9$. Signatures corresponding to different time instants are represented on a single figure to better appreciate the variability recovered from the data.

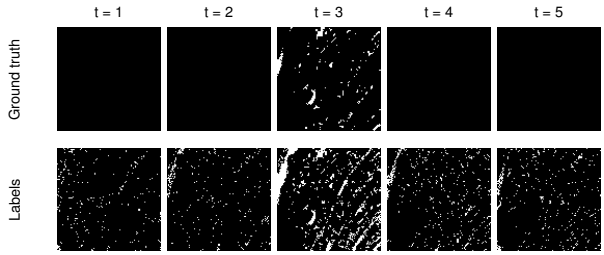


Fig. 38. Outlier labels $\mathbf{z}_t$ estimated for each image of the synthetic dataset with $R = 9$ (the different rows correspond to the true labels, and the estimated labels) [0 in black, 1 in white].

APPENDIX E
ROBUSTNESS OF THE PROPOSED APPROACH TO MODERATE SPATIAL VARIABILITY

We have generated two synthetic datasets in which the bottom left-hand size of the images are affected by moderate spatial variability (i.e., variability within each image of the sequence) and outliers (see Fig. 39). An example of the spatial variability considered in this dataset is reported in Fig. 40. The data have been further corrupted by an additive white Gaussian noise to ensure a resulting SNR between 30 and 35 dB. In the following paragraphs, only the results obtained for the mixtures composed of $R = 3$ endmembers are commented (Table VI and Figs. 41 to 44). Similar comments can be made for the case $R = 9$.

- **Endmember estimation:** the proposed method shows an interesting robustness with respect to spatially sparse outliers and moderate spatial variability, in the sense that the estimated signatures (Figs. 44p to 44r) are very close to the corresponding ground truth (Fig. 40). In comparison, the shape of the endmembers recovered by VCA, SISAL and RLMM and the variability extracted by OU are significantly affected by outliers, as exemplified in Figs. 44e, 44h, 44k and 44n respectively. These qualitative results are confirmed by the quantitative performance measures reported in Table VI.
- **Abundance estimation:** the abundance maps estimated by FCLS, RLMM and SISAL reflect the high sensitivity of VCA (used to initialize SISAL and RLMM) to the presence of outliers (see the figures delineated in red in Figs. 41 to 43). On the contrary, the abundances recovered by OU and the proposed approach are much closer to the ground truth. These observations are confirmed by the abundance estimation performance reported in Table VI.
- **Overall performance:** the performance measures reported in Table VI are globally favorable to the proposed approach. It is important to mention that the price to pay with the good performance of the proposed method is its computational complexity, which is common with MCMC methods.

We finally observed that spatial variability has an influence on the outlier maps estimated by the algorithm (Fig. 39): as naturally expected, more pixels tend to be detected as corrupted by outliers. Indeed, for a given time instant t, spatial variability induces local deviations from the linear model $(\mathbf{M} + \mathbf{dM}_t)\mathbf{A}_t$, which are then captured as outliers or noise by the proposed model. Note however that this phenomenon does not significantly affect the recovered abundances and endmembers, as illustrated in Figs. 41-43 and 44.

TABLE VI

SIMULATION RESULTS ON SYNTHETIC DATA AFFECTED BY MODERATE SPATIAL VARIABILITY (aSAM($\mathbf{M}$) IN (°), GMSE($\mathbf{A}$)$\times 10^{-2}$, GMSE($\mathbf{dM}$)$\times 10^{-4}$, RE $\times 10^{-4}$, TIME IN (S)).

		aSAM($\mathbf{M}$)	GMSE($\mathbf{A}$)	GMSE($\mathbf{dM}$)	RE	time
$R=3$	VCA/FCLS	4.83	2.10	/	2.39	**1**
	SISAL/FCLS	4.71	1.55	/	0.82	2
	rLMM	3.92	1.68	/	**0.22**	655
	OU	1.46	0.35	3.22	0.95	113
	Proposed (MCMC)	**1.31**	**0.09**	**0.33**	0.65	2425
$R=9$	VCA/FCLS	2.93	3.12 $\times 10^{-1}$	/	20.1	**5**
	SISAL/FCLS	5.89	4.67 $\times 10^{-1}$	/	0.87	6
	rLMM	2.74	2.83 $\times 10^{-1}$	/	**0.38**	1498
	OU	2.48	**1.63** $\times 10^{-1}$	**2.81**	1.20	379
	Proposed (MCMC)	**1.60**	2.00 $\times 10^{-1}$	5.65	0.84	8926

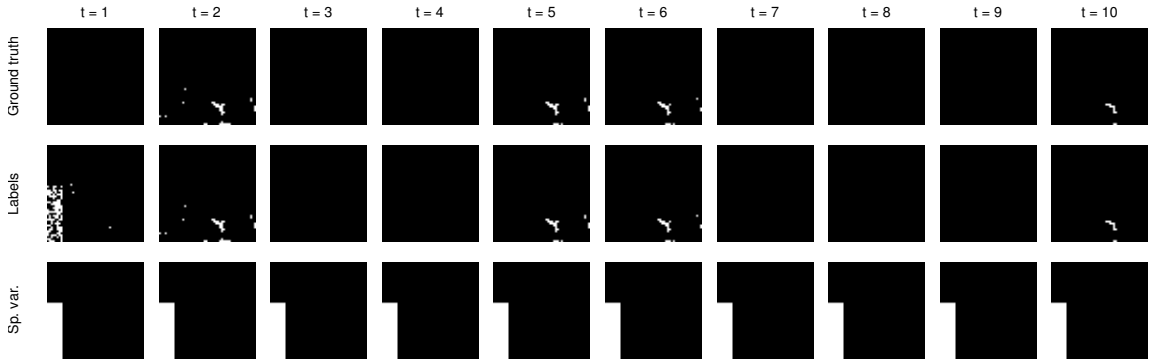
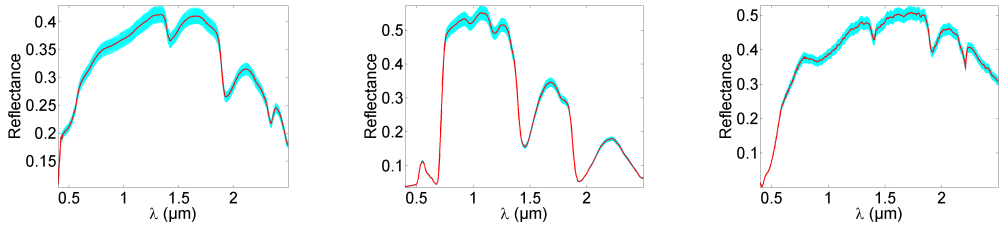


Fig. 39. mMAP estimates of the label maps recovered by the proposed approach, displayed at each time instant (the different rows correspond to the true label map where outliers appear in white, the estimated map and the areas where spatial variability is present). The influence of spatial variability can be seen on the outlier map estimated for the first image.



(a) True endmember 1 (b) True endmember 2 (c) True endmember 3

Fig. 40. True endmembers (in red) and spatial variability (in cyan) used for the synthetic dataset with $R = 3$.

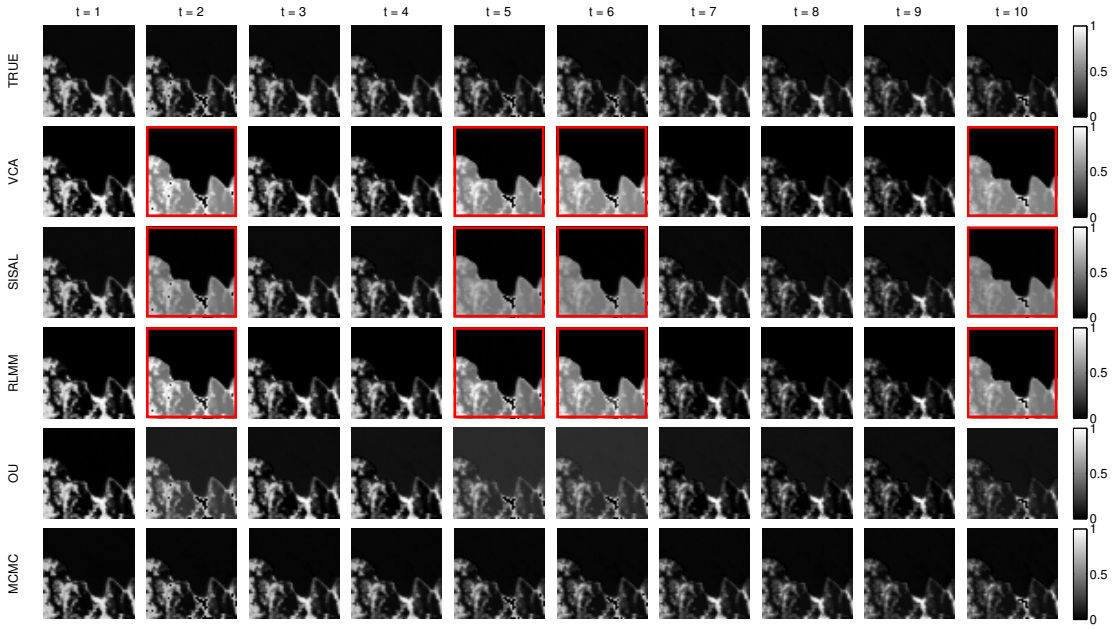


Fig. 41. Abundance map of the endmember 1 recovered by the different methods (in each row) at each time instant (given in column) for the experiment with $R = 3$ [the different rows correspond to the true abundances, VCA/FCLS, SISAL/FCLS, rLMM, OU and the proposed method]. The images delineated in red show that several methods are highly sensitive to the presence of outliers.

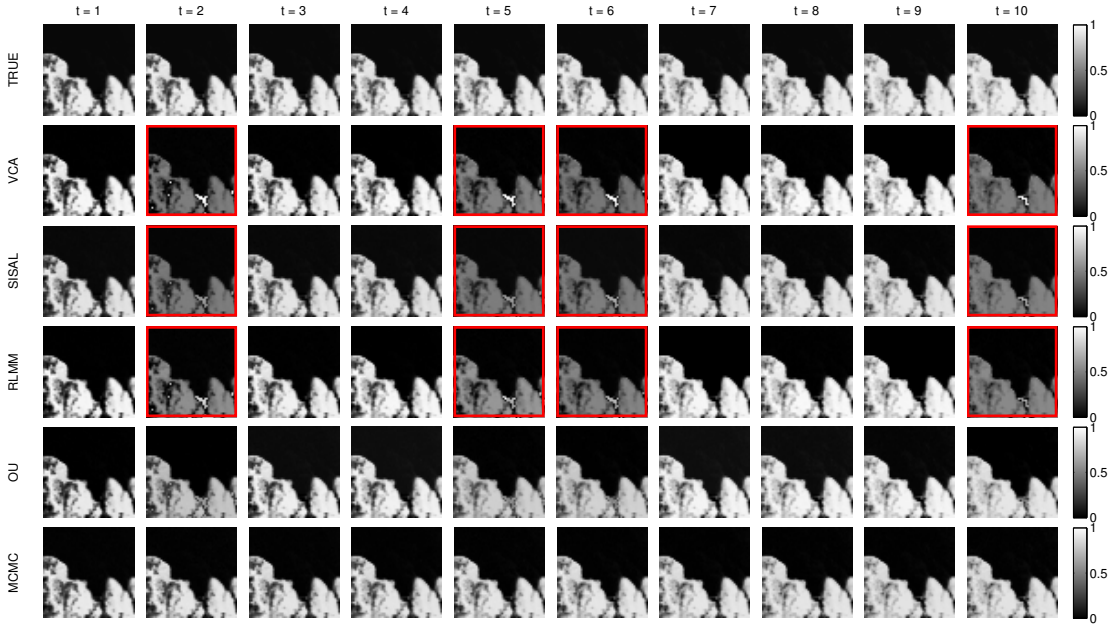


Fig. 42. Abundance map of the endmember 2 recovered by the different methods (in each row) at each time instant (given in column) for the experiment with $R = 3$ [the different rows correspond to the true abundances, VCA/FCLS, SISAL/FCLS, rLMM, OU and the proposed method]. The images delineated in red show that several methods are highly sensitive to the presence of outliers.

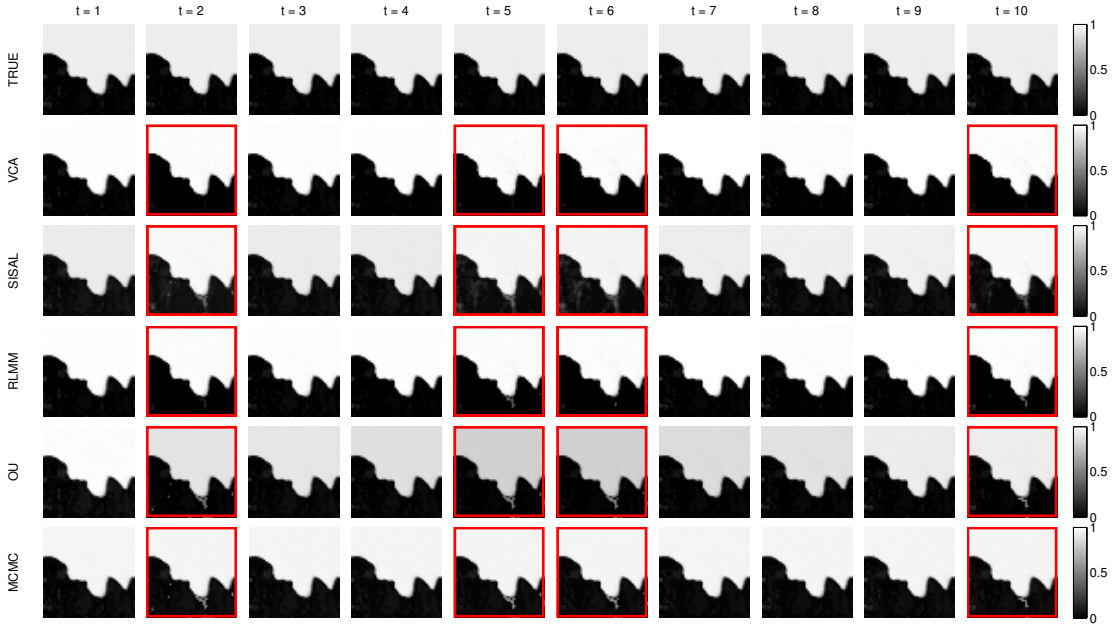
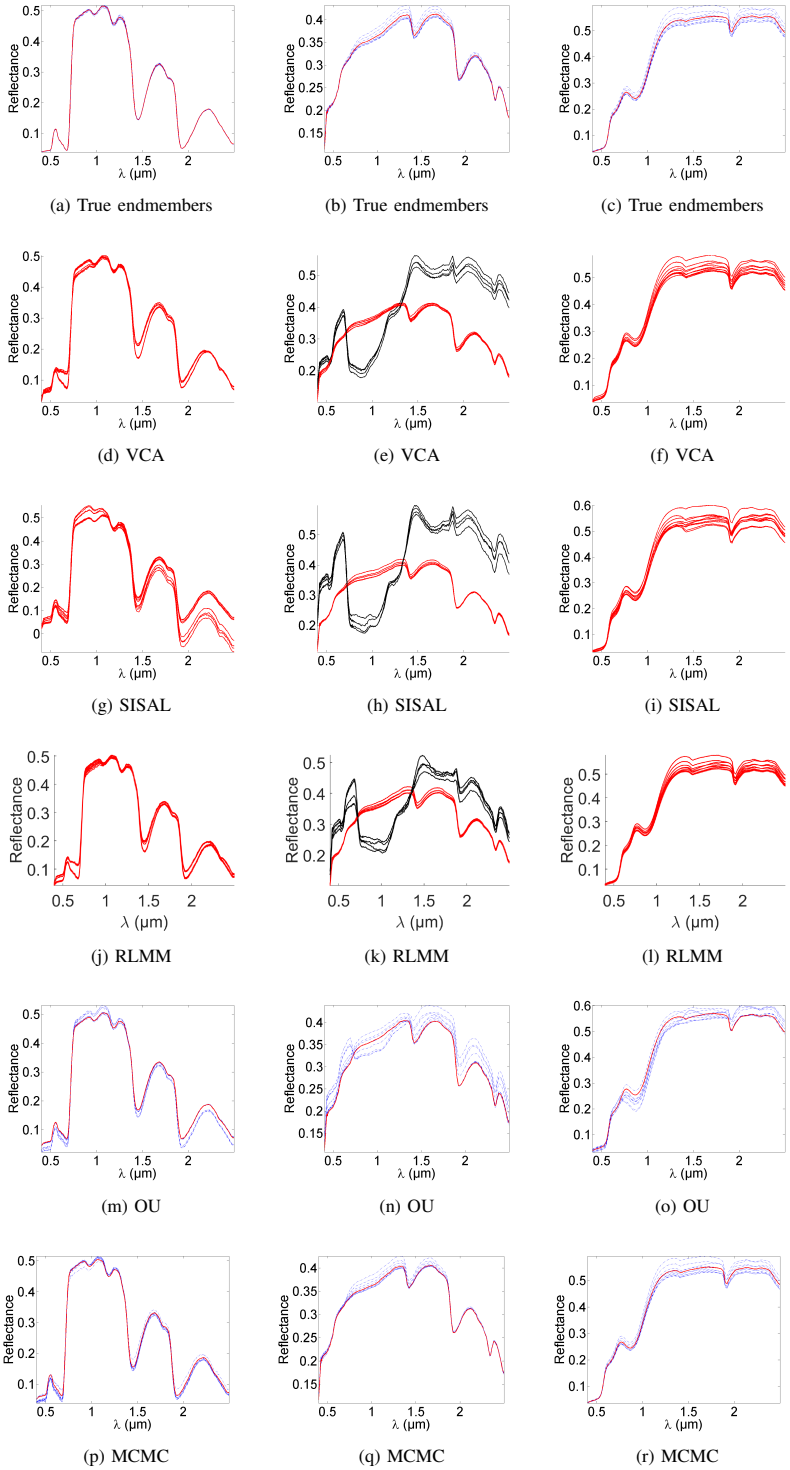


Fig. 43. Abundance map of the endmember 3 recovered by the different methods (in each row) at each time instant (given in column) for the experiment with $R = 3$ [the different rows correspond to the true abundances, VCA/FCLS, SISAL/FCLS, rLMM, OU and the proposed method]. The images delineated in red show that several methods are highly sensitive to the presence of outliers.



(a) True endmembers (b) True endmembers (c) True endmembers

(d) VCA (e) VCA (f) VCA

(g) SISAL (h) SISAL (i) SISAL

(j) RLMM (k) RLMM (l) RLMM

(m) OU (n) OU (o) OU

(p) MCMC (q) MCMC (r) MCMC

Fig. 44. Endmembers ($\mathbf{m}_r$, red lines) and their variants affected by variability ($\mathbf{m}_r + \mathbf{dm}_{r,t}$, blue dotted lines) recovered by the different methods from the synthetic dataset corrupted by spatial variability ($R = 3$). Signatures corresponding to different time instants are represented on a single figure to better appreciate the variability recovered from the data. The spectra represented in black correspond to signatures significantly affected by outliers.

REFERENCES

[1] P.-A. Thouvenin, N. Dobigeon, and J.-Y. Tourneret, "A hierarchical Bayesian model accounting for endmember variability and abrupt spectral changes to unmix multitemporal hyperspectral images," May 2017, submitted. [Online]. Available: http://thouvenin.perso.enseeiht.fr/papers/Thouvenin2017.pdf

[2] 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, Apr. 2012.

[3] N. Dobigeon, J.-Y. Tourneret, C. Richard, J. C. M. Bermudez, S. McLaughlin, and A. O. Hero, "Nonlinear unmixing of hyperspectral images: Models and algorithms," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 89–94, Jan. 2014.

[4] R. Heylen, M. Parente, and P. Gader, "A review of nonlinear hyperspectral unmixing methods," *IEEE J. Sel. Topics Appl. Earth Observ. in Remote Sens.*, vol. 7, no. 6, pp. 1844–1868, Jun. 2014.

[5] B. Somers, G. P. Asner, L. Tits, and P. Coppin, "Endmember variability in spectral mixture analysis: A review," *Remote Sens. Environment*, vol. 115, no. 7, pp. 1603–1616, Jul. 2011.

[6] A. Zare and K. C. Ho, "Endmember variability in hyperspectral imagery," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 95–104, Jan. 2014.

[7] 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, Apr. 2012.

[8] M. A. Veganzones, L. Drumetz, G. Tochon, M. D. Mura, A. Plaza, J. M. Bioucas-Dias, and J. Chanussot, "A new extended linear mixing model to address spectral variability," in *Proc. IEEE GRSS Workshop Hyperspectral Image Signal Process.: Evolution in Remote Sens. (WHISPERS)*, Lausanne, Switzerland, Jun. 2014.

[9] 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.

[10] P.-A. Thouvenin, N. Dobigeon, and J.-Y. Tourneret, "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] O. Eches, N. Dobigeon, C. Mailhes, and J.-Y. Tourneret, "Bayesian estimation of linear mixtures using the normal compositional model. Application to hyperspectral imagery," *IEEE Trans. Image Process.*, vol. 19, no. 6, pp. 1403–1413, Jun. 2010.

[12] 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, Jun. 2014.

[13] 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.

[14] A. Halimi, N. Dobigeon, J.-Y. Tourneret, S. McLaughlin, and P. Honeine, "Unmixing hyperspectral images accounting for temporal and spatial endmember variability," in *Proc. European Signal Process. Conf. (EUSIPCO)*, Nice, France, Sep. 2015, pp. 1686–1690.

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

[16] A. Halimi, C. Mailhes, and J.-Y. Tourneret, "Nonlinear regression using smooth Bayesian estimation," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*, Brisbane, Australia, Apr. 2015, pp. 2634–2638.

[17] A. Halimi, C. Mailhes, J.-Y. Tourneret, and H. Snoussi, "Bayesian estimation of smooth altimetric parameters: Application to conventional and delay/Doppler altimetry," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 4, pp. 2207–2219, Apr. 2016.

[18] 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, Jul. 2016.

[19] A. Ertürk and A. Plaza, "Informative Change Detection by Unmixing for Hyperspectral Images," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 6, pp. 1252–1256, Jun. 2015.

[20] S. Liu, L. Bruzzone, F. Bovolo, and P. Du, "Unsupervised multitemporal spectral unmixing for detecting multiple changes in hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 5, pp. 2733–2748, May 2016.

[21] 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, Jun. 2015.

[22] 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.

[23] J. M. Nascimento and J. M. Bioucas-Dias, "Vertex component analysis: a fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 898–910, Apr. 2005.

[24] 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, Jun. 2010.

[25] 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.

[26] 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.

[27] 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.

[28] R. Heylen and P. Scheunders, "A multilinear mixing model for nonlinear spectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 1, pp. 240–251, 2016.

[29] N. Dobigeon, J.-Y. Tourneret, and A. 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 Processing (ICASSP)*, Las Vegas, USA, Mar. 2008, pp. 3433–3436.

[30] V. Mazet, S. Faisan, S. Awali, M.-A. Gaveau, and L. Poisson, "Unsupervised joint decomposition of a spectroscopic signal sequence," *Signal Process.*, vol. 109, pp. 193–205, Apr. 2015.

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

[32] 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.

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

[34] 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.

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

[36] 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, Jul. 2005, pp. 811–816.

[37] C. Bazot, N. Dobigeon, and J.-Y. Tourneret, "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.

[38] L. Chaari, J.-Y. Tourneret, and C. Chaux, "Sparse signal recovery using a Bernoulli generalized Gaussian prior," in *Proc. European Signal Process. Conf. (EUSIPCO)*, Nice, France, Sep. 2015, pp. 1711–1715.

[39] O. Eches, N. Dobigeon, and J. Y. Tourneret, "Enhancing hyperspectral image unmixing with spatial correlations," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4239–4247, Nov. 2011.

[40] Y. Altmann, N. Dobigeon, S. McLaughlin, and J.-Y. Tourneret, "Residual component analysis of hyperspectral images - Application to joint nonlinear unmixing and nonlinearity detection," *IEEE Trans. Image Process.*, vol. 23, no. 5, pp. 2148–2158, Jun. 2014.

[41] J.-F. Giovannelli, "Estimation of the Ising field parameter thanks to the exact partition function," in *Proc. IEEE Int. Conf. Image Processing (ICIP)*, Hong-Kong, China, Sep. 2010, pp. 1441–1444.

[42] N. Dobigeon, S. Moussaoui, M. Coulon, J.-Y. Tourneret, and A. O. Hero, "Joint Bayesian endmember extraction and linear unmixing for hyperspectral imagery," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4355–4368, Nov. 2009.

[43] 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.

[44] J.-F. Giovannelli, "Ising field parameter estimation from incomplete and noisy data," in *Proc. IEEE Int. Conf. Image Processing (ICIP)*, Brussels, Belgium, Sep. 2011, pp. 1853–1856.

[45] A. Doucet, S. J. Godsill, and C. P. Robert, "Marginal maximum a posteriori estimation using Markov chain Monte Carlo," *Stat. Comput.*, vol. 12, no. 1, pp. 77–84, Jan. 2002.

[46] N. Dobigeon and J.-Y. Tourneret, "Efficient sampling according to a multivariate Gaussian distribution truncated on a simplex," IRIT/ENSEEIHT/TéSA, France, Tech. Rep., Mar. 2007. [Online]. Available: http://www.enseeiht.fr/~dobigeon/papers/Dobigeon_TechReport_2007b.pdf

[47] Y. Altmann, S. McLaughlin, and N. Dobigeon, "Sampling from a multivariate Gaussian distribution truncated on a simplex: a review," in *Proc. IEEE-SP Workshop Stat. and Signal Processing (SSP)*, Gold Coast, Australia, Jul. 2014, pp. 113–116.

[48] Y. Altmann, N. Dobigeon, and J.-Y. Tourneret, "Unsupervised post-nonlinear unmixing of hyperspectral images using a Hamiltonian Monte Carlo algorithm," *IEEE Trans. Image Process.*, vol. 23, no. 6, pp. 2663–2675, Jun. 2014.

[49] A. Pakman and L. Paninski, "Exact Hamiltonian Monte Carlo for truncated multivariate Gaussians," *journal of Computational and Graphical Statistics*, vol. 23, no. 2, pp. 518–542, 2014.

[50] Z. I. Botev, "The normal law under linear restrictions: Simulation and estimation via minimax tilting," *J. Roy. Stat. Soc. Ser. B*, vol. 78, no. 3, 2016, to appear.

[51] N. Chopin, "Fast simulation of truncated Gaussian distributions," *Statistics and Computing*, vol. 21, no. 2, pp. 275–288, 2011.

[52] L. Onsager, "A two-dimensional model with an order-disorder transition," *Phys. Rev.*, vol. 65, no. 3 & 4, pp. 117–149, Feb. 1944.

[53] J. Plaza, E. M. Hendrix, I. Garca, G. Martn, and A. Plaza, "On endmember identification in hyperspectral images without pure pixels: A comparison of algorithms," *J. Math. Imag. Vision*, vol. 42, pp. 163–175, Mar. 2011.

[54] A. Halimi, P. Honeine, M. Kharouf, C. Richard, and J.-Y. Tourneret, "Estimating the intrinsic dimension of hyperspectral images using a noise-whitened eigen-gap approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 7, pp. 3811–3821, Jul. 2016.

[55] P.-A. Thouvenin, N. Dobigeon, and J.-Y. Tourneret, "Online unmixing of multitemporal hyperspectral images accounting for spectral variability - complementary results and supporting material," University of Toulouse, IRIT/INP-ENSEEIHT, Tech. Rep., Oct. 2015. [Online]. Available: http://thouvenin.perso.enseeiht.fr/papers/Thouvenin_TR_2015.pdf

[56] Y. Altmann, N. Dobigeon, J.-Y. Tourneret, and J. C. M. Bermudez, "A robust test for nonlinear mixture detection in

hyperspectral images," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*, Vancouver, Canada, Jun. 2013, pp. 2149–2153.

[57] M. Pereyra, N. Dobigeon, H. Batatia, and J.-Y. Tourneret, "Computing the Cramer-Rao bound of Markov random field parameters: Application to the Ising and the Potts models," *IEEE Signal Process. Lett.*, vol. 21, no. 1, pp. 47–50, Jan. 2014.

[58] A. Gelman and D. B. Rubin, "Inference from iterative simulation using multiple sequences," *Statistical Science*, vol. 7, no. 4, pp. 457–511, 1992.

Pierre-Antoine Thouvenin (S'15–M'17) received the state engineering degree in electrical engineering from ENSEEIHT, Toulouse, France, and the M.Sc. degree in signal processing from the National Polytechnic Institute of Toulouse (INP Toulouse), both in 2014. At the time of writing this paper, he was working towards the Ph.D. degree within the Signal and Communications Group of the IRIT Laboratory, Toulouse, France. His research interests include statistical modeling, optimization techniques and hyperspectral unmixing.

Nicolas Dobigeon (S'05–M'08–SM'13) received the state engineering degree in electrical engineering from ENSEEIHT, Toulouse, France, and the M.Sc. degree in signal processing from the National Polytechnic Institute of Toulouse (INP Toulouse), both in June 2004, as well as the Ph.D. degree and Habilitation à Diriger des Recherches in Signal Processing from the INP Toulouse in 2007 and 2012, respectively. He was a Post-Doctoral Research Associate with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI, USA, from 2007 to 2008.

Since 2008, he has been with the National Polytechnic Institute of Toulouse (INP-ENSEEIHT, University of Toulouse) where he is currently a Professor. He conducts his research within the Signal and Communications Group of the IRIT Laboratory and he is also an affiliated faculty member of the Telecommunications for Space and Aeronautics (TéSA) cooperative laboratory. His current research interests include statistical signal and image processing, with a particular interest in Bayesian inverse problems with applications to remote sensing, biomedical imaging and genomics.

Jean-Yves Tourneret (SM'08) received the ingénieur degree in electrical engineering from the Ecole Nationale Supérieure d'Electronique, d'Electrotechnique, d'Informatique, d'Hydraulique et des Télécommunications (ENSEEIHT) de Toulouse in 1989 and the Ph.D. degree from the National Polytechnic Institute from Toulouse in 1992. He is currently a professor in the university of Toulouse (ENSEEIHT) and a member of the IRIT laboratory (UMR 5505 of the CNRS). His research activities are centered around statistical signal and image processing with a particular interest to Bayesian and Markov chain Monte-Carlo (MCMC) methods. He has been involved in the organization of several conferences including the European conference on signal processing EUSIPCO'02 (program chair), the international conference ICASSP'06 (plenaries), the statistical signal processing workshop SSP'12 (international liaisons), the International Workshop on Computational Advances in Multi-Sensor Adaptive Processing CAMSAP 2013 (local arrangements), the statistical signal processing workshop SSP'2014 (special sessions), the workshop on machine learning for signal processing MLSP'2014 (special sessions). He has been the general chair of the CIMI workshop on optimization and statistics in image processing hold in Toulouse in 2013 (with F. Malgouyres and D. Kouamé) and of the International Workshop on Computational Advances in Multi-Sensor Adaptive Processing CAMSAP 2015 (with P. Djuric). He has been a member of different technical committees including the Signal Processing Theory and Methods (SPTM) committee of the IEEE Signal Processing Society (2001-2007, 2010-present). He has been serving as an associate editor for the IEEE Transactions on Signal Processing (2008-2011, 2015-present) and for the EURASIP journal on Signal Processing (2013-present).