Enhancing Hyperspectral Image Unmixing With Spatial Correlations

Olivier Eches, *Member, IEEE*, Nicolas Dobigeon, *Member, IEEE*, and Jean-Yves Tourneret, *Senior Member, IEEE*

***Abstract*—This paper describes a new algorithm for hyperspectral image unmixing. Most unmixing algorithms proposed in the literature do not take into account the possible spatial correlations between the pixels. In this paper, a Bayesian model is introduced to exploit these correlations. The image to be unmixed is assumed to be partitioned into regions (or classes) where the statistical properties of the abundance coefficients are homogeneous. A Markov random field, is then proposed to model the spatial dependencies between the pixels within any class. Conditionally upon a given class, each pixel is modeled by using the classical linear mixing model with additive white Gaussian noise. For this model, the posterior distributions of the unknown parameters and hyperparameters allow the parameters of interest to be inferred. These parameters include the abundances for each pixel, the means and variances of the abundances for each class, as well as a classification map indicating the classes of all pixels in the image. To overcome the complexity of the posterior distribution, we consider a Markov chain Monte Carlo method that generates samples asymptotically distributed according the posterior. The generated samples are then used for parameter and hyperparameter estimation. The accuracy of the proposed algorithms is illustrated on synthetic and real data.**

***Index Terms*—Bayesian inference, hyperspectral images, Markov random fields (MRFs), Monte Carlo methods, spectral unmixing.**

I. Introduction

HYPERSPECTRAL image analysis involves many technical issues such as image classification, image segmentation, target detection, and the crucial step of spectral unmixing. The problem of spectral unmixing has been investigated for several decades in both the signal processing and geoscience communities, where many solutions have been proposed (see, for instance, [1] and [2] and references therein). Hyperspectral unmixing consists of decomposing the measured pixel reflectances into mixtures of pure spectra whose fractions are referred to as abundances. Assuming the image pixels are linear combinations of pure materials is very common in the unmixing framework. More precisely, the linear mixing model (LMM) considers the spectrum of a mixed pixel as a linear combination of endmembers [1]. The LMM requires to have known endmember signatures. These signatures can be obtained from a spectral library or by using an endmember extraction algorithm (EEA). Standard EEAs are reviewed in [3]. Once the endmembers that appear in a given image have been identified, the corresponding abundances have to be estimated in the so-called *inversion* step. Due to physical considerations, the abundances have to satisfy positivity and sum-to-one constraints. A lot of inversion algorithms respecting these constraints have been proposed in the literature. The fully constrained least squares (FCLS) [4] and scaled gradient [5] algorithms are two optimization techniques that ensure the positivity and sum-to-one constraints inherent to the unmixing problem. Another interesting approach introduced in [6] consists of assigning appropriate prior distributions to the abundances and to solve the unmixing problem within a Bayesian framework. However, all these inversion strategies have been developed in a pixel-by-pixel context and do not exploit the possible spatial correlations between the different pixels of the hyperspectral image. In this paper, we show that taking these spatial correlations into account allows the unmixing procedure to be improved.

Within a Bayesian estimation framework, a very popular strategy for modeling spatial information in an image is based on Markov random fields (MRFs). MRFs have been widely used in the image processing literature to properly describe neighborhood dependence between image pixels. MRFs and their pseudolikelihood approximations have been introduced by Besag in [7]. They have then been popularized by Geman in [8] by exploiting the Gibbs distribution inherent to MRFs. The hyperspectral community has also recently exploited the advantages of MRFs for hyperspectral image analysis [9]–[11]. However, to our knowledge, MRFs have not been studied for hyperspectral image unmixing. There are mainly two approaches that can be investigated to model spatial correlations between the abundances of a hyperspectral image with MRFs. The first idea is to define appropriate prior distributions for the abundances highlighting spatial correlations. This approach has been for instance adopted by Kent and Mardia in [12] where several techniques have been introduced for mixed-pixel classification of remote sensing data. These techniques rely on a fuzzy membership process, which implicitly casts the achieved classification task as a standard unmixing problem.[1] Modeling the abundance dependencies with MRFs makes this approach well adapted to unmix images with smooth abundance transition throughout the scene.

Conversely, this paper proposes to exploit the pixel correlations in an underlying membership model. This strategy allows more flexibility and appears more suited for images composed of distinct areas, as frequently encountered in remote sensing applications. Moreover, this approach has the great advantage

Manuscript received December 8, 2009; revised August 23, 2010 and January 14, 2011; accepted March 24, 2011.

The authors are with the Institut de Recherche en Informatique de Toulouse, University of Toulouse, 31071 Toulouse, France (e-mail: olivier.eches@enseeiht.fr; nicolas.dobigeon@enseeiht.fr; jean-yves.tourneret@enseeiht.fr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2011.2140119

[1]Note that, to our knowledge, Kent and Mardia's paper is one of the earliest work explicitly dealing with linear unmixing of remotely sensed images.

of easily generalizing the Bayesian algorithms previously introduced in [6] and [13], as detailed further in this paper. It consists of introducing labels that are assigned to the image pixels. To take into account the possible spatial correlations between the observed pixels, a Potts–Markov field [14] is chosen as prior for the labels. This prior enforces the neighboring pixels to belong to the same class. Potts–Markov models have been extensively used for the classification/segmentation of hyperspectral data in the remote sensing and image processing literatures [10], [11], [15], [16]. Other research works, such as [17] and [18], have proposed alternative strategies for modeling spatial correlations between pixels for the classification of hyperspectral images. All these works have shown that taking into account the spatial correlations is of real interest when analyzing hyperspectral images.

The new unmixing strategy studied in this paper assumes that the hyperspectral image to be analyzed is partitioned into homogeneous regions (or classes) in which the abundance vectors have the same first and second order statistical moments (means and covariances). This assumption implies an implicit image classification, modeled by hidden labels whose spatial dependencies follow a Potts–Markov field. Conditionally upon these labels, the abundance vectors are assigned appropriate prior distributions with unknown means and variances that depend on the pixel class. These prior distributions ensure the positivity and sum-to-one constraints of the abundance coefficients. They are based on a reparametrization of the abundance vectors and are much more flexible than the priors previously studied in [6] and [13]. Of course, the accuracy of the abundance estimation procedure depends on the hyperparameters associated with these priors. This paper proposes to estimate these hyperparameters in a fully unsupervised manner by introducing a second level of hierarchy in the Bayesian inference. Noninformative prior distributions are assigned to the hyperparameters. The unknown parameters (labels and abundance vectors) and hyperparameters (prior abundance mean and variance for each class) are then inferred from their joint posterior distribution. Since this posterior is too complex to derive closed-form expressions for the classical Bayesian estimators, Markov chain Monte Carlo (MCMC) techniques are studied to alleviate the numerical problems related to the LMM with spatial constraints. MCMC allow one to generate samples asymptotically distributed according to the joint posterior of interest. These samples are then used to approximate the Bayesian estimators, such as the minimum mean square error (MMSE) or the maximum *a posteriori* (MAP) estimators. Note that the underlying classification and abundance estimation problems are jointly solved within this Bayesian framework.

The paper is organized as follows. The unmixing problem is formulated in Section II. Section III introduces a hierarchical Bayesian model appropriate to this unmixing problem. The MCMC algorithm required to approximate the Bayesian LMM estimators is described in Section IV. Simulation results conducted on simulated and real data are provided in Sections V and VI. Finally, conclusions are reported in Section VII.

II. Problem Formulation

The LMM assumes that the spectrum of a given pixel is a linear combination of deterministic endmembers corrupted by an additive noise [1] considered here as white Gaussian. The observed L-spectrum of a given pixel p is defined as

$$\boldsymbol{y}_p = \boldsymbol{M}\boldsymbol{a}_p + \boldsymbol{n}_p \tag{1}$$

where L is the number of spectral bands, $\boldsymbol{M} = [\boldsymbol{m}_1, \ldots, \boldsymbol{m}_R]$ is a known $L \times R$ matrix containing the L-spectra of the endmembers, $\boldsymbol{a}_p$ is the $R \times 1$ abundance vector, R is the number of endmembers that are present in the image, and $\boldsymbol{n}_p$ is the noise vector. The vector $\boldsymbol{n}_p$ is classically assumed to be an independent and identically distributed (i.i.d.) zero-mean Gaussian sequence with unknown variance s^2

$$\boldsymbol{n}_p | s^2 \sim \mathcal{N}(\boldsymbol{0}_L, s^2 \boldsymbol{I}_L) \tag{2}$$

where $\boldsymbol{I}_L$ is the $L \times L$ identity matrix. Note that the noise variance is the same for all pixels of the hyperspectral image and does not vary from one pixel to another, which has been a common assumption in the hyperspectral literature [19], [20]. Considering an image of P pixels, standard matrix notations can be adopted, leading to $\boldsymbol{Y} = [\boldsymbol{y}_1, \ldots \boldsymbol{y}_P]$ and $\boldsymbol{A} = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_P]$.

A. Introducing Spatial Dependencies Between Abundances

We propose in this paper to exploit spatial correlations between the pixels of the hyperspectral image to be analyzed. More precisely, it is interesting to consider that the abundances of a given pixel are similar to the abundances of its neighbors. Formally, the hyperspectral image is assumed to be partitioned into K regions or *classes*. Let $\mathcal{I}_k \subset \{1, \ldots, P\}$ denote the subset of pixel indexes belonging to the kth class. A label vector of size $P \times 1$ denoted as $\boldsymbol{z} = [z_1, \ldots, z_P]^{\mathrm{T}}$ with $z_p \in \{1, \ldots, K\}$ is introduced to identify the class to which each pixel p belongs $(p = 1, \ldots, P)$. In other terms, $z_p = k$ if and only if $p \in \mathcal{I}_k$. In each class, the abundance vectors to be estimated are assumed to share the same first and second order statistical moments, i.e., $\forall k \in \{1, \ldots, K\},\ \forall p \in \mathcal{I}_k$

$$E[\boldsymbol{a}_p] = \boldsymbol{\mu}_k, \qquad E\left[(\boldsymbol{a}_p - \boldsymbol{\mu}_k)(\boldsymbol{a}_p - \boldsymbol{\mu}_k)^T\right] = \boldsymbol{\Lambda}_k. \tag{3}$$

Therefore, the kth class of the hyperspectral image is fully characterized by its abundance mean vector $\boldsymbol{\mu}_k$ and its abundance covariance matrix $\boldsymbol{\Lambda}_k$.

B. Markov Random Fields

To describe spatial correlations between pixels, it is important to properly define a neighborhood structure. The neighborhood relation between two pixels i and j has to be symmetric: if i is a neighbor of j, then j is a neighbor of i. This relation is applied to the nearest neighbors of the considered pixel, for example, the fourth, eighth, or twelfth nearest pixels. Fig. 1 shows two examples of neighborhood structures. The four pixel structure or *1 order neighborhood* will be considered in the rest of this paper. The associated set of neighbors, or *cliques*, has only vertical and horizontal possible configurations (see [7] and [8]).

Once the neighborhood structure has been established, the MRF can be defined. Let z_p denote a random variable associated with the pth pixel of an image (having P pixels).

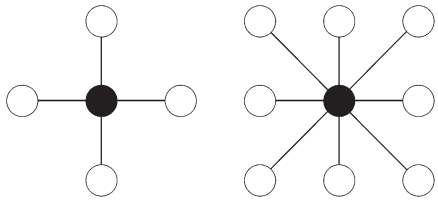

Fig. 1. 4-pixel (left) and 8-pixel (right) neighborhood structures. The considered pixel appear as a black circle whereas its neighbors are in white.

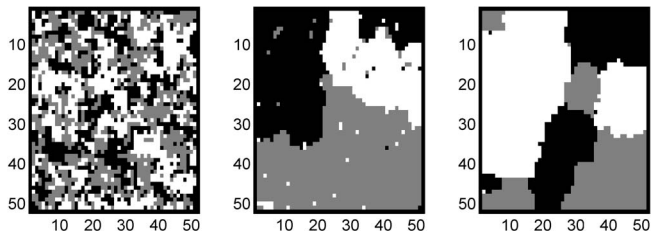

Fig. 2. Synthetic images generated from a Potts–Markov model with (from left to right) $\beta = 0.8$, 1.4, and 2.

In the context of hyperspectral image unmixing, the variables $z_1, \ldots, z_P$ indicate the pixel classes and take their values in a finite set $\{1, \ldots, K\}$, where K is the number of possible classes. The whole set of random variables $\{z_1, \ldots, z_P\}$ forms a random field. An MRF is then defined when the conditional distribution of z_i given the other pixels $\boldsymbol{z}_{-i}$ only depend on its neighbors $\boldsymbol{z}_{\mathcal{V}(i)}$, i.e.,

$$f(z_i|\boldsymbol{z}_{-i}) = f\left(z_i|\boldsymbol{z}_{\mathcal{V}(i)}\right) \tag{4}$$

where $\mathcal{V}(i)$ is the neighborhood and $\boldsymbol{z}_{-i} = \{z_j; j \neq i\}$. This paper focuses on the Potts–Markov model since it is very appropriate to hyperspectral image segmentation [11]. Given a discrete random field $\boldsymbol{z}$ attached to an image with P pixels, the Hammersley–Clifford theorem yields

$$f(\boldsymbol{z}) = \frac{1}{G(\beta)} \exp\left[\sum_{p=1}^{P} \sum_{p' \in \mathcal{V}(p)} \beta \delta(z_p - z_{p'})\right] \tag{5}$$

where β is the *granularity* coefficient, $G(\beta)$ is the normalizing constant or *partition function*, and $\delta(\cdot)$ is the Kronecker function ($\delta(x) = 1$ if $x = 0$, and $\delta(x) = 0$ if otherwise). Note that drawing a label vector $\boldsymbol{z} = [z_1, \ldots, z_P]$ from the distribution (5) can be easily achieved without knowing $G(\beta)$ by using a Gibbs sampler [21]. The hyperparameter β tunes the degree of homogeneity of each region in the image. Some simulations have been conducted to show the influence of this parameter on image homogeneity. Synthetic images have been generated from a Potts–Markov model with $K = 3$ (corresponding to three gray levels in the image) and a 1-order neighborhood structure. Fig. 2 indicates that a small value of β induces a *noisy* image with a large number of regions, contrary to a large value of β that leads to few and large homogeneous regions. It is unnecessary to consider values of $\beta \geq 2$ since, for the 1-order neighborhood structure adopted here, "*when* $\beta \geq 2$, *the Potts–Markov model is almost surely concentrated on single-color images*" [22, p. 237]. Note, however, that for larger neighborhood systems, a smaller value of β would be enough to obtain uniform patches in Potts realizations. For example, β is expected to be about twice for a 2-order neighborhood [23]. In this paper, the granularity coefficient β is fixed *a priori*.

C. Abundance Reparametrization

As explained before, the fraction vectors $\boldsymbol{a}_p$ should satisfy positivity and sum-to-one constraints defined as

$$\sum_{r=1}^{R} a_r = 1, \qquad a_r > 0, \qquad \forall r = 1, \ldots, R. \tag{6}$$

To ensure that these constraints are satisfied, we have considered a reparametrization for positive parameters summing to one that was introduced in [12] for the spectral unmixing of satellite images. Note that this reparametrization has also shown interesting results for a pharmacokinetic problem [24] and has been recently applied to hyperspectral unmixing [25]. This reparametrization consists of rewriting the abundances as a function of random variables that will be referred to as *logistic coefficients* in the rest of this paper. A logistic coefficient vector $\boldsymbol{t}_p = [t_{1,p} \ldots, t_{R,p}]^{\mathrm{T}}$ is assigned to each abundance vector $\boldsymbol{a}_p$, according to the relationship

$$a_{r,p} = \frac{\exp(t_{r,p})}{\sum_{r=1}^{R} \exp(t_{r,p})}. \tag{7}$$

Initially, the spatial dependencies resulting from the image partitioning described in Section II-A are based on the first and second order moments of the abundance vectors $\boldsymbol{a}_p$. However, the spatial constraints defined in (3) can be easily adapted when using logistic coefficient vectors. Indeed, in each class, the unknown logistic coefficient vectors are assumed to share the same first and second order moments, i.e., $\forall k \in \{1, \ldots, K\}$, $\forall p \in \mathcal{I}_k$

$$\begin{aligned} \boldsymbol{\psi}_k &= E[\boldsymbol{t}_p | z_p = k] \\ \boldsymbol{\Sigma}_k &= E\left[(\boldsymbol{t}_p - \boldsymbol{\psi}_k)(\boldsymbol{t}_p - \boldsymbol{\psi}_k)^{\mathrm{T}}\right]. \end{aligned} \tag{8}$$

With this reparametrization, the kth class is fully characterized by the unknown hyperparameters $\boldsymbol{\psi}_k$ and $\boldsymbol{\Sigma}_k$.

III. Hierarchical Bayesian Model

This section investigates the likelihood and the priors inherent to the LMM for the spectral unmixing of hyperspectral images, based on Potts–Markov random fields and logistic coefficients. The unknown parameter vector associated with the proposed Bayesian model is denoted as $\boldsymbol{\Upsilon} = \{\boldsymbol{T}, \boldsymbol{z}, s^2\}$, where s^2 is the noise variance, $\boldsymbol{z}$ is the label vector, and $\boldsymbol{T} = [\boldsymbol{t}_1, \ldots, \boldsymbol{t}_P]$ with $\boldsymbol{t}_p = [t_{1,p}, \ldots, t_{R,p}]^{\mathrm{T}} (p = 1, \ldots, P)$ is the logistic coefficient matrix. Note that the noise variance s^2 is unknown in this paper, contrary to the model considered in [12].

A. Likelihood

The additive white Gaussian noise sequence of the LMM allows one to write[2] $\boldsymbol{y}_p|\boldsymbol{t}_p, s^2 \sim \mathcal{N}(\boldsymbol{M}\boldsymbol{a}_p(\boldsymbol{t}_p), s^2\boldsymbol{I}_L)(p = 1, \ldots, P)$. Therefore, the likelihood function of $\boldsymbol{y}_p$ is

$$f(\boldsymbol{y}_p|\boldsymbol{t}_p, s^2) \propto \frac{1}{s^L} \exp\left[-\frac{\left\|\boldsymbol{y}_p - \boldsymbol{M}\boldsymbol{a}_p(\boldsymbol{t}_p)\right\|^2}{2s^2}\right] \tag{9}$$

where $\propto$ means proportional to and $\|\boldsymbol{x}\| = \sqrt{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}}$ is the standard ℓ^2 norm. By assuming independence between the noise sequences $\boldsymbol{n}_p(p = 1, \ldots, P)$, the likelihood of the P image pixels is

$$f(\boldsymbol{Y}|\boldsymbol{T}, s^2) = \prod_{p=1}^{P} f(\boldsymbol{y}_p|\boldsymbol{t}_p, s^2). \tag{10}$$

B. Parameter Priors

This section defines the prior distributions of the unknown parameters and their associated hyperparameters that will be used for the LMM.

1) Label Prior: The prior distribution for the label vector $\boldsymbol{z} = [z_1, \ldots, z_P]^{\mathrm{T}}$ introduced in Section II-B is a Potts–Markov random field with a 1-order neighborhood and a known granularity coefficient β. The resulting prior distribution can be written as in (5), where $\mathcal{V}(p)$ is the 1-order neighborhood shown in Fig. 1 (left).

2) Logistic Coefficient Prior: Following the approach described in Section II-A, each component of $\boldsymbol{t}_p$ is assumed to be distributed according to a Gaussian distribution. In addition, as highlighted in Section II-C [see (8)], the mean and variance of the logistic coefficients depend on the class to which the corresponding pixel belongs. Therefore, the prior distribution for $\boldsymbol{t}_p$ is defined conditionally upon the pixel label

$$t_{r,p}|z_p = k, \psi_{r,k}, \sigma^2_{r,k} \sim \mathcal{N}\left(\psi_{r,k}, \sigma^2_{r,k}\right) \tag{11}$$

where the hyperparameters $\psi_{r,k}$ and $\sigma^2_{r,k}$ depend on their pixel class k. As suggested in Section I, a hierarchical Bayesian algorithm will be used to estimate these hyperparameters. For a given pixel p, by assuming prior independence between the coefficients $t_{1,p}, \ldots, t_{R,p}$, the prior distribution for the vector $\boldsymbol{t} = [t_{1,p}, \ldots, t_{R,p}]^{\mathrm{T}}$ is

$$f(\boldsymbol{t}_p|z_p = k, \boldsymbol{\psi}_k, \boldsymbol{\Sigma}_k) \sim \mathcal{N}(\boldsymbol{\psi}_k, \boldsymbol{\Sigma}_k) \tag{12}$$

where $\boldsymbol{\psi}_k = [\psi_{1,k}, \ldots, \psi_{R,k}]^{\mathrm{T}}$ and $\boldsymbol{\Sigma}_k = \mathrm{diag}(\sigma^2_{r,k})$ is the $R \times R$ diagonal matrix whose diagonal elements are $\sigma^2_{r,k}$.

By assuming prior independence between the P vectors $\boldsymbol{t}_1, \ldots, \boldsymbol{t}_P$, the full posterior distribution for the logistic coefficient matrix $\boldsymbol{T}$ is

$$f(\boldsymbol{T}|\boldsymbol{z}, \boldsymbol{\Psi}, \boldsymbol{\Sigma}) = \prod_{k=1}^{K} \prod_{p \in \mathcal{I}_k} f(\boldsymbol{t}_p|z_p = k, \boldsymbol{\psi}_k, \boldsymbol{\Sigma}_k) \tag{13}$$

with $\boldsymbol{\Psi} = [\boldsymbol{\psi}_1, \ldots, \boldsymbol{\psi}_K]$ and $\boldsymbol{\Sigma} = \{\boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_K\}$.

[2]Note that the dependence of the abundance vector $\boldsymbol{a}_p$ on the logistic coefficient vector $\boldsymbol{t}_p$ through (7) has been explicitly mentioned by denoting $\boldsymbol{a}_p = \boldsymbol{a}_p(\boldsymbol{t}_p)$.

3) Noise Variance Prior: A conjugate inverse-gamma distribution is assigned to the noise variance

$$s^2|\nu, \delta \sim \mathcal{IG}(\nu, \delta) \tag{14}$$

where ν and δ are adjustable hyperparameters. This paper assumes that $\nu = 1$ (as in [6]) and estimates δ jointly with the other unknown parameters and hyperparameters (using a hierarchical Bayesian algorithm).

C. Hyperparameter Priors

Hierarchical Bayesian algorithms require one to define prior distributions for the hyperparameters. Note that the hyperparameters $\psi_{r,k}$ and $\sigma^2_{r,k}$ fully describe the different classes partitioning the image. The prior distributions for $\psi_{r,k}$ and $\sigma^2_{r,k}$ are conjugate distributions. More precisely, a vague inverse-gamma distribution is chosen for the logistic coefficient variance $\sigma^2_{r,k}$, i.e.,

$$\sigma^2_{r,k}|\xi, \gamma \sim \mathcal{IG}(\xi, \gamma) \tag{15}$$

where $\xi = 1$ and $\gamma = 5$ (in order to obtain a large variance). Moreover, a centered Gaussian distribution with unknown variance has been chosen as prior for the logistic coefficient mean

$$\psi_{r,k}|\upsilon^2 \sim \mathcal{N}(0, \upsilon^2) \tag{16}$$

where υ^2 is another adjustable hyperparameter. By assuming independence between the different mean vectors $\boldsymbol{\psi}_k$, as well as between the covariance matrices $\boldsymbol{\Sigma}_k$ for $k = 1, \ldots, K$, the priors for $\boldsymbol{\Psi}$ and $\boldsymbol{\Sigma}$ can be expressed as

$$f(\boldsymbol{\Psi}|\upsilon^2) \propto \prod_{k=1}^{K} \prod_{r=1}^{R} \left(\frac{1}{\upsilon^2}\right)^{\frac{1}{2}} \exp\left(-\frac{\psi^2_{r,k}}{2\upsilon^2}\right) \tag{17}$$

$$f(\boldsymbol{\Sigma}|\xi, \gamma) \propto \prod_{k=1}^{K} \prod_{r=1}^{R} \frac{\gamma^{\xi}}{\Gamma(\xi)} \left(\sigma^2_{r,k}\right)^{-(\xi+1)} \exp\left(-\frac{\gamma}{\sigma^2_{r,k}}\right). \tag{18}$$

Jeffreys' priors are chosen for the hyperparameters δ and υ^2 (see, e.g., [26, p. 131] for the details including computations)

$$f(\delta) \propto \frac{1}{\delta}\mathbf{1}_{\mathbb{R}^+}(\delta) \qquad f(\upsilon^2) \propto \frac{1}{\upsilon^2}\mathbf{1}_{\mathbb{R}^+}(\upsilon^2) \tag{19}$$

where $\mathbf{1}_{\mathbb{R}^+}(\cdot)$ denotes the indicator function defined on $\mathbb{R}^+$. These choices reflect the lack of knowledge regarding these two hyperparameters. At this last hierarchy level within the Bayesian inference, the hyperparameter vector can be defined as $\boldsymbol{\Omega} = \{\boldsymbol{\Psi}, \boldsymbol{\Sigma}, \upsilon^2, \delta\}$.

D. Joint Distribution

The joint posterior of the unknown parameter vector $\boldsymbol{\Theta} = (\boldsymbol{\Upsilon}, \boldsymbol{\Omega})$ is classically defined using the hierarchical structure

$$f(\boldsymbol{\Theta}|\boldsymbol{Y}) = f(\boldsymbol{Y}|\boldsymbol{\Upsilon})f(\boldsymbol{\Upsilon}|\boldsymbol{\Omega})f(\boldsymbol{\Omega}) \tag{20}$$

i.e., $f(\boldsymbol{\Theta}|\boldsymbol{Y})$ is proportional to the product of the likelihood by the priors and hyperpriors. The posterior distribution (20)

associated with the LMM is too complex to obtain closed-form expressions for the MMSE or MAP estimators of the unknown parameter vector $\boldsymbol{\Theta}$. To alleviate this problem, we propose to use MCMC methods to generate samples that are asymptotically distributed according to (20). The generated samples are then used to approximate the Bayesian estimators. The next section studies a hybrid Gibbs sampler that generates samples asymptotically distributed according to the posterior distribution (20).

IV. Hybrid Gibbs Sampler

This section studies a Metropolis-within-Gibbs sampler that generates samples according to $f(\boldsymbol{\Theta}|\boldsymbol{Y})$. The proposed sampler iteratively generates samples distributed according to the conditional distributions detailed below.

A. Conditional Distribution of the Label Vector $\boldsymbol{z}$

For each pixel p ($p = 1, \ldots, P$), the class label z_p is a discrete random variable whose conditional distribution is fully characterized by the probabilities

$$P\left[z_p = k|\boldsymbol{\Theta}_{-z_p}\right] \propto f(\boldsymbol{t}_p|z_p = k, \boldsymbol{\psi}_k, \boldsymbol{\Sigma}_k) f(z_p|\boldsymbol{z}_{-p})$$

where $\boldsymbol{\Theta}_{-z_p}$ denotes $\boldsymbol{\Theta}$ without z_p, $k = 1, \ldots, K$ (K is the number of classes), and $\boldsymbol{z}_{-p}$ denotes the vector $\boldsymbol{z}$ whose pth element has been removed. These posterior probabilities can be expressed as

$$\begin{aligned}\mathrm{P}\left[z_p = k|\boldsymbol{\Theta}_{-z_p}\right] \propto \exp\left[\sum_{p=1}^{P}\sum_{p'\in\mathcal{V}(p)} \beta\delta(z_p - z_{p'})\right] \\ \times|\boldsymbol{\Sigma}_k|^{-1/2}\exp\left[-\frac{1}{2}(\boldsymbol{t}_p - \boldsymbol{\psi}_k)^{\mathrm{T}}\boldsymbol{\Sigma}_k^{-1}(\boldsymbol{t}_p - \boldsymbol{\psi}_k)\right]\end{aligned} \tag{21}$$

where $|\boldsymbol{\Sigma}_k| = \prod_{r=1}^{R}\sigma_{r,k}^2$. Note that the posterior probabilities of the label vector $\boldsymbol{z}$ in (21) define an MRF. Consequently, sampling from this conditional distribution can be achieved using the scheme detailed in [21], i.e., by drawing a discrete value in the finite set $\{1, \ldots, K\}$ with the probabilities (21).

B. Conditional Distribution of Logistic Coefficient Matrix $\boldsymbol{T}$

For each pixel p, the Bayes theorem yields

$$\begin{aligned}&f\left(\boldsymbol{t}_p|z_p = k, \boldsymbol{\psi}_k, \boldsymbol{\Sigma}_k, \boldsymbol{y}_p, s^2\right) \\ &\quad\propto f(\boldsymbol{y}_p|\boldsymbol{t}_p, s^2) f(\boldsymbol{t}_p|z_p = k, \boldsymbol{\psi}_k, \boldsymbol{\Sigma}_k) \\ &\quad\propto \left(\frac{1}{s^2}\right)^{\frac{L}{2}}\exp\left\{-\frac{1}{2s^2}\left\|\boldsymbol{y}_p - \boldsymbol{M}\boldsymbol{a}_p(\boldsymbol{t}_p)\right\|^2\right\} \\ &\qquad\times|\boldsymbol{\Sigma}_k|^{-\frac{1}{2}}\exp\left[-\frac{1}{2}(\boldsymbol{t}_p - \boldsymbol{\psi}_k)^{\mathrm{T}}\boldsymbol{\Sigma}_k^{-1}(\boldsymbol{t}_p - \boldsymbol{\psi}_k)\right].\end{aligned} \tag{22}$$

Unfortunately, it is too difficult to generate samples distributed according to (22). Therefore, a Metropolis–Hastings step is used, based on a random walk method [27, p. 245] with a Gaussian proposal distribution $\mathcal{N}(0, u_r^2)$. The variance u_r^2 of the instrumental distribution has been fixed to obtain an acceptance rate between 0.15 and 0.5 as recommended in [28].

C. Conditional Distributions of the Noise Variance

The Bayes theorem yields

$$f(s^2|\boldsymbol{Y}, \boldsymbol{T}, \delta) \propto f(s^2|\delta)\prod_{p=1}^{P} f(\boldsymbol{y}_p|\boldsymbol{t}_p, s^2).$$

As a consequence, $s^2|\boldsymbol{Y}, \boldsymbol{T}, \delta$ is distributed according to the following inverse-gamma distribution

$$s^2|\boldsymbol{Y}, \boldsymbol{T}, \delta \sim \mathcal{IG}\left(\frac{LP}{2} + 1, \delta + \sum_{p=1}^{P}\frac{\left\|\boldsymbol{y}_p - \boldsymbol{M}\boldsymbol{a}_p(\boldsymbol{t}_p)\right\|^2}{2}\right).$$

D. Conditional Distribution of $\boldsymbol{\Psi}$ and $\boldsymbol{\Sigma}$

For each endmember r ($r = 1, \ldots, R$) and each class k ($k = 1, \ldots, K$), the conditional distribution of $\psi_{r,k}$ can be written as

$$f\left(\psi_{r,k}|\boldsymbol{\Theta}_{-\psi_{r,k}}\right) \propto f(\psi_{r,k}|v^2)\prod_{p\in\mathcal{I}_k} f\left(t_{r,p}|z_p = k, \psi_{r,k}, \sigma_{r,k}^2\right).$$

Similarly, the conditional distribution of $\sigma_{r,k}^2$ is

$$f\left(\sigma_{r,k}^2|\boldsymbol{\Theta}_{-\sigma_{r,k}^2}\right) \propto f\left(\sigma_{r,k}^2\right)\prod_{p\in\mathcal{I}_k} f\left(t_{r,p}|z_p = k, \psi_{r,k}, \sigma_{r,k}^2\right).$$

Straightforward computations allow the following results to be obtained

$$\psi_{r,k}|\boldsymbol{\Theta}_{-\psi_{r,k}} \sim \mathcal{N}\left(\frac{v^2 n_k \bar{t}_{r,k}}{\sigma_{r,k}^2 + v^2 n_k}, \frac{v^2\sigma_{r,k}^2}{\sigma_{r,k}^2 + v^2 n_k}\right)$$

$$z_{r,k}|\boldsymbol{\Theta}_{-\sigma_{r,k}^2} \sim \mathcal{IG}\left(\frac{n_k}{2} + 1, \gamma + \sum_{p\in\mathcal{I}_k}\frac{(t_{r,p} - \psi_{r,k})^2}{2}\right)$$

with $\bar{t}_{r,k} = (1/n_k)\sum_{p\in\mathcal{I}_k} t_{r,p}$

E. Conditional Distribution of v^2 and δ

The conditional distributions of v^2 and δ are the following inverse-gamma and gamma distributions, respectively

$$v^2|\boldsymbol{\Psi} \sim \mathcal{IG}\left(\frac{RK}{2}, \frac{1}{2}\sum_{k=1}^{K}\boldsymbol{\psi}_k^{\mathrm{T}}\boldsymbol{\psi}_k\right) \qquad \delta|s^2 \sim \mathcal{G}\left(1, \frac{1}{s^2}\right).$$

V. Simulation Results on Synthetic Data

Many simulations have been conducted to illustrate the accuracy of the proposed algorithm. The first experiment considers a 25 × 25 synthetic image with $K = 3$ different classes. The image contains $R = 3$ mixed components (construction concrete, green grass, and micaceous loam) whose spectra

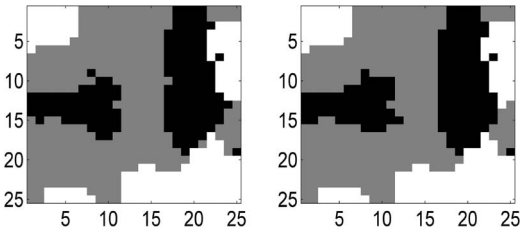

Fig. 3. Left: actual labels. Right: estimated labels.

TABLE I
ESTIMATION OF ABUNDANCE MEANS AND VARIANCES ($\times 10^{-3}$)

		Actual	Estimated
Class 1	$\boldsymbol{\mu}_1$	$[0.6, 0.3, 0.1]^T$	$[0.57, 0.3, 0.13]^T$
	$\text{diag}(\boldsymbol{\Lambda}_1)$	$[5, 5, 5]^T$	$[5.6, 6.7, 6.7]^T$
Class 2	$\boldsymbol{\mu}_2$	$[0.3, 0.5, 0.2]^T$	$[0.29, 0.49, 0.2]^T$
	$\text{diag}(\boldsymbol{\Lambda}_2)$	$[5, 5, 5]^T$	$[4.5, 5.2, 8.1]^T$
Class 3	$\boldsymbol{\mu}_3$	$[0.3, 0.2, 0.5]^T$	$[0.3, 0.2, 0.5]^T$
	$\text{diag}(\boldsymbol{\Lambda}_3)$	$[5, 5, 5]^T$	$[4.6, 5.7, 10.2]^T$

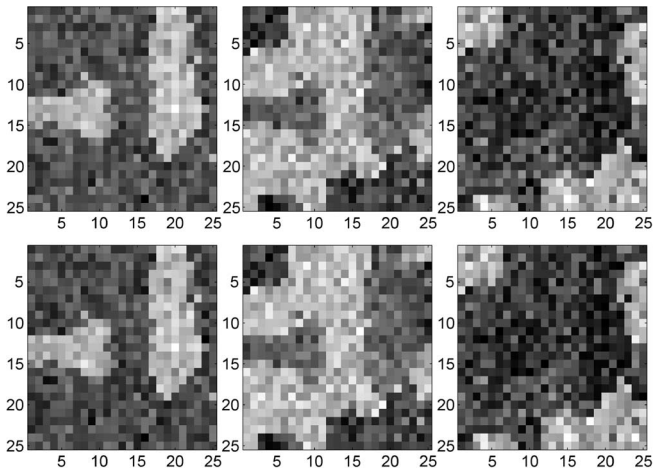

Fig. 4. Top: abundance maps of the 3 pure materials for LMM. Bottom: abundance maps of the 3 pure materials estimated by the hybrid Gibbs sampler (left to right: construction concrete, green grass, micaceous loam).

($L = 413$ spectral bands) have been extracted from the spectral libraries distributed with the ENVI package [29]. A label map generated using (5) with $\beta = 1.1$ is shown in Fig. 3 (left). The mean and variance of the abundances have been chosen for each class, as reported in Table I. These values reflect the fact that the first endmember is more present in Class 1 (with an average concentration of 60%), the second endmember is more present in Class 2 (with an average concentration of 50%), and the third endmember is more present in Class 3 (with an average concentration of 50%). In this simulation scenario, the abundance variance has been fixed to a common value of 0.005 for all endmembers, pixels, and classes. The generated abundance maps for the LMM are depicted in Fig. 4. Note that a white (black) pixel in the fraction map indicates a large (small) value of the abundance coefficient. The noise variance is chosen such that the average signal-to-noise ratio (SNR) is equal to $SNR = 19$ dB, i.e., $s^2 = 0.001$.

The MMSE and MAP estimators for the unknown parameters can be computed from samples generated with the Gibbs samplers presented in Section IV. For instance, the marginal MAP estimates of the label vector $\hat{\boldsymbol{z}}_{\text{MAP}}$ are shown in Fig. 3 (right) for the proposed hybrid Gibbs algorithm. The MMSE estimates of the abundances conditioned upon $\hat{\boldsymbol{z}}_{\text{MAP}}$ are shown in Fig. 4. A number of $N_{\text{MC}} = 5000$ iterations (including 500 burn-in iterations) have been necessary to obtain these results. The proposed algorithm generates samples distributed according to the full posterior of interest. Then, these samples can be used to compute the posterior distributions of the parameters of interest. For instance, the estimated abundance means and variances have been reported in Table I (last row). The estimated classes, abundance coefficients, and abundance mean vectors are clearly in accordance with their actual values.

TABLE II
GLOBAL mses OF EACH ABUNDANCE COMPONENT

	FCLS	Bayesian	Spatial
MSE_1^2	0.0019	0.0016	3.1×10^{-4}
MSE_2^2	4.3×10^{-4}	4.1×10^{-4}	8.98×10^{-5}
MSE_3^2	0.0014	0.0013	2.35×10^{-4}

The LMM hybrid Gibbs algorithm is compared respectively with its non-spatial constrained Bayesian counterpart developed in [6]. The synthetic image shown in Fig. 3 has been analyzed by the initial algorithm in [6] with the same number of iterations N_{MC} in addition to the FCLS algorithm [4]. As a criterion, the global MSE of the rth estimated abundances has been computed for each algorithm. This global MSE is defined as

$$\text{MSE}_r^2 = \frac{1}{P}\sum_{p=1}^{P}(\hat{a}_{r,p} - a_{r,p})^2 \tag{23}$$

where $\hat{a}_{r,p}$ denotes the MMSE estimate of the abundance $a_{r,p}$. Table II reports the different results, showing that the algorithm developed in this paper (referred to as "Spatial") performs better than the non-spatial constrained algorithms (referred to as "Bayesian" and "FCLS").

VI. SIMULATION RESULTS ON AVIRIS IMAGES

A. Performance of the Proposed Algorithm

This section illustrates the performance of the proposed spatial algorithm on a real hyperspectral dataset, acquired over Moffett Field (CA, USA) in 1997 by the JPL spectro-imager Airborne Visible/Infrared Imaging Spectrometer (AVIRIS). Many previous works have used this image to illustrate and compare algorithm performances with hyperspectral images [30]. The first region of interest, shown in Fig. 5, is a 50×50 image. The dataset has been reduced from the original 224 bands to $L = 189$ bands by removing water absorption bands. As in [6], a principal component analysis has been conducted as a preprocessing step to determine the number of endmembers present in the scene. Then, the endmember spectra have been extracted with the help of N-FINDR algorithm proposed by Winter in [31]. The $R = 3$ extracted endmembers, shown in Fig. 6, corresponds to soil, vegetation and water.[3] The

[3]Note that the influence of the endmember extraction step on the unmixing results has been investigated in [21] by coupling the proposed algorithm with other EEAs.

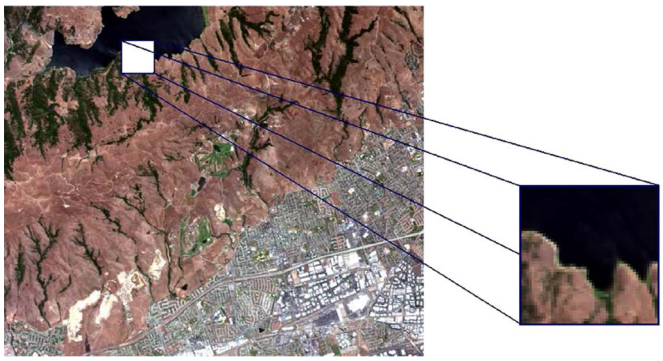

Fig. 5. Real hyperspectral data: Moffett field acquired by AVIRIS in 1997 (left) and the region of interest shown in true colors (right).

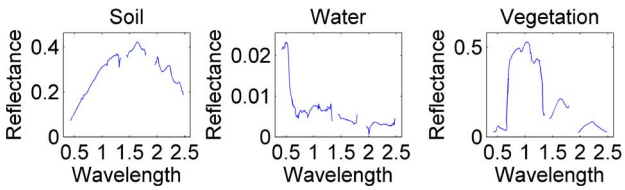

Fig. 6. $R = 3$ endmember spectra obtained by the N-FINDR algorithm.

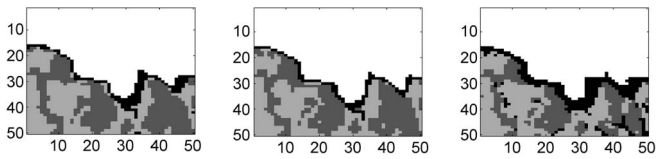

Fig. 7. Label map estimated by the LMM-based proposed algorithm for $R = 3$ (left), $R = 4$ (middle) and $R = 5$ (right).

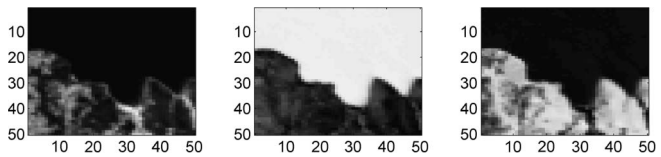

Fig. 8. Abundance maps estimated by the proposed algorithm (from left to right: vegetation, water and soil).

algorithm proposed in Section IV has been applied on this image with $N_{\mathrm{MC}} = 5000$ iterations (with 500 burn-in iterations). The number of classes has been set to $K = 4$ since prior knowledge on the scene allows 4 areas in the image to be identified: water point, lake shore, vegetation and soil.

The estimated classification and abundance maps for the proposed hybrid Gibbs algorithm are shown in Figs. 7 (left) and 8. The results are very similar to those obtained with the LMM-based Bayesian algorithm (see [6, Fig. 9 (top)]) or the FCLS algorithm (see [6, Fig. 9 (bottom)]).

The performance of the proposed algorithm has been also evaluated for different values of the number of endmembers R. The resulting classification maps for $R = 4$ and $R = 5$ are given in Fig. 7 (middle and right). These maps show that the classification results are quite robust with respect to the number of endmembers. The corresponding abundance maps can be found in [21], as well as the results of the proposed algorithm when the number of classes varies.

The execution time of the proposed method (combined with the N-FINDR procedure) has been compared with those of

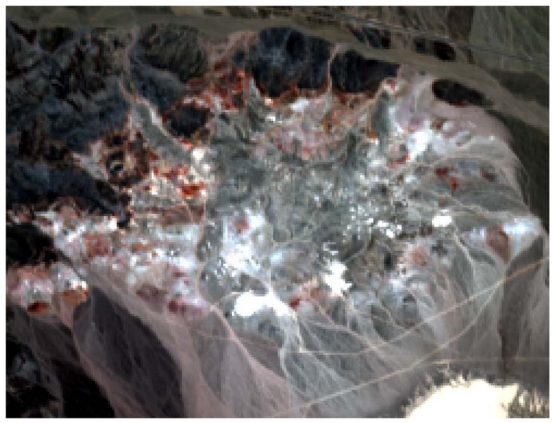

Fig. 9. AVIRIS image of 190 $\times$ 250 pixels extracted from Cuprite scene observed in composite natural colors.

TABLE III
COMPUTATIONAL TIMES OF LMM-BASED UNMIXING ALGORITHMS

	FCLS	cNMF	Spatial
Times (s.)	0.388	2.5×10^3	8.4×10^3

two other unmixing algorithms when applied on this Moffett image: the FCLS algorithm (combined with N-FINDR) and the constrained nonnegative matrix factorization (cNMF) algorithm that jointly estimates the endmember matrix and the abundances [32]. The results are reported in Table III.[4] The proposed method (referred to as "Spatial") has the higher computational cost, mainly due to the joint estimation of the labels and the abundance vectors. However, it provides more information about unmixing. In particular, the samples generated by the proposed Gibbs sampler can be used to determine confidence intervals for the estimated parameters.

B. Simulation on a Larger Image

The performance of the proposed Bayesian algorithm has also been evaluated on a larger real hyperspectral image. The selected scene has been extracted from the AVIRIS Cuprite image, acquired over a mining site in Nevada in 1997. The geologic characteristics of the complete data have been mapped in [33]. The area of interest of size 190 $\times$ 250 is shown in Fig. 9 and has been previously studied in [34] to test the VCA algorithm with $R = 14$. Therefore, in this experiment, the same number of endmembers has been extracted by the VCA algorithm. The number of classes has been set to $K = 14$, which seems to be sufficient to capture the natural diversity of the scene. The proposed algorithm has been used to estimate the unknown parameters related to the analyzed scene. As an example, Fig. 10 shows the estimated classification map (see [21] for more results). The proposed Bayesian inversion algorithm has been able to identify some regions similar to those recovered in [34]. For instance, the composition of two areas (marked as colored rectangles in Fig. 10) is investigated. Table IV reports the abundance means for the most significant

[4]These simulations have been carried out with an unoptimized MATLAB 2007b 32-b implementation on a Core(TM)2Duo 2.66-GHz computer.

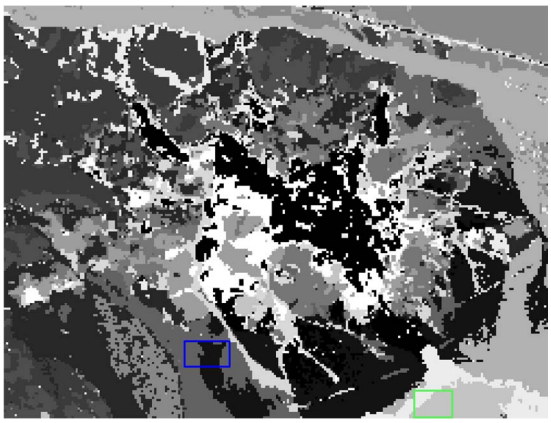

Fig. 10. Classification map for the 190×250 Cuprite area ($K = 14$).

TABLE IV
ABUNDANCE MEANS FOR THE MOST SIGNIFICANT ENDMEMBERS IN EACH HIGHLIGHTED REGION

	Green area	
	light gray	white
Endm. 1	0.001	0.225
Endm. 3	0.045	0.000
Endm. 5	0.098	0.027
Endm. 6	0.839	0.528

	Blue area	
	black	dark gray
Endm. 1	0.135	0.044
Endm. 9	0.155	0.158
Endm. 10	0.159	0.127
Endm. 13	0.187	0.206

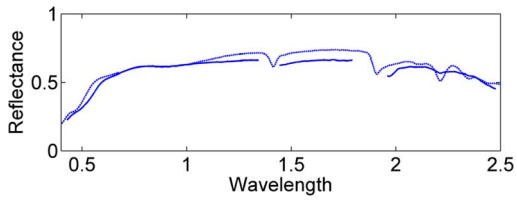

Fig. 11. Comparison of the 6th endmember spectrum extracted by the VCA algorithm (solid line) with the Montmorillonite signature extracted from the USGS spectral library (dashed line).

endmembers that appear in the two highlighted regions. From these tables, one can conclude that the two classes represented in black and dark gray of the "blue" area are composed of very mixed pixels (the abundance of the most significant endmember is 0.201). On the other hand, both classes in the "green" area are clearly dominated by the sixth endmember. By comparing its corresponding signature with the materials included in the USGS library spectra, this sixth endmember matches the Montmorillonite spectrum (see Fig. 11). This result is in agreement with the ground truth. Montmorillonnite is the most commonly found material in this area [33].

VII. CONCLUSION

A new hierarchical Bayesian algorithm was proposed for hyperspectral image unmixing. MRFs were introduced to model spatial correlations between the pixels of the image. A hidden discrete label was introduced for each pixel of the image to identify several classes defined by homogeneous abundances (with constant first and second order statistical moments). The positivity and sum-to-one constraints on the abundances were handled by using an appropriate reparametrization defined by logistic coefficient vectors. We derived the joint posterior distribution of the unknown parameters and hyperparameters associated to the proposed Bayesian LMM. An MCMC method was then studied to generate samples asymptotically distributed according to this posterior. The generated samples were used to estimate the abundance maps as well as the underlying image labels. The results obtained on simulated data and on real AVIRIS images are very promising. Future works include the estimation of the granularity coefficient involved in Potts–Markov random fields.

ACKNOWLEDGMENT

The authors would like to thank one of the reviewers for pointing out the relevant paper [12] and for his valuable suggestions that helped to improve the manuscript.

REFERENCES

[1] N. Keshava and J. Mustard, "Spectral unmixing," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 44–57, Jan. 2002.
[2] C.-I. Chang, *Hyperspectral Data Exploitation: Theory and Applications*. Hoboken, NJ: Wiley, 2007.
[3] P. J. Martinez, R. M. Perez, A. Plaza, P. L. Aguilar, M. C. Cantero, and J. Plaza, "Endmember extraction algorithms from hyperspectral images," *Ann. Geophys.*, vol. 49, no. 1, pp. 93–101, Feb. 2006.
[4] D. C. Heinz and C.-I. Chang, "Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 3, pp. 529–545, Mar. 2001.
[5] C. Theys, N. Dobigeon, J.-Y. Tourneret, and H. Lantéri, "Linear unmixing of hyperspectral images using a scaled gradient method," in *Proc. IEEE-SP Workshop SSP*, Cardiff, U.K., Aug. 2009, pp. 729–732.
[6] N. Dobigeon, J.-Y. Tourneret, and C.-I. Chang, "Semi-supervised linear spectral using a hierarchical Bayesian model for hyperspectral imagery," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 2684–2695, Jul. 2008.
[7] J. Besag, "Spatial interaction and the statistical analysis of lattice systems," *J. Roy. Stat. Soc. Ser. B*, vol. 36, no. 2, pp. 192–236, 1974.
[8] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-6, no. 6, pp. 721–741, Nov. 1984.
[9] R. S. Rand and D. M. Keenan, "Spatially smooth partitioning of hyperspectral imagery using spectral/spatial measures of disparity," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 6, pp. 1479–1490, Jun. 2003.
[10] G. Rellier, X. Descombes, F. Falzon, and J. Zerubia, "Texture feature analysis using a Gauss–Markov model in hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 7, pp. 1543–1551, Jul. 2004.
[11] N. Bali and A. Mohammad-Djafari, "Bayesian approach with hidden Markov modeling and mean field approximation for hyperspectral data analysis," *IEEE Trans. Image Process.*, vol. 17, no. 2, pp. 217–225, Feb. 2008.
[12] J. T. Kent and K. V. Mardia, "Spatial classification using fuzzy membership models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 10, no. 5, pp. 659–671, Sep. 1988.
[13] 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.
[14] F. Wu, "The Potts model," *Rev. Modern Phys.*, vol. 54, no. 1, pp. 235–268, Jan. 1982.
[15] R. Neher and A. Srivastava, "A Bayesian MRF framework for labeling terrain using hyperspectral imaging," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 6, pp. 1363–1374, Jun. 2005.
[16] J. Li, J. M. Bioucas-Dias, and A. Plaza, "Semi-supervised hyperspectral image segmentation using multinomial logistic regression with active learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 11, pp. 4085–4098, Nov. 2010.
[17] M. Fauvel, J. A. Benediktsson, J. Chanussot, and J. R. Sveinsson, "Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 11, pp. 3804–3814, Nov. 2008.

[18] Y. Tarabalka, M. Fauvel, J. Chanussot, and J. A. Benediktsson, "SVM and MRF-based method for accurate classification of hyperspectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 4, pp. 736–740, Oct. 2010.
[19] J. Harsanyi and C.-I. Chang, "Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 4, pp. 779–785, Jul. 1994.
[20] C.-I. Chang, X.-L. Zhao, M. L. G. Althouse, and J. J. Pan, "Least squares subspace projection approach to mixed pixel classification for hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 3, pp. 898–912, May 1998.
[21] O. Eches, N. Dobigeon, and J.-Y. Tourneret, Enhancing hyperspectral image unmixing with spatial correlations, Univ. Toulouse, Toulouse, France, Tech. Rep. [Online]. Available: http://dobigeon.perso.enseeiht.fr/
[22] J.-M. Marin and C. P. Robert, *Bayesian Core: A Practical Approach to Computational Bayesian Statistics*. New-York: Springer-Verlag, 2007.
[23] B. D. Ripley, *Statistical Inference for Spatial Processes*. Cambridge, U.K.: Cambridge Univ. Press, 1988.
[24] A. Gelman, F. Bois, and J. Jiang, "Physiological pharmacokinetic analysis using population modeling and informative prior distributions," *J. Amer. Math. Soc.*, vol. 91, no. 436, pp. 1400–1412, Dec. 1996.
[25] K. Themelis and A. A. Rontogiannis, "A soft constrained MAP estimator for supervised hyperspectral signal unmixing," in *Proc. EUSIPCO*, Lausanne, Switzerland, Aug. 2008.
[26] C. P. Robert, *The Bayesian Choice: From Decision-Theoretic Motivations to Computational Implementation,* 2nd ed. New York: Springer-Verlag, 2007, ser. Springer Texts in Statistics.
[27] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods,* 2nd ed. New York: Springer-Verlag, 2004.
[28] G. O. Roberts, "Markov chain concepts related to sampling algorithms," in *Markov Chain Monte Carlo in Practice*, W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, Eds. London, U.K.: Chapman & Hall, 1996, pp. 259–273.
[29] *ENVI User's Guide Version 4.0*, RSI, Boulder, CO, 2003.
[30] E. Christophe, D. Léger, and C. Mailhes, "Quality criteria benchmark for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 9, pp. 2103–2114, Sep. 2005.
[31] M. E. Winter, "Fast autonomous spectral endmember determination in hyperspectral data," in *Proc. 13th Int. Conf. Appl. Geol. Remote Sens.*, Vancouver, BC, Canada, Apr. 1999, vol. 2, pp. 337–344.
[32] P. Sajda, S. Du, T. R. Brown, R. Stoyanova, D. C. Shungu, X. Mao, and L. C. Parra, "Nonnegative matrix factorization for rapid recovery of constituent spectra in magnetic resonance chemical shift imaging of the brain," *IEEE Trans. Med. Imag.*, vol. 23, no. 12, pp. 1453–1465, Dec. 2004.
[33] R. N. Clark, G. A. Swayze, K. E. Livo, R. F. Kokaly, S. J. Sutley, J. B. Dalton, R. R. McDougal, and C. A. Gent, "Imaging spectroscopy: Earth and planetary remote sensing with the USGS Tetracorder and expert systems," *J. Geophys. Res.*, vol. 108, no. E12, pp. 5-1–5-44, Dec. 2003.
[34] 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.

Olivier Eches was born in Villefranche-de-Rouergue, France, in 1984. He received the Eng. degree in electrical engineering from ENSEEIHT, Toulouse, France, in 2007 and the M.Sc. and Ph.D. degrees in signal processing from the National Polytechnic Institute of Toulouse, University of Toulouse, in 2007 and 2010, respectively.

He is currently a Postdoctoral Research Associate with the Department of Electrical and Computer Engineering, University of Iceland. He is working on segmentation and unmixing of hyperspectral images.

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

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

Jean-Yves Tourneret (SM'08) received the Ingénieur degree in electrical engineering from ENSEEIHT (Toulouse, France) in 1989 and the Ph.D. degree from the National Polytechnic Institute of Toulouse in 1992.

He is currently a professor in the University of Toulouse and a member of the IRIT laboratory (UMR 5505 of the CNRS). His research activities are centered around statistical signal processing with a particular interest to MCMC methods.

Prof. Tourneret was the program chair of the European conference on signal processing (EUSIPCO). He was also member of the organizing committee for the international conference ICASSP'06 which was held in Toulouse (France) in 2006. 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 is currently serving as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING.