Kalman Filtering and Expectation Maximization for Multitemporal Spectral Unmixing

Ricardo A. Borsoi, *Graduate Student Member, IEEE*, Tales Imbiriba, *Member, IEEE*, Pau Closas, *Senior Member, IEEE*, José Carlos M. Bermudez, and Cédric Richard, *Senior Member, IEEE*

***Abstract*— The recent evolution of hyperspectral imaging technology and the proliferation of new emerging applications press for the processing of multiple temporal hyperspectral images. In this work, we propose a novel spectral unmixing (SU) strategy using physically motivated parametric endmember (EME) representations to account for temporal spectral variability. By representing the multitemporal mixing process using a state-space formulation, we are able to exploit the Bayesian filtering machinery to estimate the EME variability coefficients. Moreover, by assuming that the temporal variability of the abundances is small over short intervals, an efficient implementation of the expectation–maximization (EM) algorithm is employed to estimate the abundances and the other model parameters. Simulation results indicate that the proposed strategy outperforms state-of-the-art multi-temporal SU (MTSU) algorithms.**

***Index Terms*— Endmember (EME) variability, expectation–maximization (EM), hyperspectral, Kalman filter, multitemporal unmixing.**

I. Introduction

SPECTRAL unmixing (SU) aims to decompose a hyperspectral image (HI) into its pure spectral components, termed endmembers (EMEs), and the proportional abundances to which they contribute to the reflectance in each pixel [1]. Although the interaction between light and the EMEs can be complex and nonlinear [2]–[4], the linear mixing model (LMM) is still widely used due to its simplicity and good performance [1].

Manuscript received May 21, 2020; revised August 12, 2020; accepted September 18, 2020. Date of publication October 9, 2020; date of current version December 27, 2021. This work was supported in part by the National Council for Scientific and Technological Development (CNPq) under Grant 304250/2017-1, Grant 409044/2018-0, Grant 141271/2017-5, and Grant 204991/2018-8; in part by the Foundation for Research Support of the State of Rio Grande do Sul (FAPERGS) under Grant 19/2551-0001844-4; and in part by the National Science Foundation under Award CNS-1815349 and Award ECCS-1845833. *(Corresponding author: Ricardo A. Borsoi.)*

Ricardo A. Borsoi is with the Department of Electrical Engineering, Universidade Federal de Santa Catarina (DEE-UFSC), Florianópolis 88040-370, Brazil, and also with the Lagrange Laboratory (CNRS, OCA), Université Côte d'Azur (UCA), 06108 Nice, France (e-mail: raborsoi@gmail.com).

Tales Imbiriba and Pau Closas are with the Department of Electrical and Computer Engineering, Northeastern University, Boston, MA 02115 USA (e-mail: talesim@gmail.com; closas@ece.neu.edu).

José Carlos M. Bermudez is with the Department of Electrical Engineering, Universidade Federal de Santa Catarina (DEE-UFSC), Florianópolis 88040-370, Brazil, and also with the Graduate Program on Electronic Engineering and Computing, Catholic University of Pelotas (UCPel), Pelotas 96010-000, Brazil (e-mail: j.bermudez@ieee.org).

Cédric Richard is with the Lagrange Laboratory (CNRS, OCA), Université Côte d'Azur, 06108 Nice, France (e-mail: cedric.richard@unice.fr).

This article has supplementary downloadable material available at https://ieeexplore.ieee.org, provided by the authors.

Digital Object Identifier 10.1109/LGRS.2020.3025781

Spectral variability (SV) consists of changes in EME spectra occurring both within a single image and between images acquired at different time instants. They can be caused by differences in atmospheric, illumination, or seasonal conditions [5], [6]. Early approaches have considered large libraries of spectral signatures to represent variable EME spectra [5], [7]–[9]. More recently, different extensions of the LMM have been proposed to account for the SV within a given HI, by considering, e.g., additive [10] and multiplicative [11]–[13] scaling factors or by parameterizing SV using deep generative models [14].

Multi-temporal SU (MTSU) has recently become a subject of great interest due to the possibility of leveraging time information in HI sequences, allowing for monitoring the dynamical evolution of the materials and their distributions [15]–[17]. However, the influence of SV in multitemporal scenarios can be significantly stronger than in the case of a single HI. This introduces a challenge to multitemporal SU since EME variability must be carefully modeled to achieve a good performance. Previous works have considered different strategies to incorporate dynamical information about the EMEs, often based on parametric models originally devised to account for variations within a single HI. These include constraining the EMEs in adjacent time instants to be scaled versions of each other [18] or to be represented as a mean EME matrix with small, additive perturbations [19]–[21]. However, these works disregard important information as they do not account for the low-dimensional structure that often underlies the changes observed in EME spectra when representing its evolution.

In this letter, we propose a new algorithm for a MTSU that is based on a dynamical model for the EME time variability. Especially, we couple the representation power of recently proposed parametric EME models (which were originally devised to operate within a single HI, such as [12]) with a Bayesian filtering methodology to reliably estimate the EMEs in HI sequences. Instead of operating directly on the EME spectral space, we make use of a parametric EME model to represent EME dynamics indirectly through vectors of parameters that capture the time variations of each material. The Bayesian filtering and smoothing are combined with the expectation–maximization (EM) algorithm to estimate the required parameters given a window of observations in time. The initialization of the resulting Kalman filter is also estimated in the process, which improves convergence for short image sequences. Under some approximations about the temporal variation of the abundances, the proposed algorithm is able to blindly estimate the EMEs, the average abundances, and the remaining model parameters from the observed HI data. Finally, a unique abundance matrix is estimated for each time instant using the resulting EME model. Simulation results

show that, for small abundance variations over time (which can be usually satisfied in small time windows), the proposed method is able to outperform state-of-the-art algorithms in both EME and abundance estimation accuracy.

II. Multitemporal SU

The multitemporal LMM [1] represents an HI with L-band and N pixels at time t as

$$\boldsymbol{Y}_t = \boldsymbol{M}_t \boldsymbol{A}_t + \boldsymbol{E}_t, \quad \text{s.t. } \mathbf{1}^\top \boldsymbol{A}_t = \mathbf{1}^\top, \quad \boldsymbol{A}_t \geq \mathbf{0} \tag{1}$$

where $\boldsymbol{Y}_t \in \mathbb{R}^{L\times N}$ is the observed HI, the columns of $\boldsymbol{M}_t \in \mathbb{R}^{L\times P}$ are the P EME spectral signatures, $\boldsymbol{A}_t \in \mathbb{R}^{P\times N}$ contains the abundances for each pixel, and $\boldsymbol{E}_t$ represents additive noise, all indexed at time $t \in \{1, \ldots, T\}$.

An important challenge related to the use of representation (1) regards the consideration of SV, which causes the signatures of the EMEs in $\boldsymbol{M}_t$ to change due to, e.g., seasonal, illumination, or acquisition variations [5]. SV occurs both in space (within the same HI) and time. Spatial-domain SV has been addressed in several works (see [5], [6], [10]–[14] and references therein). For simplicity, this work assumes only variations of EMEs in time. EME variation within the same HI can be later incorporated into the proposed model, for instance, by adapting models, such as the one in [12], to represent the space–time dynamical behavior of the EMEs.

A straightforward way to perform SU under time variability is to do it for each image separately. However, such an approach disregards the temporal information and the time dynamics of the SV, which can be exploited to enhance both the abundance and EME estimation performance. Different SU algorithms accounting for EME time variability have been recently proposed; most of them are inspired by models designed to account for SV within a single image. For instance, Henrot *et al.* [18] constrain the EME matrices at each time instant to be scaled versions of a reference EME matrix. Thouvenin *et al.* [19] model the EMEs at each time instant by a mean EME matrix plus small perturbations, which are assumed to be temporally smooth. All variables are then estimated using a stochastic approach. This latter model was later extended for distributed unmixing with additional sparsity constraints in [20] and to include sparse additive residual terms to represent abrupt spectral variations in the HI using a hierarchical Bayesian framework in [21]. However, these works do not provide a satisfactory means of modeling the dynamical evolution of the EMEs since they operate directly in the input spectral space, ignoring the fact that SV can often be represented more accurately using physically meaningful parameterizations of EME spectra.

Different models have been recently proposed to model EME spatial variability as a parametric function of reference spectral signatures as

$$\boldsymbol{M} = f(\boldsymbol{M}_0, \boldsymbol{\psi}) \tag{2}$$

where f is a parametric function, $\boldsymbol{M}_0 \in \mathbb{R}^{L\times P}$ contains reference/average spectral signatures, and $\boldsymbol{\psi}$ is a vector of parameters of the variability model. Such models include additive perturbations [10], spectrally uniform [11] or spectrally varying [12], [13] scaling factors, and parameterizations using deep neural networks learned from the observed HI [14]. Such parametric models are especially interesting for building a dynamical model to consider EME time variability.

III. Dynamical Parametric EME Model

In this letter, we consider a multitemporal extension of the parametric EME model (2). We assume a fixed reference EME matrix $\boldsymbol{M}_0$ and model the time variations in $\boldsymbol{M}_t$ through a time varying $\boldsymbol{\psi}_t$, $t = 1, \ldots, T$. By assuming that temporally adjacent images are acquired at reasonably short time intervals, we model the difference $\boldsymbol{\psi}_t - \boldsymbol{\psi}_{t-1}$ as a small zero-mean vector. Thus, we assume the following model for $\boldsymbol{\psi}_t$:

$$\boldsymbol{\psi}_t = \boldsymbol{\psi}_{t-1} + \boldsymbol{q}_t \tag{3}$$

where $\boldsymbol{\psi}_t$ is a vector containing the parameters of the EME model at time t and $\boldsymbol{q}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{Q})$ contains the innovations that describe its dynamical evolution. Note that $\boldsymbol{q}_t$ is only constrained to be zero mean on statistical and not temporal average, which means that each realization of the sequence $\{\widehat{\boldsymbol{q}}_t\}$, which is learned from the observed HIs, can exhibit behavior, such as trends and complex dynamic evolution. Moreover, the Gaussian assumption is only made in the model parameters $\boldsymbol{\psi}_t$ and not on the EME signatures themselves, which allows for the use of complex EME distributions through the pushforward measure obtained using the function f, as done in [14]. This generalizes the parametric EME model (2) to the multitemporal setting as $\boldsymbol{M}_t = f(\boldsymbol{M}_0, \boldsymbol{\psi}_t)$, where the parametric function f now relates the EME matrices and the vectors of parameters at each time instant. Considering this model, the multitemporal LMM can be represented as

$$\boldsymbol{Y}_t = f(\boldsymbol{M}_0, \boldsymbol{\psi}_t)\boldsymbol{A}_t + \boldsymbol{E}_t. \tag{4}$$

Next, one must choose a function f for (4) that establishes a good compromise between mathematical tractability and performance. The generalized LMM (GLMM) [12], [22] is able to represent arbitrary SV by considering spectrally varying multiplicative scaling factors, introducing a connection between the amount of SV and the amplitude of EME reflectance spectra at each band. The GLMM introduces a matrix $\boldsymbol{\Psi} \in \mathbb{R}^{L\times P}$ of scaling factors with nonnegative entries $[\boldsymbol{\Psi}]_{\ell,k} \geq 0$ acting individually at each wavelength. This leads to the following representation for the tth observed HI:

$$\boldsymbol{Y}_t = (\boldsymbol{M}_0 \odot \boldsymbol{\Psi}_t)\boldsymbol{A}_t + \boldsymbol{E}_t \tag{5}$$

where $\odot$ is the Hadamard (elementwise) product. Using the vectorization property, (5) can be expressed as

$$\boldsymbol{y}_t = \mathrm{vec}(\boldsymbol{Y}_t) = \left(\boldsymbol{A}_t^\top \otimes \boldsymbol{I}_L\right) \mathrm{diag}(\boldsymbol{m}_0)\boldsymbol{\psi}_t + \boldsymbol{e}_t \tag{6}$$

with $\boldsymbol{m}_0 = \mathrm{vec}(\boldsymbol{M}_0)$ and $\boldsymbol{\psi}_t = \mathrm{vec}(\boldsymbol{\Psi}_t)$ and $\boldsymbol{e}_t = \mathrm{vec}(\boldsymbol{E}_t)$.

We write the abundance matrix $\boldsymbol{A}_t$ as $\boldsymbol{A}_t = \boldsymbol{A} + \boldsymbol{\Delta A}_t$, where $\boldsymbol{\Delta A}_t$ represents small random fluctuations over the average abundance matrix $\boldsymbol{A}$. Considering $\boldsymbol{\Delta A}_t$ to be small for a time window $t \in \{t_0, \ldots, t_0 + T\}$ $\forall t_0$, these variations can be incorporated into the observation noise, leading to the following model:

$$\boldsymbol{y}_t = \boldsymbol{H}(\boldsymbol{A})\,\mathrm{diag}(\boldsymbol{m}_0)\boldsymbol{\psi}_t + \boldsymbol{r}_t \tag{7}$$

where $\boldsymbol{H}(\boldsymbol{A}) = \boldsymbol{A}^\top \otimes \boldsymbol{I}_L$ and $\boldsymbol{r}_t = \boldsymbol{e}_t + (\boldsymbol{\Delta A}_t^\top \otimes \boldsymbol{I}_L)\,\mathrm{diag}(\boldsymbol{m}_0)\boldsymbol{\psi}_t$. Note that the observation noise $\boldsymbol{r}_t$ in (7) is correlated with the state $\boldsymbol{\psi}_t$. In the following, we will use a signal-independent noise approximation, which provides competitive performance at a modest computational cost. Further discussion on the impact of such an approximation can be found in the Supplementary Material, also available in [23].

IV. Proposed Method

In this section, we present the proposed dynamical methodology that connects the Kalman smoother with EM approach. For this, we assume that, for a given time window of duration T, the abundance variation is small, but the EMEs can vary due to different seasonal or acquisition conditions. Then, we employ a time-varying state-space formulation to model the SV, which naturally leads to a Kalman filter-based

formulation. We couple a Kalman smoother, used to obtain accurate estimations for the state variables, with the EM estimation of model parameters, such as the abundance matrix and the noise power. Assuming the abundances fixed over a time window $t \in \{t_0, \ldots, t_0 + T\}$, we use (3) and (7) to form the linear state-space model

$$\boldsymbol{\psi}_t = \boldsymbol{\psi}_{t-1} + \boldsymbol{q}_t, \quad \boldsymbol{y}_t = \boldsymbol{H}(\boldsymbol{A}) \operatorname{diag}(\boldsymbol{m}_0)\boldsymbol{\psi}_t + \boldsymbol{r}_t. \tag{8}$$

Neglecting the dependence of $\boldsymbol{r}_t$ on $\boldsymbol{\psi}_t$ and assuming $\boldsymbol{q}_t$ and $\boldsymbol{r}_t$ to be Gaussian, this system can be solved using the classical Kalman filter and smoothing equations. Next, we present the Kalman filter and smoother equations followed by the EM strategy to estimate the abundances and noise power.

A. Kalman Filter EME Model

The Bayesian filtering computes marginal posterior distributions of the states by assuming Markovity over the state sequence. When the dynamical and measurement models are linear and Gaussian, the solution is given in the form of the Kalman filter, which can be expressed in a set of equations for the Prediction

$$\boldsymbol{\psi}_{t|t-1} = \boldsymbol{\psi}_{t-1|t-1}, \quad \boldsymbol{P}_{t|t-1} = \boldsymbol{P}_{t-1|t-1} + \boldsymbol{Q} \tag{9}$$

and for the Update step

$$\begin{aligned} \boldsymbol{v}_t &= \boldsymbol{y}_t - \boldsymbol{B}\boldsymbol{\psi}_{t|t-1} \\ \boldsymbol{S}_t &= \boldsymbol{B}\boldsymbol{P}_{t|t-1}\boldsymbol{B}^\top + \boldsymbol{R} \\ \boldsymbol{K}_t &= \boldsymbol{P}_{t|t-1}\boldsymbol{B}^\top \boldsymbol{S}_t^{-1} \\ \boldsymbol{\psi}_{t|t} &= \boldsymbol{\psi}_{t|t-1} + \boldsymbol{K}_t \boldsymbol{v}_t \\ \boldsymbol{P}_{t|t} &= \boldsymbol{P}_{t|t-1} - \boldsymbol{K}_t \boldsymbol{S}_t \boldsymbol{K}_t^\top \end{aligned} \tag{10}$$

where $\boldsymbol{P}_{t_1|t_2}$ is the covariance matrix of $\boldsymbol{\psi}_{t_1}$ conditioned on $\boldsymbol{y}_t$ for $t_1 \leq t_2$, $\boldsymbol{B} = \boldsymbol{H}(\boldsymbol{A}) \operatorname{diag}(\boldsymbol{m}_0) = (\boldsymbol{A}^\top \otimes \boldsymbol{I}_L) \operatorname{diag}(\boldsymbol{m}_0)$, and $\boldsymbol{R}$ is the covariance matrix of $\boldsymbol{r}_t$ in (8). Solving (10) requires to construct and invert matrix $\boldsymbol{S}_t$ of size $NL \times NL$, which is impractical. To circumvent this issue, we assume that the noise covariance matrix satisfies $\boldsymbol{R} = \sigma_r^2 \boldsymbol{I}_{NL}$. Thus, using the Woodbury identity for the inverse of sum of matrices, the right part of the third term in (10) is written as

$$\boldsymbol{B}^\top \boldsymbol{S}_t^{-1} = \sigma_r^{-2}\boldsymbol{B}^\top - \sigma_r^{-4}\boldsymbol{B}^\top \boldsymbol{B}\left(\boldsymbol{P}_{t|t-1}^{-1} + \sigma_r^{-2}\boldsymbol{B}^\top \boldsymbol{B}\right)^{-1}\boldsymbol{B}^\top$$

which now involves only the inverse of a $PL \times PL$ matrix.

B. Kalman Smoother

The objective of the Bayesian smoothers is to provide a marginal posterior distribution of state $\boldsymbol{\psi}_t$ assuming knowledge of the measurements $\boldsymbol{y}_t$ in an observation window of duration T, that is, $p(\boldsymbol{\psi}_t | \boldsymbol{y}_{t_0}, \ldots, \boldsymbol{y}_{t_0+T})$. For the model in (8), the smoother solution can be implemented very efficiently by iteratively updating the conditional distributions obtained by the Kalman filter backwards in time, for $t_0 + T, \ldots, t_0$. In this case, the smoother equations are given by

$$\begin{aligned} \boldsymbol{\psi}_{t+1|t} &= \boldsymbol{\psi}_{t|t}, \quad \boldsymbol{P}_{t+1|t} = \boldsymbol{P}_{t|t} + \boldsymbol{Q}, \ \boldsymbol{G}_t = \boldsymbol{P}_{t|t}\boldsymbol{P}_{t+1|t}^{-1} \\ \boldsymbol{\psi}_t^s &= \boldsymbol{\psi}_{t|t} + \boldsymbol{G}_t\left[\boldsymbol{\psi}_{t+1}^s - \boldsymbol{\psi}_{t+1|t}\right] \\ \boldsymbol{P}_t^s &= \boldsymbol{P}_{t|t} + \boldsymbol{G}_t\left[\boldsymbol{P}_{t+1}^s - \boldsymbol{P}_{t+1|t}\right]\boldsymbol{G}_t^\top. \end{aligned} \tag{11}$$

C. EM Algorithm

Estimation of the sequence $\{\boldsymbol{\psi}_t\}$ of EME model parameters using (9)–(11) requires that $\boldsymbol{A}$, $\boldsymbol{Q}$, and $\boldsymbol{R}$, as well as the initializations $\boldsymbol{P}_{0|0}$ and $\boldsymbol{\psi}_{0|0}$, be known in advance. Let us denote these parameters by $\boldsymbol{\theta} = \{\boldsymbol{A}, \ \boldsymbol{P}_{0|0}, \ \boldsymbol{Q}, \ \boldsymbol{R}, \ \boldsymbol{\psi}_{0|0}\}$. Instead of fixing $\boldsymbol{\theta}$ with values known *a priori*, we can view it as unobserved latent variables of model (8), which can be estimated by maximizing the conditional marginal likelihood $p(\boldsymbol{y}_{t_0}, \ldots, \boldsymbol{y}_{t_0+T}|\boldsymbol{\theta})$ using the EM algorithm. Starting with an initial guess $\boldsymbol{\theta}^{(0)}$, the EM algorithm finds a local maximum of $p(\boldsymbol{y}_{t_0}, \ldots, \boldsymbol{y}_{t_0+T}|\boldsymbol{\theta})$ by repeating the following steps:

$$\begin{aligned} &a) \text{ E-step: compute } \mathcal{Q}(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)}) \\ &b) \text{ M-step: compute } \boldsymbol{\theta}^{(k+1)} = \arg\max_{\boldsymbol{\theta}} \ \mathcal{Q}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(k)}) \end{aligned} \tag{12}$$

for $k = 1, \ldots, K_{max}$, with K_{max} being the number of iterations and $\mathcal{Q}(\boldsymbol{\theta}|\boldsymbol{\theta}^{(k)}) = \mathrm{E}_\varsigma\{\log p(\boldsymbol{\psi}_{t_0}, \ldots, \boldsymbol{\psi}_{t_0+T}, \boldsymbol{y}_{t_0}, \ldots, \boldsymbol{y}_{t_0+T}|\boldsymbol{\theta})\}$, with $\varsigma = p(\boldsymbol{\psi}_{t_0}, \ldots, \boldsymbol{\psi}_{t_0+T}, \boldsymbol{y}_{t_0}, \ldots, \boldsymbol{y}_{t_0+T}|\boldsymbol{\theta}^{(k)})$ being the expectation of the logarithm of the data likelihood, taken with respect to the full joint posterior given the parameters $\boldsymbol{\theta}^{(k)}$. Although the EM algorithm is very general and not always easy to solve, for a linear model such as (8), we can find closed-form solutions, leading to more efficient implementation in high-dimensional settings. Furthermore, for the linear Gaussian model, $\mathcal{Q}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(k)})$ can be computed based on the Kalman smoother results obtained using $\boldsymbol{\theta}^{(k)}$ as the system parameters. This leads to an elegant solution that consists of the successive application of the smoother and estimation of the parameters. For the model (8), $\mathcal{Q}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(t)})$ is given by Särkkä [24]

$$\begin{aligned} &\mathcal{Q}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(t)}) \\ &\quad = -\frac{1}{2}\Big(\operatorname{tr}\left\{\boldsymbol{P}_{0|0}^{-1}\left[\boldsymbol{P}_0^s + \left(\boldsymbol{\psi}_0^s - \boldsymbol{\psi}_{0|0}\right)\left(\boldsymbol{\psi}_0^s - \boldsymbol{\psi}_{0|0}\right)^\top\right]\right\} \\ &\qquad + \operatorname{tr}\left\{\boldsymbol{R}^{-1}\left[\boldsymbol{\Sigma}_5 - \boldsymbol{\Sigma}_3\boldsymbol{H}(\boldsymbol{A})^\top - \boldsymbol{H}(\boldsymbol{A})\boldsymbol{\Sigma}_3^\top \right.\right. \\ &\qquad \left.\left. + \boldsymbol{H}(\boldsymbol{A})\boldsymbol{\Sigma}_1\boldsymbol{H}(\boldsymbol{A})^\top\right]\right\} \\ &\qquad + \operatorname{tr}\left\{\boldsymbol{Q}^{-1}\left[\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_4 - \boldsymbol{\Sigma}_4^\top + \boldsymbol{\Sigma}_2\right]\right\} \\ &\qquad + \log\{|\boldsymbol{Q}\boldsymbol{R}|^T |\boldsymbol{P}_{0|0}|\}\Big) + C \end{aligned}$$

where C is a constant term and

$$\boldsymbol{\Sigma}_1 = \sum_{t=1}^{T} \boldsymbol{P}_t^s + \boldsymbol{\psi}_t^s \boldsymbol{\psi}_t^{s\top}, \quad \boldsymbol{\Sigma}_4 = \sum_{t=1}^{T} \boldsymbol{P}_t^s \boldsymbol{G}^{s\top}_{t-1} + \boldsymbol{\psi}_t^s \boldsymbol{\psi}^{s\top}_{t-1}$$

$$\boldsymbol{\Sigma}_2 = \sum_{t=1}^{T} \boldsymbol{P}_{t-1}^s + \boldsymbol{\psi}_{t-1}^s \boldsymbol{\psi}^{s\top}_{t-1}, \quad \boldsymbol{\Sigma}_3 = \sum_{t=1}^{T} \boldsymbol{y}_t \boldsymbol{\psi}_t^{s\top}$$

$$\boldsymbol{\Sigma}_5 = \sum_{t=1}^{T} \boldsymbol{y}_t \boldsymbol{y}_t^\top$$

Under the assumption that $\boldsymbol{R} = \sigma_r^2 \boldsymbol{I}_{NL}$, optimizing $\mathcal{Q}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(t)})$ with respect to $\boldsymbol{P}_{0|0}$, $\boldsymbol{Q}$, $\boldsymbol{R}$, and $\boldsymbol{\psi}_{0|0}$ is relatively straightforward and can be done as [24]

$$\boldsymbol{P}_{0|0}^* = \boldsymbol{P}_0^s + \left(\boldsymbol{\psi}_0^s - \boldsymbol{\psi}_{0|0}\right)\left(\boldsymbol{\psi}_0^s - \boldsymbol{\psi}_{0|0}\right)^\top \tag{13}$$

$$\boldsymbol{Q}^* = \boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_4 - \boldsymbol{\Sigma}_4^\top + \boldsymbol{\Sigma}_2 \tag{14}$$

$$\sigma_r^* = \operatorname{tr}\left\{\boldsymbol{\Sigma}_5 - 2\boldsymbol{H}(\boldsymbol{A})\boldsymbol{\Sigma}_3^\top + \boldsymbol{H}(\boldsymbol{A})\boldsymbol{\Sigma}_1\boldsymbol{H}(\boldsymbol{A})^\top\right\}/(LN) \tag{15}$$

$$\boldsymbol{\psi}_{0|0}^* = \boldsymbol{\psi}_0^s. \tag{16}$$

The optimization with respect to $\boldsymbol{H}(\boldsymbol{A})$ is, however, more complex due to the structure of this matrix. Since $\boldsymbol{R} = \sigma_r^2 \boldsymbol{I}$ for some $\sigma_r > 0$, the problem can be stated as

$$\widehat{\boldsymbol{A}} = \arg\min_{\boldsymbol{A}} \ \operatorname{tr}\{\boldsymbol{H}(\boldsymbol{A})\boldsymbol{\Sigma}_1\boldsymbol{H}(\boldsymbol{A})^\top - 2\,\boldsymbol{\Sigma}_3\boldsymbol{H}(\boldsymbol{A})^\top\} \tag{17}$$

Algorithm 1 Kalman Filter and Smoother for MTSU

Input : $\{\boldsymbol{y}_t\}_{t=1}^{T}$, $\boldsymbol{A}^{(0)}$, $\boldsymbol{\psi}^{(0)}$, $\boldsymbol{M}_0$. $\boldsymbol{\psi}_{0|0}^{(0)}$, $\boldsymbol{Q}^{(0)}$, $\sigma_r^{(0)}$, $\boldsymbol{P}_{0|0}^{(0)}$, λ

for $i = 1, \ldots, K_{\max}$ **do**
  Estimate $\boldsymbol{\psi}_t$ using (9)–(10) for $t = 1, \ldots, T$;
  Estimate $\boldsymbol{\psi}_t^s$ and $\boldsymbol{P}_t^s$ using (11) for $t = T, \ldots, 1$;
  Estimate $\boldsymbol{P}_{0|0}^{(i)}$, $\boldsymbol{Q}^{(i)}$, $\sigma_r^{(i)}$, $\boldsymbol{\psi}_{0|0}^{(i)}$, $\boldsymbol{A}^{(i)}$ using (13)–(16), (21);
end
Estimate the temporal abundance variations according to (22);
return $\boldsymbol{A}_t^* = \widehat{\boldsymbol{A}}_t$, $\boldsymbol{M}_t^* = \boldsymbol{M}_0 \odot \text{vec}^{-1}(\boldsymbol{\psi}_t^s)$, for $t = 1, \ldots, T$

In order to solve (17) efficiently, we rewrite its terms in the following to explore the structure of $\boldsymbol{H}(\boldsymbol{A})$. For the first term

$$\begin{aligned}\text{tr}\{\boldsymbol{H}(\boldsymbol{A})\boldsymbol{\Sigma}_1\boldsymbol{H}(\boldsymbol{A})^\top\} &= \text{tr}\{(\boldsymbol{A}\otimes\boldsymbol{I}_L)(\boldsymbol{A}^\top\otimes\boldsymbol{I}_L)\widetilde{\boldsymbol{\Sigma}}_1\}\\ &= \text{tr}\{(\boldsymbol{A}\boldsymbol{A}^\top\otimes\boldsymbol{I}_L)\widetilde{\boldsymbol{\Sigma}}_1\}\end{aligned} \tag{18}$$

where $\widetilde{\boldsymbol{\Sigma}}_1 = \text{diag}(\boldsymbol{m}_0)\boldsymbol{\Sigma}_1\,\text{diag}(\boldsymbol{m}_0)$. To explore the properties of the Kronecker product and simplify the solution to this problem, we assume that $\widetilde{\boldsymbol{\Sigma}}_1$ can be decomposed as [25], [26] $\widetilde{\boldsymbol{\Sigma}}_1 = \sum_{k=1}^{K_1}\boldsymbol{C}_k\otimes\boldsymbol{D}_k$, $\boldsymbol{C}_k \in \mathbb{R}^{P\times P}$, $\boldsymbol{D}_k \in \mathbb{R}^{L\times L}$, and using the properties of the Kronecker product, we have

$$\text{tr}\{(\boldsymbol{A}\boldsymbol{A}^\top\otimes\boldsymbol{I}_L)\widetilde{\boldsymbol{\Sigma}}_1\} = \sum_{k=1}^{K_1}\text{tr}\{(\boldsymbol{A}\boldsymbol{A}^\top\boldsymbol{C}_k)\}\,\text{tr}\{\boldsymbol{D}_k\}. \tag{19}$$

Similarly, for the second term, we have $\text{tr}\{\boldsymbol{H}(\boldsymbol{A})\boldsymbol{\Sigma}_3^\top\} = \text{tr}\{(\boldsymbol{A}^\top\otimes\boldsymbol{I}_L)\widetilde{\boldsymbol{\Sigma}}_3^\top\}$, where $\widetilde{\boldsymbol{\Sigma}}_3 = \boldsymbol{\Sigma}_3\,\text{diag}(\boldsymbol{m}_0)$. Decomposing $\widetilde{\boldsymbol{\Sigma}}_3$ as in [25] and [26] leads to $\widetilde{\boldsymbol{\Sigma}}_3 = \sum_{k=1}^{K_2}\widetilde{\boldsymbol{C}}_k\otimes\widetilde{\boldsymbol{D}}_k$, $\widetilde{\boldsymbol{C}}_k \in \mathbb{R}^{N\times P}$, $\widetilde{\boldsymbol{D}}_k \in \mathbb{R}^{L\times L}$, and using the properties of the Kronecker product, we have

$$\text{tr}\{(\boldsymbol{A}^\top\otimes\boldsymbol{I}_L)\widetilde{\boldsymbol{\Sigma}}_3\} = \sum_{k=1}^{K_2}\text{tr}\left\{\left(\boldsymbol{A}^\top\widetilde{\boldsymbol{C}}_k^\top\right)\right\}\text{tr}\left\{\widetilde{\boldsymbol{D}}_k^\top\right\}. \tag{20}$$

By substituting (19) and (20) in (17), taking the derivative of the cost function with respect to $\boldsymbol{A}$ and setting it equal to zero, we obtain the following solution for $\boldsymbol{A}$:

$$\widehat{\boldsymbol{A}} = \left[\sum_{k=1}^{K_1}\text{tr}\{\boldsymbol{D}_k\}\left(\boldsymbol{C}_k+\boldsymbol{C}_k^\top\right)\right]^{-1}\left[2\sum_{k=1}^{K_2}\text{tr}\left\{\widetilde{\boldsymbol{D}}_k^\top\right\}\widetilde{\boldsymbol{C}}_k^\top\right]. \tag{21}$$

Although the approach presented in Sections IV-A–IV-C provides an estimate $\widehat{\boldsymbol{A}}$ of the average abundances, the temporal abundance variations $\Delta\boldsymbol{A}_t$ can make $\widehat{\boldsymbol{A}}$ an inaccurate approximation of $\boldsymbol{A}_t$ for some image sequences (e.g., when sudden changes are present). To mitigate this issue, we compensate the abundance variations $\Delta\boldsymbol{A}_t$ by solving using a fully constrained least-squares (FCLS) problem

$$\begin{aligned}&\min_{\boldsymbol{A}_t}\ \|\boldsymbol{Y}_t - (\boldsymbol{M}_0\odot\widehat{\boldsymbol{\Psi}}_t)\boldsymbol{A}_t\|_F^2 + \lambda\|\boldsymbol{A}_t - \widehat{\boldsymbol{A}}\|_F^2\\ &\text{s.t. } \boldsymbol{A}_t \geq 0,\quad \boldsymbol{1}^\top\boldsymbol{A}_t = \boldsymbol{1}^\top\end{aligned} \tag{22}$$

for $t = 1, \ldots, T$, where $\widehat{\boldsymbol{\Psi}}_t$ is the matrix-ordered version of the estimated states $\boldsymbol{\psi}_t^s$ and $\lambda \geq 0$ is a regularization parameter. The proposed methodology is summarized in Algorithm 1.

V. Experimental Results

In this section, we evaluate the performance of the proposed method by comparing it with the FCLSs and with the online unmixing (OU) strategy proposed in [19]. To illustrate how the use of temporal information improves SU, we have also applied the GLMM algorithm [12] (which considers SV only within a single HI) to process each HI independently. In all experiments, the reference EME matrix $\boldsymbol{M}_0$ was extracted from the observed HI at $t = 1$ using the vertex component analysis (VCA) algorithm [27], and the abundances were initialized with the corresponding FCLS result. The other parameters were initialized as $\boldsymbol{\psi}_{0|0} = \boldsymbol{1}$, $\boldsymbol{Q} = 0.1\boldsymbol{I}$, $\sigma_r = 0.01$, $\boldsymbol{P}_{0|0} = \boldsymbol{I}$, and $\lambda = 10^{-8}$, and five EM iterations were considered. The parameters of the OU algorithm were searched in the ranges detailed in the original publication [19].

TABLE I
Quantitative Simulation Results (Values × 100)

	Synthetic Data (average results)				Real Data
	$\text{NRMSE}_{\boldsymbol{A}}$	$\text{NRMSE}_{\boldsymbol{M}}$	$\text{SAM}_{\boldsymbol{M}}$	$\text{NRMSE}_{\boldsymbol{Y}}$	$\text{NRMSE}_{\boldsymbol{Y}}$
FCLS	4.60	3.30	2.47	4.03	8.73
GLMM	4.87	4.25	3.69	**2.93**	**0.15**
OU	3.40	2.82	2.05	3.09	3.48
Proposed	**2.65**	**2.06**	**1.83**	3.12	9.02

The performance of the methods is evaluated using the average normalized root mean squared error (NRMSE) between the estimated abundances ($\text{NRMSE}_{\boldsymbol{A}}$), EMEs ($\text{NRMSE}_{\boldsymbol{M}}$) and between the reconstructed HIs. NRMSE is defined as $\text{NRMSE}_{\boldsymbol{X}} = (1/T)\sum_{t=1}^{T}(\|\boldsymbol{X}_t - \boldsymbol{X}_t^*\|_F^2 \,/\, \|\boldsymbol{X}_t\|_F^2)^{1/2}$, where $\boldsymbol{X}_t$ and $\boldsymbol{X}_t^*$ are a true and an estimated variable, respectively, at time instant t. We also consider the average spectral angle mapper (SAM) between the estimated EMEs, defined as $\text{SAM}_{\boldsymbol{M}} = (1/T)\sum_{t=1}^{T}\sum_{k=1}^{P}\arccos(\boldsymbol{m}_{k,t}^\top\boldsymbol{m}_{k,t}^*/\|\boldsymbol{m}_{k,t}\|\|\boldsymbol{m}_{k,t}^*\|)$.

A synthetic data set with $L = 173$ bands, $N = 50$ pixels, and $T = 10$ frames was created by generating abundance values sampled from a Dirichlet distribution. HIs containing three EMEs (vegetation, water, and soil) were generated following the model (5) to generate one different EME matrix for each time instant. Temporal SV was introduced by performing a random walk according to the model $\boldsymbol{\psi}_t = \boldsymbol{F}\boldsymbol{\psi}_{t-1} + \boldsymbol{q}_t$ with $\boldsymbol{F} = 0.9\boldsymbol{I}$, $\boldsymbol{\psi}_0 = \boldsymbol{1}_{LP}$, and covariance $0.01\boldsymbol{I}$. Finally, the white Gaussian noise was added to the images, resulting in an SNR of 30 dB. The $\text{SAM}_{\boldsymbol{M}}$ metric for the GLMM method was computed by considering the average EME matrix (across all pixels) for each time instant. In order to evaluate the performance of the algorithms, we performed 900 Monte Carlo runs. In addition, we also simulated different amounts of temporal abundance variability, with the temporal standard deviation of each pixel being, on average, approximately 3×10^{-3}.

The average metrics (across all MC runs and variance values) are shown in Table I. It can be seen that the proposed method performs significantly better than other algorithms. Improvements over OU can be found in $\text{NRMSE}_{\boldsymbol{A}}$ (22%), $\text{NRMSE}_{\boldsymbol{M}}$ (27%), and $\text{SAM}_{\boldsymbol{M}}$ (11%). The GLMM provided the smallest reconstruction error $\text{NRMSE}_{\boldsymbol{Y}}$ since it has the largest amount of degrees of freedom. However, this did not translate into better abundance or EME estimates since spatial SV was not present in the data, and time information was not taken into account.

For the simulations with real data, we consider three images from the Lake Tahoe sequence, originally studied in [19], acquired on April 10, 2014, June 02, 2014, and April 29, 2015. These images can be seen in Fig. 1 and are composed

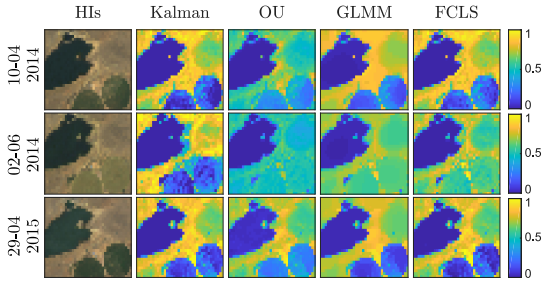

Fig. 1. (Right) Abundance maps for the soil EME and (Left) visual representation of the Lake Tahoe HI sequence.

of three EMEs: water (a lake) and soil and vegetation (two crop circles, whose aspect varies considerably between the three HIs). The images were first downsized to 28×38 pixels for faster processing and contained $L = 173$ bands. $\boldsymbol{M}_0$ and the OU EME initialization were constructed using the same signatures as in [19]. The OU parameters were the same as those used in [19]. The results can be seen in Fig. 1 and Table I. Due to space limitations, only the soil EME is shown (more results can be found in the Supplementary Material, also available in [23]). It can be seen that the proposed algorithm provides higher abundance values in the regions corresponding to the soil in the HI, with significantly less confusion in the vegetation EME compared with the other methods, especially for $t = 2$. Similar improvements could be noticed for the other EMEs as well. The reconstruction error $\text{NRMSE}_{\boldsymbol{Y}}$ of the algorithms is generally related to their overall amount of degrees of freedom, thus being much larger for both the proposed method and FCLS than for GLMM.

VI. Conclusion

We proposed a new MTSU algorithm accounting for SV. A state-space dynamical model was proposed for the time evolution of the coefficients encoding the SV of the EMEs. Bayesian filtering was used to estimate the state variables. Assuming small abundance variations in short time intervals, EM employed to efficiently estimate the remaining parameters, including the fractional abundances. Simulation results indicate that the proposed method can outperform state-of-the-art MTSU algorithms. A future perspective is the extension of the method to properly handle abrupt abundance changes.

References

[1] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 44–57, Jan. 2002.

[2] 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. 82–94, Jan. 2014.

[3] T. Imbiriba, J. C. M. Bermudez, C. Richard, and J.-Y. Tourneret, "Nonparametric detection of nonlinearly mixed pixels and endmember estimation in hyperspectral images," *IEEE Trans. Image Process.*, vol. 25, no. 3, pp. 1136–1151, Mar. 2016.

[4] R. A. Borsoi, T. Imbiriba, J. C. M. Bermudez, and C. Richard, "A blind multiscale spatial regularization framework for kernel-based spectral unmixing," *IEEE Trans. Image Process.*, vol. 29, pp. 4965–4979, 2020.

[5] A. Zare and K. C. Ho, "Endmember variability in hyperspectral analysis: Addressing spectral variability during spectral unmixing," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 95–104, Jan. 2014.

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

[7] D. A. Roberts, M. Gardner, R. Church, S. Ustin, G. Scheer, and R. O. Green, "Mapping chaparral in the santa monica mountains using multiple endmember spectral mixture models," *Remote Sens. Environ.*, vol. 65, no. 3, pp. 267–279, Sep. 1998.

[8] R. A. Borsoi, T. Imbiriba, J. C. M. Bermudez, and C. Richard, "A fast multiscale spatial regularization for sparse hyperspectral unmixing," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 4, pp. 598–602, Apr. 2019.

[9] R. A. Borsoi, T. Imbiriba, J. C. M. Bermudez, and C. Richard, "Deep generative models for library augmentation in multiple endmember spectral mixture analysis," *IEEE Geosci. Remote Sens. Lett.*, early access, Jul. 22, 2020, doi: 10.1109/lgrs.2020.3007161.

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

[12] T. Imbiriba, R. A. Borsoi, and J. C. Moreira Bermudez, "Generalized linear mixing model accounting for endmember variability," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Calgary, AB, Canada, Apr. 2018, pp. 1862–1866.

[13] R. A. Borsoi, T. Imbiriba, and J. C. Moreira Bermudez, "Improved hyperspectral unmixing with endmember variability parametrized using an interpolated scaling tensor," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Brighton, U.K., May 2019, pp. 2177–2181.

[14] R. A. Borsoi, T. Imbiriba, and J. C. M. Bermudez, "Deep generative endmember modeling: An application to unsupervised spectral unmixing," *IEEE Trans. Comput. Imag.*, vol. 6, pp. 374–384, 2020.

[15] B. Somers and G. P. Asner, "Invasive species mapping in Hawaiian rainforests using multi-temporal hyperion spaceborne imaging spectroscopy," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 351–359, Apr. 2013.

[16] C. L. Lippitt, D. A. Stow, D. A. Roberts, and L. L. Coulter, "Multidate MESMA for monitoring vegetation growth forms in southern California shrublands," *Int. J. Remote Sens.*, vol. 39, no. 3, pp. 655–683, Feb. 2018.

[17] B. Somers and G. P. Asner, "Multi-temporal hyperspectral mixture analysis and feature selection for invasive species mapping in rainforests," *Remote Sens. Environ.*, vol. 136, pp. 14–27, Sep. 2013.

[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] 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.

[20] J. Sigurdsson, M. O. Ulfarsson, J. R. Sveinsson, and J. M. Bioucas-Dias, "Sparse distributed multitemporal hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 11, pp. 6069–6084, Nov. 2017.

[21] 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," *IEEE Trans. Comput. Imag.*, vol. 4, no. 1, pp. 32–45, Mar. 2018.

[22] R. A. Borsoi, T. Imbiriba, and J. C. M. Bermudez, "Super-resolution for hyperspectral and multispectral image fusion accounting for seasonal spectral variability," *IEEE Trans. Image Process.*, vol. 29, pp. 116–127, 2020.

[23] R. Augusto Borsoi, T. Imbiriba, P. Closas, J. Carlos Moreira Bermudez, and C. Richard, "Kalman filtering and expectation maximization for multitemporal spectral unmixing," 2020, *arXiv:2001.00425*. [Online]. Available: http://arxiv.org/abs/2001.00425

[24] S. Särkkä, *Bayesian Filtering Smoothing*, vol. 3. Cambridge, U.K.: Cambridge Univ. Press, 2013.

[25] C. F. Van Loan and N. Pitsianis, "Approximation with Kronecker products," in *Linear Algebra for Large Scale and Real-Time Applications*. Dordrecht, The Netherlands: Springer, 1993, pp. 293–314.

[26] K. Batselier and N. Wong, "A constructive arbitrary-degree kronecker product decomposition of tensors," *Numer. Linear Algebra with Appl.*, vol. 24, no. 5, p. e2097, Oct. 2017.

[27] J. M. P. Nascimento and J. M. B. 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.