A Data Dependent Multiscale Model for Hyperspectral Unmixing With Spectral Variability

Ricardo Augusto Borsoi, Tales Imbiriba, Member, IEEE, José Carlos Moreira Bermudez, Senior Member, IEEE

Abstract—Spectral variability in hyperspectral images can result from factors including environmental, illumination, atmospheric and temporal changes. Its occurrence may lead to the propagation of significant estimation errors in the unmixing process. To address this issue, extended linear mixing models have been proposed which lead to large scale nonsmooth ill-posed inverse problems. Furthermore, the regularization strategies used to obtain meaningful results have introduced interdependencies among abundance solutions that further increase the complexity of the resulting optimization problem. In this paper we present a novel data dependent multiscale model for hyperspectral unmixing accounting for spectral variability. The new method incorporates spatial contextual information to the abundances in the Extended Linear Mixing Model by using a multiscale transform based on superpixels. The proposed method results in a fast algorithm that solves the abundance problem only once in each scale during each iteration. Simulation results using synthetic and real images compare the performances, both in accuracy and execution time, of the proposed algorithm and other state-of-the-art solutions.

Index Terms—Hyperspectral data, spectral variability, spatial regularization, multiscale, superpixels.

I. INTRODUCTION

Hyperspectral devices acquire hundreds of contiguous reflectance samples from the observed electromagnetic spectra. This observed reflectance is often mixed at the pixel level and requires unmixing strategies to correctly unveil important information about the materials and their proportion in a target scene [1]. Hyperspectral unmixing (HU) aims at decomposing the observed reflectance in pure spectral components, i.e., endmembers, and their proportions [1], commonly referred as fractional abundances. Different models and strategies have been proposed to solve this problem [2], [3], [4]. The vast majority of methods considers the Linear Mixing Model (LMM) [1], which assumes that the observed reflectance vector (i.e. an HI pixel) can be modeled as a convex combination of the endmembers present in the scene. Although this assumption naturally leads to fast and reliable unmixing strategies, the intrinsic limitation of the LMM cannot cope with relevant non-ideal effects intrinsic to practical applications [3], [5], [6]. One such important non-ideal effect is endmember variability [4], [7].

Endmember variability can be caused by a myriad of factors including environmental conditions, illumination, atmospheric and temporal changes [4]. Its occurrence may result in significant estimation errors being propagated throughout the unmixing process [8]. The most common approaches to deal with spectral variability can be divided into three basic classes. 1) to group endmembers in variational sets, 2) to model endmembers as statistical distributions, and 3) to incorporate the variability in the mixing model, often using physically motivated concepts [7]. This work follows the third approach. Recently, [8], [9] and [10] introduced variations to the LMM to cope with spectral variability. The Perturbed LMM model (PLMM) [8] introduces an additive perturbation to the endmember matrix. Such perturbation matrix then needs to be estimated jointly with the abundances. Though the perturbation matrix can model arbitrary endmember variations, it lacks physical motivation. The Extended Linear Mixing Model (ELMM) proposed in [9] increased the flexibility of the LMM model by associating a pixel-dependent multiplicative term to each endmember. This generalization can efficiently model changes in the observed reflectances due to illumination, an important effect [9]. This model addresses a physically motivated problem, with the advantage of estimating a variability parameter vector of much lower dimension when compared with the additive perturbation matrix in PLMM. Although the ELMM performs well in situations where spectral variability is mainly caused by illumination variations, it lacks flexibility when the endmembers are subject to more complex spectral distortions. In [10] a physically-motivated generalization to the ELMM was proposed resulting in the Generalized Linear Mixing Model. The GLMM considered a three-dimensional tensor of pixel- and band-dependent multiplicative factors allowing arbitrary variations to each band of each endmember.

Though the above described models were shown to be capable to model endmember variability effects with good accuracy, their use in HU leads to severely ill-posed inverse problems, which require sound regularization strategies to yield meaningful solutions. One way to mitigate this ill-posedness is to explore spatial correlations found in typical abundance [11] and variability [9] maps. For instance, spatial information has been employed both for endmember extraction [12], [13] and for regularization in linear [14], nonlinear [15], Bayesian [11], [16], [17] and sparse [18] unmixing strategies. Total variation (TV) deserves special mention as a spatial regularization approach that promotes spatially piecewise homogeneous solutions without compromising sharp discontinuities between neighboring pixels. This property is important to handle the type of spatial correlation.
found in many hyperspectral unmixing applications [19], [20].

Although important to mitigate the ill-posedness of the inverse problem, the use of spatial regularization in spectral-variability-aware HU introduces interdependencies among abundance solutions for different image pixels. This in turn leads to intricate, large scale and computationally demanding optimization problems, whose solution requires the use of variable splitting (e.g. ADMM) techniques [8], [9], [10]. Such complexity is usually incompatible with recent demands for timely processing of vast amounts of remotely sensed data required by many modern real world applications [21], [22]. Thus, it is desirable to search for faster and lower complexity strategies that yield comparable unmixing performances.

Two recent works have proposed new regularization techniques for ill-posed HU problems aimed at avoiding the interdependency between pixels introduced by standard regularization methods. In [23] a low-rank tensor regularization strategy named ULTRA was proposed for regularizing ill-posed HU problems. Although ULTRA avoids the pixel interdependency, it requires a canonical polyadic decomposition at every algorithm iteration, which may negatively impact the complexity of the problem for large datasets.

In [24] a multiscale spatial regularization approach was proposed for sparse unmixing. The method uses a signal-adaptive spatial multiscale decomposition to break the unmixing problem down into two simpler problems, one in an approximation domain and another in the original domain. The spatial contextual information is obtained by solving an unregularized unmixing problem in the approximation domain. This information is then mapped back to the original image domain and used to regularize the original unmixing problem. The multiscale approach resulted in a fast algorithm that outperformed competing methods, both in accuracy and in execution time, and promoted piecewise homogeneity in the estimated abundances without compromising sharp discontinuities among neighboring pixels.

Motivated by the excellent results in [24], we propose in this paper a novel data dependent multiscale mixture model for use in hyperspectral unmixing accounting for spectral variability of the endmembers. The new model uses a multiscale transform to incorporate spatial contextual information into the abundances of the ELMM model. While in [24] the static endmember matrix allowed the easy separation of the estimation process in two distinct problems in different domains, the variability of the endmember matrix considered in this paper ties the abundances in the approximation and original image domains. We show, however, that under a few mild assumptions the new abundance estimation problem can also be solved separately in the two domains.

When compared with approaches that rely on standard spatial regularization strategies and on variable splitting techniques such as ADMM, the proposed strategy leads to a faster iterative algorithm that at each iteration solves the abundance problem only once in each domain. The new algorithm is named Multiscale Unmixing Algorithm Accounting for Spectral Variability (MUA-SV). Simulation results clearly show the advantage of the proposed algorithm, both in accuracy and in execution time, over the competing methods.

The paper is organized as follows. Section II briefly reviews the linear mixing models and its variants accounting for spectral variability. In Section III, we present the proposed multiscale formulation for the unmixing problem. The proposed MUA-SV algorithm is discussed in Section IV. Simulation results with synthetic and real data are presented in Section V. Section VI presents the conclusions.

II. LINEAR MIXING MODELS

The Linear Mixing Model (LMM) [1] assumes that an L-band pixel $y_n = [y_{n,1}, \ldots, y_{n,L}]^\top$ is represented as

$$y_n = Ma_n + e_n$$

subject to $1^\top a_n = 1$ and $a_n \geq 0$ (1)

where $M \in \mathbb{R}^{L \times P}$ is the endmember matrix whose columns $m_i = [m_{i,1}, \ldots, m_{i,L}]^\top$ are the spectral signatures of pure materials, $a_n = [a_{n,1}, \ldots, a_{n,P}]^\top$ is the abundance vector, $e_n \sim \mathcal{N}(0, \sigma^2_n I)$ is an additive white Gaussian noise (WGN), and $I$ is the identity matrix. The LMM assumes that the endmember spectra are fixed for all pixels $y_n$, $n = 1, \ldots, N$, in the HI. This assumption can compromise the accuracy of estimated abundances in many circumstances due to the spectral variability existing in a typical scene. Recently, variations of the LMM have been proposed to cope with the variability phenomenon. Next, we review the two most prominent models.

A. Linear Mixing Models Considering Spectral Variability

The most general form of the LMM considering spectral variability generalizes (1) to allow a different endmember matrix for each pixel. This results in the following observation model for the $n$-th pixel

$$y_n = M_n a_n + e_n, \quad n = 1, \ldots, N$$

where $M_n \in \mathbb{R}^{L \times P}$ is the $n$-th pixel endmember matrix.

The Perturbed Linear Mixing Model (PLMM) proposed in [25] models $M_n$ as a fixed matrix $M_0$ plus a pixel-dependent variation matrix that can accommodate generic spatial variations. Mathematically,

$$y_n = (M_0 + dM_n) a_n + e_n.$$  (3)

This model is not physically motivated. Hence, in most cases all elements of $dM_n$ must be included as independent variables in the solution of the ill-posed unmixing problem, making the inverse problem very hard to solve. This limitation motivated the development of simpler, physically motivated variability models.

The extended linear mixing model (ELMM) is a simpler model proposed in [9]. It incorporates a multiplicative diagonal matrix to LMM, which maintains the directional information of the reference endmembers, but allows them to be independently scaled. The ELMM is expressed as

$$y_n = M_0 \psi_n a_n + e_n,$$  (4)

where $\psi_n \in \mathbb{R}^{P \times P}$ is a diagonal matrix containing a (positive) scaling factor for each endmember. This model is a particular case of (2) that can model typical endmember variations,
such as those caused by illumination variability due to the topography of the scene [9]. If all pixels are considered, (4) can be written compactly as [9]

\[
Y = M_0(Ψ \circ A) + E
\]

(5)

where \( \circ \) is the Hadamard product, \( Y = [y_1, \ldots, y_N] \) is the matrix with all observed pixels, \( A = [a_1, \ldots, a_N] \) is the matrix of abundances, \( E = [e_1, \ldots, e_N] \) is the noise, and \( Ψ = [\text{diag}(ψ_1), \ldots, \text{diag}(ψ_N)] \) is a matrix whose \( n \)-th column contains the diagonal elements of the associated inverse problem. The optimization problem that needs to be solved using model (4) is much less ill-posed than that generated using model (3) due to the reduced number of parameters to be estimated. This simplicity is obtained at the price of reduced generality.

For both the PLMM and ELMM, the problem of estimating the fractional abundances and the spectral signatures of the endmembers was cast as a large scale, non-convex inverse problem, which was solved using variable splitting procedures [25], [9]. The computational cost of these solutions is very high, making them unsuited for processing large amounts of data. Furthermore, the introduction of \textit{a priori} information about the spatial regularity of the abundance maps, which is essential to reduce the ill-posedness of the inverse problem, results in an optimization problem that is not separable per pixel. This significantly increases the computational cost of the solution. Considering this limitation of the models described above, it is of interest to develop new mixture models that combine the generality of the endmember variability patterns that can be considered with the possibility of an efficient solution of the associated inverse problem. In the next section, we introduce a new mixture model that represents separately the image components at different scales using a data-dependent transformation learned from the observed HI \( Y \). The use of this new model results in a method that is able to provide more accurate solutions at a much lower computational cost than the existing methods.

III. A Multiscale Spatial Mixture Model

To constrain the set of possible solutions, we propose to separately represent the mixture process in two distinct image scales, namely, the coarse scale containing rough spatial structures, and the fine spatial scale containing small structures and details. By doing so, the conditions for spatial smoothness can be imposed on the relevant parameters in the much simpler coarse scale, and then be translated into the fine scale for further processing. It will become clear that encoding spatial information directly in the observation model leads to the formulation of a spatial regularization that permits numerically efficient solutions and results in improved estimation quality when compared to existing approaches.

Consider a transformation based on relevant contextual inter-pixel information present in the observed image \( Y \) to be applied to both \( Y \) and \( A \). The transformed matrices are given by

\[
Y_C = YW; \quad A_C = AW,
\]

(6)

where \( Y_C = [y_{C1}, \ldots, y_{CN}] \) ∈ \( \mathbb{R}^{L \times S} \) and \( A_C = [a_{C1}, \ldots, a_{CN}] \) ∈ \( \mathbb{R}^{P \times S} \) with \( S \ll N \) are, respectively, the HI and the abundance matrix in the coarse approximation scale, denoted by \( C \). \( W \) ∈ \( \mathbb{R}^{N \times S} \) is the transformation matrix that unveils the low-level image structures.

The spatial details of the image are represented in the detail scale, denoted by \( D \), which is obtained by computing the complement to the transformation \( W \). Mathematically,

\[
Y_D = Y(I - WW^*); \quad A_D = A(I - WW^*),
\]

(7)

where \( Y_D = [y_{D1}, \ldots, y_{DN}] \) ∈ \( \mathbb{R}^{L \times N} \) and \( A_D = [a_{D1}, \ldots, a_{DN}] \) ∈ \( \mathbb{R}^{P \times N} \) are the input image and the abundance matrix in the detail scale. Matrix \( W^* \) ∈ \( \mathbb{R}^{S \times N} \) is a conjugate transformation to \( W \), and takes the images from the coarse domain \( C \) back to the original image domain. \( Y_D \) and \( A_D \) are coarse versions of \( Y \) and \( A \) in the original image domain. Transformation \( W \) captures the spatial correlation of the input image, whereas its complement \( (I - WW^*) \) captures existing fine spatial variabilities. This way it is possible to introduce spatial correlation into the abundance map solutions by separately controlling the regularization strength in each of the scales \( C \) and \( D \). This is computationally much simpler than to use more complex penalties. By imposing a smaller penalty in the coarse scale \( C \) and a larger penalty in the details scale \( D \), we effectively favor smooth solutions to the optimization problem.

We can define a composite transformation as

\[
W = [W \quad I - WW^*],
\]

(8)

which decomposes the input image into the coarse approximation \( C \) and its complement \( D \). Note that the transformation is invertible, with a right inverse given by

\[
W^{-1} = (W^* \quad I)^T.
\]

(9)

Writing (2) for all pixels as \( Y = [M_1a_1 \ldots M_Na_N] + E \), and multiplying it from the right by \( W \) yields \( YW = [Y_C Y_D] \) with

\[
Y_C = [M_1a_1 \ldots M_Na_N]W + E_C
\]

\[
Y_D = [M_1a_1 \ldots M_Na_N](I - WW^*) + E_D
\]

(10)

where \( E_C = EW \) and \( E_D = E(I - WW^*) \).

A. The multiscale transformation

The choice of the transformation \( W \) is important for the proposed methodology to achieve a good reconstruction accuracy. Desirable features for this transform are 1) to group image pixels that are semantically similar, i.e., that belong to homogeneous regions, and 2) to respect image borders by not grouping pixels that belong to different image structures or features. Additionally, it must be fast to compute.

In [24], a superpixel decomposition of the image was considered for transformation \( W \). Superpixels satisfy the aforementioned criteria, and have recently been widely applied to hyperspectral imaging tasks, including classification [26], segmentation [27], endmember detection [28], and multiscale regularization [24]. Superpixel algorithms group image pixels into regions with contextually similar spatial information [29],
decomposing the image into a set of contiguous homogeneous regions. The sizes and regularity of the regions are controlled by adjusting a set of parameters. A particularly fast and efficient algorithm is the Simple Linear Iterative Clustering (SLIC) algorithm [29], also considered in [24]. The SLIC algorithm is an adaptation of the k-means algorithm that considers a reduced search space to lower the computational complexity, and a properly defined metric that balances spectral and spatial contributions. The superpixels (clusters) are initialized almost uniformly at low-gradient image neighborhoods to reduce the influence of noise. The number of clusters \( S \) is determined as a function of the average superpixel size defined by the user. The clustering employs a normalized distance function that considers both spatial and spectral (color) similarities among pixels. The relative contributions of spatial and spectral components are controlled by a regularity parameter \( \gamma \). The parameter \( \gamma \) can be increased to emphasize the spatial distance and obtain more compact (lower area to perimeter ratio) superpixels, or decreased to emphasize spectral distances and yield a tighter adherence to image borders (with more irregular shapes). See Appendix B for more details.

The decomposition \( YW \) of the image \( Y \) returns a set of superpixels. The value of each superpixel is equal to the average of all original pixel values inside that superpixel region. The conjugate transform, \( Y_C W^* \), takes each superpixel in \( Y_C \) and attributes its value to all pixels of the uniform image sampling grid that lie inside its corresponding superpixel region. The successive application of both transforms, \( WW^* \), effectively consists in averaging all pixels inside each superpixel of the input image. The superpixel decomposition of the Cuprite hyperspectral image using the SLIC algorithm is illustrated in Figure 1.

![Image (bands 50, 80 and 100)](image)

Superpixels

Figure 1: Superpixel decomposition of a section of the Cuprite image using the algorithm in [29] with \( S = 5 \) and \( \gamma = 0.005 \).

### B. The optimization problem

The global optimization problem for unmixing with spectral variability can be formulated as follows

\[
\hat{A}, \hat{M}, \hat{\Psi} = \arg \min_{A, M, \Psi} \frac{1}{2} \| Y - [M_1 a_1 \ldots M_N a_N] W W^* \|_F^2 + \lambda_A R(A) + \frac{\lambda_M}{2} \sum_{n=1}^N \| M_n - M_0 \psi_n \|_F^2 + \lambda_H R(\Psi)
\]

subject to \( a_n \geq 0, 1^\top a_n = 1 \),

\[
M_n \geq 0, n = 1, \ldots, N
\]

where \( M \) is an \( L \times P \times N \) tensor containing the endmember matrices, with entries given by \( [M]_{:, :, n} = M_n \). The parameter \( \lambda_M \) allows for a trade-off between allowing matrices \( M_n \) to vary more freely or more strictly enforcing the ELMMS variability model (4). The regularization functionals \( R(A) \) and \( R(\Psi) \) incorporate prior information about the spatial smoothness of the abundance and scaling maps.

The scaling maps constraint is given by

\[
R(\Psi) = \| H_h(\Psi) \|_F^2 + \| H_v(\Psi) \|_F^2
\]

where \( H_h \) and \( H_v \) are linear operators that compute the vertical and horizontal gradients of a bi-dimensional signal, acting separately for each material. The abundance maps constraint introduces spatial regularity indirectly through transformation \( W \). The constraint is given by

\[
\begin{align*}
R(A) &= \frac{\rho}{2} \| AW \|_F^2 + \frac{1}{2} \| A(I - WW^*) \|_F^2 \\
&= \frac{\rho}{2} \| Ac \|_F^2 + \frac{1}{2} \| Ad \|_F^2,
\end{align*}
\]

where the constant \( \rho \) allows one to strengthen the regularization in the fine (spatially high-frequency) image scale \( D \), resulting in piecewise smooth abundance maps. The parameters \( \lambda_A \) and \( \lambda_H \) control the balance between the different terms in the cost function.

Writing the optimization problem (11) explicitly as a function of the multiscale transformation \( \mathcal{V} \) yields

\[
\begin{align*}
\hat{A}, \hat{M}, \hat{\Psi} &= \arg \min_{A, M, \Psi} \frac{1}{2} \| Y_C W^* - [M_1 a_1 \ldots M_N a_N] WW^* \|_F^2 + \frac{1}{2} \| Y_D - [M_1 a_1 \ldots M_N a_N] (I - WW^*) \|_F^2 \\
&+ 2 \text{tr} \left( (Y_C W^* - [M_1 a_1 \ldots M_N a_N] WW^*) \times (Y_D - [M_1 a_1 \ldots M_N a_N] (I - WW^*)) \right) \\
&+ \lambda_M \sum_{n=1}^N \| M_n - M_0 \psi_n \|_F^2 + \lambda_H R(\Psi) + \lambda_A R(A) \\
\text{subject to } a_n &\geq 0, 1^\top a_n = 1, \\
M_n &\geq 0, n = 1, \ldots, N
\end{align*}
\]

where \( \text{tr}(\cdot) \) is the matrix trace operator.

This problem can be solved using an Alternating Least Squares method, which optimizes (14) iteratively with respect to one variable at a time. However, (14) imposes interdependency among the different pixels of \( A \), especially if TV [19] is employed in the regularization term \( R(A) \). Next, we present a multiscale formulation that eliminates this interdependency allowing the solution to be computed faster and in parallel.

### IV. MUA-SV ALGORITHM FOR SOLVING (11)

We propose to solve (11) using an alternating minimization strategy similar to that of [9]. We minimize the cost function successively with respect to each of the variables \( M, A \) and \( \Psi \).
Each of these steps consists of solving a convex optimization sub-problem with cost functions \( J_M, J_A \) and \( J_\Psi \), which can be solved efficiently under the proposed formulation. The global MUA-SV algorithm is presented in Algorithm 1.

### Algorithm 1: Global MUA-SV algorithm

**Input**: Image \( Y \), parameters \( \lambda_M, \lambda_A, \lambda_\Psi, \rho \) and matrices \( A^{(0)}, \Psi^{(0)} \) and \( M_0 \).

**Output**: Estimated matrices \( \hat{A}, \hat{\Psi} \) and tensor \( \hat{M} \).

1. Compute the superpixel decomposition of \( HI \) and the corresponding transformation matrices \( W, W^* \), \( \mathcal{W} \) and \( \mathcal{W}^\dagger \) using the SLIC algorithm [29];
2. Compute the decomposition of the \( HI \) into approximation and detail domains \( Y_C \) and \( Y_D \) using (6) and (7);
3. Set \( i = 1 \);
4. while stopping criterion is not satisfied do
   5. \( \hat{M}^{(i)} = \arg\min_{M} J_M(A^{(i-1)}, \Psi^{(i-1)}) \);
   6. \( A^{(i)} = \arg\min_{A} J_A(\hat{M}^{(i)}) \);
   7. \( \Psi^{(i)} = \arg\min_{\Psi} J_\Psi(\hat{M}^{(i)}) \);
   8. \( i = i + 1 \);
5. return \( \hat{A} = A^{(i-1)}, \hat{M} = M^{(i-1)}, \hat{\Psi} = \Psi^{(i-1)} \);

**A. Optimizing with respect to \( M \)**

The cost function \( J_M(A^{(i-1)}, \Psi^{(i-1)}) \) consists of those terms in (11) that depend on \( M \), and is given by

\[
J_M(A, \Psi) = \frac{1}{2} \sum_{n=1}^{N} \left( \|y_n - M_n a_n\|_2^2 + \lambda_M \|M_n - M_0 \Psi_n\|_F^2 \right)
\]

subject to \( M_n \geq 0, n = 1, \ldots, N \) (15)

Similarly to [9], we compute an approximate solution to minimize (15) (line 5 of Algorithm 1) for each image pixel as

\[
\hat{M}_n = \mathcal{P}_+ \left( (y_n a_n^\top + \lambda_M M_0 \Psi_n) (a_n a_n^\top + \lambda_M I_p)^{-1} \right)
\]

where \( \mathcal{P}_+ (\cdot) \) is an operator that projects each element of a matrix to the nonnegative orthant by thresholding any negative element to zero.

**B. Optimizing with respect to \( \Psi \)**

The cost function \( J_\Psi(\hat{M}^{(i)}) \) consists of those terms in (11) that depend on \( \Psi \), and is given by

\[
J_\Psi(M) = \frac{\lambda_\Psi}{2} \sum_{n=1}^{N} \|M_n - M_0 \Psi_n\|_F^2 + \lambda_\Psi R(\Psi)
\]

We follow the approach detailed in [9, Eqs. (20)-(23)] to minimize (17).

**C. Optimizing with respect to \( A \)**

Optimization with respect to \( A \) consists of minimizing those terms in (14) that depend only on \( A \), given by

\[
J_A(A, \hat{M}) = \frac{1}{2} \|Y_C W^* - [M_1 a_1 \ldots M_N a_N] W W^*\|_F^2
\]

\[
+ \frac{1}{2} \|Y_D - [M_1 a_1 \ldots M_N a_N] (I - W W^*)\|_F^2
\]

\[
+ 2 \rho \{ (Y_C W^* - [M_1 a_1 \ldots M_N a_N] W W^*) \cdot (Y_D - [M_1 a_1 \ldots M_N a_N] (I - W W^*)) \}
\]

\[
+ \frac{\rho \lambda_A}{2} \|A\|_2^2 + \frac{\lambda_A}{2} \|A_D\|_F^2
\]

(18)

subject to \( A \geq 0, 1^\top A = 1^\top \)

The cost function (18) is not separable with respect to the abundance matrices \( A_C \) and \( A_D \) in the coarse and detail scales. This can severely impact the required computational load and the convergence time to a meaningful result. To mitigate this issue, we make few simplifying assumptions that turn the minimization of (18) into a separable optimization problem.

1) **Simplifying assumptions**: The following assumptions lead to an approximate problem to (18) which is separable with respect to \( A_C \) and \( A_D \).

**A1 - Zero-mean, uncorrelated residuals**: We define residuals as the \( HI \) reconstruction errors in each scale \( C \) and \( D \). Hence,

\[
RE_C = Y_C W^* - [M_1 a_1 \ldots M_N a_N] W W^*
\]

\[
RE_D = Y_D - [M_1 a_1 \ldots M_N a_N] (I - W W^*)
\]

We assume that for \( A \) a critical point of (18), \( RE_C \) and \( RE_D \) are spatially zero-mean and uncorrelated across scales. This is reasonable if (4) represents the data with reasonable accuracy, in which case the main contribution towards the residual error comes from the observation noise \( e_n \), which is white and spatially uncorrelated.

**A2 - Spatially smooth endmember signatures**: We assume that the pixel-by-pixel endmember signatures \( M_n \) are similar in small, compact spatial neighborhoods. More precisely, if \( N \) is a set of pixels comprising a compact spatial neighborhood, we assume that the endmember signature of any pixel in \( N \) does not deviate significantly from the average signature, so that the quantity

\[
\frac{1}{|N|} \sum_{n \in N} M_n
\]

is small for all \( j \in N \), where \( |N| \) is the cardinality of \( N \).

**A3 - Homogeneity of the superpixel sizes**: We assume that the number of pixels in each superpixel, denoted by \( \Omega(s) \), \( s = 1, \ldots, S \) does not vary significantly from the average superpixel size \( N/S \). This property is expected from superpixel algorithms since it is a general requirement that superpixels should be compact (except for image boundaries) and placed regularly across the image [30]. Furthermore, the superpixel decomposition performed by the SLIC algorithm consists of a clustering procedure based on both spatial and spectral distances. Therefore, if the contribution of the spatial distance in the superpixel clustering metric is not excessively small, it is safe to assume that each \( \Omega(s) \) does not deviate significantly from the average superpixel size \( N/S \) for any \( s = 1, \ldots, S \).
2) Residuals inner product: The third term in the cost function in optimization problem (14) consists of the inner product \( \langle R E_c, R E_D \rangle \) between the reconstruction error/residuals at the coarse and detail scales. If A1 is satisfied, then this term can be neglected when compared to the first two terms without significantly altering the critical point.

3) Approximate Mixture Model: Consider (18) after neglecting its third term. Both \( W \) (in the first term) and \( I - WW^* \) (in the second term) act upon all the products \( M_n a_n \) instead of just upon \( a_n \), for \( n = 1, \ldots, N \). This precludes the separation of (18) in a sum of non-negative functions exclusively dependent on \( A_C \) or \( A_D \), which could be independently minimized. However, combining A2 and the fact that transformation \( W \) groups pixels that are in spatially adjacent regions, we now propose an approximate separable mixing model.

We initially express each pixel \( y_{ci} \) and \( y_{Di} \) of (10) as

\[
y_{ci} = \sum_{j=1}^{N} W_{j,i} M_j a_j + e_{ci} \tag{20}
\]

and

\[
y_{Di} = M_i a_i - \sum_{j=1}^{S} \sum_{l=1}^{N} W_{j,i} W_{l,j} M_l a_l + e_{Di} \tag{21}
\]

where \( W_{j,i} \) and \( W_{j,i}^* \) are the \((j, i)\)-th elements of \( W \) and \( W^* \), respectively. Then, using A2 and the fact that \( W \) is a localized transform, we approximate every endmember matrix \( M_j \) in (20) by

\[
M_j \approx M_{ci} = \sum_{i=1}^{N} \frac{1_{W_{j,i}}}{|\text{supp}(W_{j,i})|} M_i
\tag{22}
\]

where \( 1_{W_{j,i}} \) is the indicator function of \( W_{j,i} \) (i.e. \( 1_{W_{j,i}} = 1 \) if \( W_{j,i} \neq 0 \) and \( 1_{W_{j,i}} = 0 \) otherwise), and \( |\text{supp}(f)| \) denotes the cardinality of the support of \( f \) as a function of \( l \).

Equivalently, we approximate every matrix \( M_j \) in (21) by

\[
M_i \approx M_{ci} = \sum_{n=1}^{S} \sum_{m=1}^{N} \frac{1_{W_{n,m}}}{|\text{supp}(W_{n,m})|} M_{m,n} \tag{23}
\]

where \( |\text{supp}(m,n)(f)| \) denotes the cardinality of the support of \( f \) as a function of both \( m \) and \( n \). Thus, (20) and (21) can be approximated as (details in Appendix A)

\[
y_{ci} \approx M_{ci} a_{ci} + e_{ci} \tag{24}
\]

and

\[
y_{Di} \approx M_i a_i + M_{Di, [A_C W^*]}^* + e_{Di} \tag{25}
\]

where \( [\cdot] \) denotes the \( i \)-th column of a matrix, and \( M_{Di} = M_j - M_{ci} \) reflects the variability of \( M_i \) with respect to \( M_{ci} \), the average endmember matrix of its neighborhood. According to A2, \( M_{Di} \approx 0 \). Note that, since \( W \) is a localized transform, we average \( a_n \) and \( M_n \) only in small spatial neighborhoods where their variability is small.

Selecting \( W \) and \( W^* \) according to the superpixels decomposition, we have that:

- \( M_{ci} \) is the average of all \( M_j \) inside the \( i \)-th superpixel.
- \( M_{ci}^* \) is the average of all \( M_j \) inside the \( i \)-th superpixel.

Note that \( W^* \) is also a localized transform, as it attributes the superpixel value to all pixels in the original domain that lie inside that superpixel, which encompasses a compact spatial neighborhood.

Writing (24) and (25) for all pixels, we write (20) and (21) in the matrix form as:

\[
Y_c = [M_C, a_{c1}, \ldots, M_C a_{cS}] + E_c
\]

\[
Y_D = [M_{D1}, [A_C W^*]_1, \ldots, M_{DN}, [A_C W^*]_N] + [M_1 a_{d1}, \ldots, M_N a_{dN}] + E_D \tag{26}
\]

where \( E_c \) and \( E_D \) include additive noise and modeling errors.

4) Abundance constraints: The two constraints in (18) are functions of \( A \), and thus must be considered in the optimization with respect to \( A_C \) and \( A_D \). Assuming \( W \) in (8) to be of full row rank, the sum-to-one constraint can be expressed as

\[
1^T A W = 1^T W \quad \iff \quad 1^T A_C = 1^T W , \quad 1^T A_D = 1^T (I - WW^*) . \tag{27}
\]

Considering the positivity constraint we have

\[
A \geq 0 \Rightarrow A W W^\dagger \geq 0 \quad \iff \quad [A_C, A_D] W^\dagger \geq 0 \iff A_C W^* + A_D \geq 0 . \tag{28}
\]

If \( W^* \geq 0 \), which is true if \( W \) is selected as the superpixel decomposition, we can further state that

\[
A_C W^* \geq 0 \iff A_C \geq 0 \tag{29}
\]

what simplifies the constraint by removing possible interdependencies between different pixels, and makes the problem separable for all pixels in the coarse scale \( C \).

5) Reformulating the optimization problem: Using the results obtained in Sections IV-C2 to IV-C4, minimizing (18) with respect to \( A \) can be restated as determining \( A_C \) and \( A_D \) that minimize

\[
\mathcal{J}(A_C, A_D) = \frac{1}{2} \| Y_c - [M_C, a_{c1}, \ldots, M_C a_{cS}] W^* \|_F^2 + \frac{1}{2} \| Y_D - [M_1 a_{d1}, \ldots, M_N a_{dN}] \|_F^2 - [M_{D1}, [A_C W^*]_1, \ldots, M_{DN}, [A_C W^*]_N] \|_F^2 + \frac{\rho}{2} \| A_C \|_F^2 + \frac{\lambda}{2} \| A_D \|_F^2
\]

subject to \( A_C W^* + A_D \geq 0 \), \( 1^T A_C = 1^T W \), \( 1^T A_D = 1^T (I - WW^*) \). \tag{30}

This problem is still not separable with respect to \( A_C \) and \( A_D \) since the second term of (30) and the positivity constraint include both matrices. To obtain a separable problem we make the following considerations:

1) A2 implies that the entries of \( M_{Di} \) are small when compared to those of \( M_n \);
2) \( A2 \) also implies that the entries of \( Y_c W^* \) are usually much larger than the entries of \( Y_D \).

These considerations imply that the contribution of the terms \( M_D \) [\( A_C W^* \)] in the second term of (30) can be neglected when compared to \( Y_c W^* \). Using this approximation and (29), the optimization with respect to \( A_c \) can be stated as the minimization of

\[
\tilde{J}(A_c) = \frac{1}{2} \| Y_c W^* - [M_{C_1} a_{C_1}, \ldots, M_{C_S} a_{C_S}] W^* \|_F^2 \\
+ \frac{\rho \lambda A}{2} \| A_c \|_F^2 \\
\text{subject to } A_c \geq 0, \quad 1^T A_c = 1^T W \quad (31)
\]

For \( W \) based on the superpixel decomposition, \( W^* \) assigns to each pixel in the original image domain the value of the superpixel to which it belongs. Using this property, the cost function (31) simplifies to

\[
\tilde{J}(A_c) = \frac{1}{2} \sum_{n=1}^S \left( \Omega_s(n) \| y_{c_n} - M_{c_n} a_{c_n} \|_2^2 + \frac{\rho \lambda A}{2} \| a_{c_n} \|_2^2 \right) \\
\text{subject to } a_{c_n} \geq 0, \quad 1^T a_{c_n} = 1^T |W|_n \\
n = 1, \ldots, S 
\]

where \( |W|_n \) is the \( n \)-th column of \( W \), and \( \Omega_s(n) \) is the number of pixels in the \( n \)-th superpixel. Using this property, we can further simplify (32) by approximating \( \Omega_s(n) \) by the average superpixel size as \( \Omega_s(n) \approx N/S \), making the optimization problem easier since \( S \) is specified a priori by the user.

Doing this results in the following cost function

\[
\tilde{J}(A_c) = \frac{N^2}{2S^2} \sum_{n=1}^S \left( \| y_{c_n} - M_{c_n} a_{c_n} \|_2^2 + \frac{\rho \lambda A}{2} \| a_{c_n} \|_2^2 \right) \\
\text{subject to } a_{c_n} \geq 0, \quad 1^T a_{c_n} = 1^T |W|_n \\
n = 1, \ldots, S. 
\]

Note that this cost function is equivalent to

\[
\tilde{J}(A_c) = 2 \frac{1}{2} \sum_{n=1}^S \left( \left[ y_{c_n} \right] - \left[ \begin{array}{c} M_{c_n} \\ 0 \end{array} \right] \right) \left( \begin{array}{c} \frac{M_{c_n}}{(S/N)\sqrt{\rho \lambda A I_P}} \\ 0 \end{array} \right) a_{c_n} \|_2^2 \\
\text{subject to } a_{c_n} \geq 0, \quad 1^T a_{c_n} = 1^T |W|_n \\
n = 1, \ldots, S. 
\]

Minimization of (34) is a standard FCLS problem, which can be solved efficiently.

Solving with respect to \( A_D \):

Considering the detail scale \( D \) of the image, the optimization problem with respect to \( A_D \) is obtained from (30) by selecting the terms depending on \( A_D \) as

\[
\tilde{J}(A_D) = \frac{1}{2} \| Y_D - [M_1 a_{D_1}, \ldots, M_N a_{D_N}] \\
- [M_D, [A_C W^*]_1, \ldots, M_D, [A_C W^*]_N] \|_F^2 \\
+ \frac{\lambda D}{2} \| A_D \|_F^2 \\
\text{subject to } A_c W^* + A_D \geq 0, \quad 1^T A_D = 1^T (I - WW^*). 
\]

As we are minimizing (30) with respect to \( A_C \) and \( A_D \) sequentially/iteratively, we replace matrix \( A_C \) in (35) with its current estimate, denoted by \( A_C^{(i)} \). This results in the minimization of the following cost function:

\[
\tilde{A}_C(A_D) = \frac{1}{2} \sum_{n=1}^N \left( \| y_{D_n} - M_n a_{D_n} - M_D [A_C^{(i)} W^*]_n \|_2^2 \\
+ \frac{\lambda A}{2} \| a_{D_n} \|_2^2 \right) \\
\text{subject to } [A_C^{(i)} W^*]_n + a_{D_n} \geq 0, \quad 1^T a_{D_n} = 1^T (I - WW^*)_n, \quad n = 1, \ldots, N. 
\]

This is again a standard FCLS problem, which can be solved efficiently. The algorithm for computing the solution to the \( A \) subproblem is summarized in Algorithm (2).

**Algorithm 2: Algorithm for solving subproblem \( A \)**

**Input:** \( M^{(i)}, Y_C, Y_D, W, W^*, \Psi, V^T, \lambda_A \) and \( \rho \).

**Output:** Estimated abundances for the \( i \)-th iteration \( A^{(i)} \).

1. \( A_C^{(i)} = \arg \min J(A_C^{(i)}; A_D) \);
2. \( A_D^{(i)} = \arg \min J(A_D^{(i)}; \Psi) \);
3. \( A^{(i)} = [A_C^{(i)}, A_D^{(i)}] V^T \);
4. return \( A^{(i)} \).

**V. RESULTS**

In this section, we compare the unmixing performances achieved using the proposed MUA-SV algorithm, the Fully Constrained Least Squares (FCLS), the Scaled Constrained Least Squares (SCLS), the PLMM-based solution [25] and the ELM-based solution [9], the latter two designed to tackle spectral variability. The SCLS algorithm is a particular case of the ELMM model that employs the same scaling factors \( \psi_n \) for all endmembers in each pixel (i.e. \( M_n = \psi_n M_0 \), where \( \psi_n \in \mathbb{R}_+ \)) [31]. It is a low complexity algorithm that can be used as a baseline method to account for spectral variability.

For all simulations, the reference endmember signatures \( M_0 \) were extracted from the observed image using the Vertex Component Analysis (VCA) algorithm [32]. The abundance maps were initialized with the SCLS result for all algorithms. The scaling factors \( \Psi \) for ELMM and MUA-SV were initialized with ones. The matrix \( M \) for the PLMM was initialized with the results of the VCA. The alternating least squares loop in Algorithm 1 is terminated when the norm of the relative variation of the three variables between two successive iterations is smaller than \( \epsilon_A = \epsilon_\psi = \epsilon_M = 2 \times 10^{-3} \).
Experiments were performed for three synthetic and two real data sets. For the synthetic data, the regularization parameters were selected for each algorithm to provide the best abundance estimation performance. The complete set of parameters, comprising the SLIC (\(S\) and \(\gamma\)) and the regularization parameters (\(\rho\), \(\lambda_M\), \(\lambda_A\), and \(\lambda_\psi\)), were searched in appropriate intervals. For instance, \(\gamma \in \{0.001, 0.0025, 0.005, 0.01, 0.025, 0.05\}\), \(S\) assumed an integer value in the interval [2, 9], \(\rho\) was selected so that \(\rho S^2/N^2 \in \{0.001, 0.01, 0.025, 0.05, 0.1, 0.15, 0.2, 0.25, 0.35, 0.5\}\), while \(\lambda_M\), \(\lambda_A\), and \(\lambda_\psi\) were searched in the range \([5 \times 10^{-4}, 100]\), with 12 points sampled uniformly.

### A. Synthetic data sets

Three synthetic data sets were built. The first data cube (DC1) was built from the ELMM model to verify how MUA-SV performs when the actual endmembers closely follows the adopted model. The second data cube (DC2) was built using the more challenging additive perturbation model of [25]. The third data cube (DC3) was based on a realistic simulation of endmember variability caused by illumination conditions following the Hapke’s model [33].

The data cube DC1 contains \(50 \times 50\) pixels and three materials, also randomly selected from the USGS spectral library and used as the reference endmember matrix \(M_0\), with 224 spectral bands. The abundance maps are piecewise smooth images generated by sampling a Gaussian Random field [34], and are depicted in Figure 2a. Spectral variability was added to the reference endmembers using the same model as in [9], where the endmember instances for each pixel were generated by applying a constant scaling factor to the reference endmembers with amplitude limited to the interval [0.75, 1.25]. Finally, a white Gaussian noise with a 25dB SNR was added to the already scaled endmembers. The true scaling factors applied to each endmember were generated using a Gaussian Random field, and thus exhibit spatial correlation.

The data cube DC2 contains \(70 \times 70\) pixels and three materials, also randomly selected from the USGS spectral library to compose matrix \(M_0\) with 224 spectral bands. The abundance maps (shown in Fig. 2b) are composed by square regions distributed uniformly over a background, containing pure pixels (first row) and mixtures of two and three endmembers (second and third rows). The background pixels are mixtures of the same three endmembers, with abundances 0.2744, 0.1055 and 0.62. Spectral variability was added following the model proposed in [25], which considered a per-pixel variability given by random piecewise linear functions to scale individually the spectrum of each endmember by a factor in the interval [0.8, 1.2]. Such a variability model does not match the ELMM, as it yields different variabilities across the spectral bands, and is not designed to produce spatial correlation. Nevertheless, it provides a good ground for comparison with more flexible models such as the PLMM.

The data cube DC3 contains \(50 \times 50\) pixels and three materials, and is based on a simulation originally presented in [9]. This data cube is devised to realistically represent the spectral variability introduced due to changes in the illumination conditions caused by the topography of the scene, and is generated according to a physical model proposed by Hapke [33]. Hapke’s model is able to represent the reflectance of a material as a function of its single scattering albedo, photometric parameters and geometric characteristics of the scene, namely, the incidence, emergence and azimuth angles during acquisition [9], [33]. Thus, pixel dependent reflectance signatures for each endmember can be obtained given its single scattering albedo and the scene topography.

In this example, the scene was composed of three materials, namely, basalt, palagonite and tephra, which are frequently present on small bodies of the Solar System, and contained 16 spectral bands. Afterwards, a digital terrain model simulating a hilly region was generated, which is shown in Fig. 3, and from this model the acquisition angles associated with each pixel were derived (as a function of the scene topography) by considering the angle between the sun and the horizontal plane as 18°, and the sensor to be placed vertically downward. Finally, the pixel dependent endmember signatures for the scene were generated from the single scattering albedo of the materials, and from the geometric characteristics of the scene using Hapke’s model. The abundance maps used for DC2 were the same used for DC1, as shown in Fig. 2c.

The resulting hyperspectral images for all data cubes were generated from the pixel-dependent endmember signatures and abundance maps following the LMM, and were later contaminated by white Gaussian noise, with signal-to-noise ratios (SNR) of 20, 30, and 40dB. The regularization parameters for all algorithms and all examples were selected using a grid search procedure in order to provide best abundance estimation performance, and are presented in Appendix C.

The unmixing accuracy metrics used are the abundances mean squared error (MSE)

\[
\text{MSE}_A = \frac{1}{NLP} \parallel A - \hat{A} \parallel_F^2, \quad (38)
\]

the mean squared error of the estimated spectra

\[
\text{MSE}_M = \frac{1}{NLP} \sum_{n=1}^{N} \parallel M_n - \hat{M}_n \parallel_F^2, \quad (39)
\]

and the mean squared reconstruction error

\[
\text{MSE}_Y = \frac{1}{NL} \sum_{n=1}^{N} \parallel y_n - \hat{M}_n \hat{a}_n \parallel^2. \quad (40)
\]

We also evaluate the estimates of the endmember signatures using the average Spectral Angle Mapper (SAM), defined by

\[
\text{SAM}_M = \frac{1}{N} \sum_{n=1}^{N} \sum_{k=1}^{P} \arccos \left( \frac{m_{k,n}^T \hat{m}_{k,n}}{\parallel m_{k,n} \parallel \parallel \hat{m}_{k,n} \parallel} \right). \quad (41)
\]

where \(m_{k,n}\) and \(\hat{m}_{k,n}\) are the \(k\)-th columns of \(M_n\) and \(\hat{M}_n\), respectively.

The quantitative results achieved by all algorithms are displayed in Table I for all tested SNR values. The reconstructed

---

1 Generated using the code in http://www.ehu.es/ccwintco/index.php/Hyperspectral_Imagery_Synthesis_tools_for_MATLAB

2 Most of the data for this simulation was generously provided by Lucas Dr่นetz and his collaborators.
abundance maps for the three data cubes and an SNR of 30dB are shown in Figs. 2a, 2b and 2c for a qualitative comparison.

The computational complexity of the algorithms was evaluated through their execution times, which are shown in Table II. The algorithms were implemented in Matlab on a desktop computer equipped with an Intel Core i7 4.2Ghz processor and 16Gb RAM. ELMM, PLMM and SLIC were implemented using the codes made available by the respective authors.

1) Discussion: Table I shows a significantly better MSE\(_A\) performance of MUA-SV for all three data cubes and SNR values when compared with the other algorithms. This indicates that MUA-SV effectively exploits the spatial properties of the abundance maps, even when the actual spectral variability does not follow exactly the model in (4).

Figs. 2a, 2b and 2c show the true and reconstructed abundance maps for all algorithms and 30dB SNR. As expected, models accounting for spectral variability tend to yield better reconstruction quality than FCLS, with EELMM yielding piecewise smooth solutions. In general, the solution provided by MUA-SV approaches better the ground-truth, in that it estimates the intensity of the abundance maps with better accuracy than the other algorithms. This can be most clearly seen for the results for DC2 (Fig. 2b), where the regions with pure pixels are better represented by MUA-SV.

Regarding the spectral performances, as measured by the MSE\(_S\) and SAM\(_M\), the results varied among the algorithms, with no method performing uniformly better than the others. There is also a significant discrepancy between the Euclidean metric and the spectral angle in many examples, highlighting the different characteristics of the two metrics.

The ELMM model yielded the smallest reconstruction error MSE\(_Y\) in most cases (6), followed by MUA-SV (4 cases). However, the large discrepancies between the MSE\(_Y\) and the MSE\(_A\) results among the methods that address spectral variability indicates that there is no clear relationship between these two variables.

The execution times shown in Table II indicate that MUA-SV is 2.2 times faster than ELMM and 7.5 times faster than PLMM, a significant gain in computational efficiency. This difference is more accentuated when processing larger datasets, as will be verified in the following.

B. Sensitivity analysis

To evaluate the sensitivity of MUA-SV MSE\(_A\) to variations in the algorithm parameters\(^3\), we initially set all regularization parameters (\(\lambda_M\), \(\lambda_A\), \(\lambda_\Psi\) and \(\rho\)) equal to their optimal values\(^4\). Then, we varied one parameter at a time within a range from \(-80\%\) to \(+80\%\) of its optimal value. Fig. 4 presents the MSE\(_A\) values obtained by varying each parameter. It can be seen that small variations about the optimal values do not affect the MSE\(_A\) significantly, and that the maximum values obtained for the whole parameter ranges tested are still lower than those achieved by the other algorithms.

To evaluate the sensitivity of the MUA-SV results to variations in the SLIC parameters, we plotted the resulting

\(^3\)For conciseness, we present only the results for the DC1 data cube with a 30dB SNR. The results for the other data cubes and SNRs are available in Appendix D, and corroborate the conclusions presented in this section.

\(^4\)Note that the operating point of MUA-SV is not optimal for this case due to the relatively coarse grid employed in the parameter search procedure.
Table I: Quantitative results of all algorithms for data cubes DC1, DC2 and DC3 (with parameters selected to yield best abundance estimates). All values are multiplied by 10³.

| SNR  | Method | MSE_A | MSE_est | SAM_est | MSE_Y | MSE_A | MSE_est | SAM_est | MSE_Y | MSE_A | MSE_est | SAM_est | MSE_Y |
|------|--------|-------|---------|---------|-------|-------|---------|---------|-------|-------|---------|---------|-------|
| 20dB | FCLS   | 21.97 | x       | x       | 6.91  | 66.47 | x       | x       | 6.45  | 74.14 | x       | x       | 2.63  |
|      | SCLS   | 28.79 | 6.87    | 190.50  | 6.86  | 73.35 | 4.07    | 171.01  | 6.20  | 73.18 | 3.02    | 214.56  | 0.50  |
|      | PLMM   | 24.64 | 5.42    | 188.80  | 3.50  | 85.65 | 3.19    | 174.35  | 3.33  | 39.07 | 1.44    | 122.66  | 0.39  |
|      | ELMM   | 17.61 | 5.42    | 196.70  | 5.59  | 65.11 | 3.09    | 178.05  | 6.69  | 59.54 | 2.80    | 317.44  | 0.0001|
|      | MUA-SV | 12.90 | 5.24    | 212.20  | 1.56  | 29.80 | 3.36    | 185.67  | 3.28  | 28.11 | 1.84    | 308.57  | 0.0002|
| 30dB | FCLS   | 28.10 | x       | x       | 1.76  | 60.28 | x       | x       | 0.93  | 172.31| x       | x       | 1.41  |
|      | SCLS   | 12.37 | 4.53    | 187.60  | 1.63  | 62.23 | 3.84    | 161.20  | 0.71  | 21.41 | 2.42    | 68.73   | 0.05  |
|      | PLMM   | 19.61 | 4.88    | 173.00  | 0.86  | 49.38 | 3.95    | 162.54  | 0.41  | 38.00 | 1.53    | 65.38   | 0.10  |
|      | ELMM   | 10.71 | 3.70    | 170.20  | 0.59  | 40.16 | 3.05    | 177.91  | 0.001 | 18.47 | 1.73    | 101.51  | 0.00002|
|      | MUA-SV | 7.07  | 3.46    | 166.90  | 0.35  | 24.30 | 2.83    | 161.52  | 0.33  | 14.70 | 1.75    | 68.62   | 0.07  |
| 40dB | FCLS   | 20.04 | x       | x       | 1.23  | 71.37 | x       | x       | 0.44  | 256.20| x       | x       | 1.39  |
|      | SCLS   | 7.38  | 3.88    | 186.30  | 1.10  | 69.48 | 3.52    | 160.10  | 0.17  | 8.98  | 2.40    | 30.90   | 0.01  |
|      | PLMM   | 13.44 | 3.64    | 170.30  | 0.56  | 44.73 | 3.02    | 140.74  | 0.11  | 34.38 | 1.47    | 74.15   | 0.08  |
|      | ELMM   | 5.36  | 2.51    | 149.70  | 0.02  | 46.83 | 2.63    | 159.21  | 0.0002| 8.12  | 1.28    | 43.14   | 0.01  |
|      | MUA-SV | 3.98  | 2.52    | 149.90  | 0.02  | 26.01 | 2.97    | 155.96  | 0.31  | 7.94  | 1.81    | 30.66   | 0.02  |

Table II: Execution time (in seconds) of the unmixing algorithms, averaged for all SNR values considered.

|         | FCLS | SCLS | PLMM | MUA-SV |
|---------|------|------|------|--------|
| DC1     | 0.14 | 0.42 | 14.76| 2.57   |
| DC2     | 0.27 | 0.83 | 37.52| 18.29  |
| DC3     | 0.17 | 0.35 | 15.82| 9.59   |
| Houston | 0.82 | 2.31 | 174.53| 36.29 |
| Cuprite | 6.63 | 15.61| 527.89| 95.54 |

Table III: Reconstruction errors (MSE_Y) for the Houston and Cuprite data sets (all values are multiplied by 10³).

|         | FCLS | SCLS | PLMM | MUA-SV |
|---------|------|------|------|--------|
| Houston | 2.283| 0.037| 0.010| 0.014  |
| Cuprite | 0.050| 0.044| 0.040| 0.079  |

MSE_A as a function of √N/S and γ, with the algorithm parameters λ_M, λ_A, λ_Ψ and ρ fixed at their optimal values. The results are shown in Fig. 4. It is seen that the MSE_A performance does not deviate significantly from its optimal value unless the superpixel size √N/S becomes too large. This is expected since very large superpixels may contain semantically different pixels, hindering the capability of the transform W to adequately capture coarse scale information. Furthermore, large values of √N/S may violate hypothesis A2, which has been used thoroughly in the derivation of the MUA-SV algorithm, and thus represent a bad design choice.

C. Simulations with real images

In this experiment, we consider two data sets obtained from real hyperspectral images. The first data set is comprised of a 152×108 pixels subset of the Houston hyperspectral image, with 144 spectral bands. The second data set is a 250×191 pixels subregion of the Cuprite image, with 188 spectral bands. Spectral bands presenting water absorption and low SNR were removed from both images. The parameters of the algorithms are shown in Appendix C. They were selected empirically.
for the proposed method, and set identically to those reported in [9] for the ELMM and PLMM. The number of endmembers was selected as \( P = 4 \) for the Houston data set, and as \( P = 14 \) for the Cuprite data set, following the observations in [9]. The endmembers were extracted using the VCA algorithm [32].

Since the true abundance maps are unavailable for those hyperspectral images, we make a qualitative assessment of the recovered abundance maps based on knowledge of materials present in prominent fashion in those scenes. The reconstructed abundance maps for the Houston data set are depicted in Fig. 5. The four materials which are prominently present in this dataset are vegetation, red metallic roofs, concrete stands, and asphalt. It can be seen that ELMM and MUA-SV yield the best results for the overall abundances of all materials, with smaller proportion indeterminacy in regions known to have mostly pure materials such as the football field, the square metallic roofs and the concrete stands in the stadium. However, MUA-SV provides better results, more clearly observed in the purer areas such as the concrete stands of the stadium, which appear to be more mixed with the asphalt abundances in the ELMM results. This evidences the better performance of the MUA-SV algorithm.

The reconstructed abundance maps for the Alunite, Sphene, Buddingtonite and Muscovite materials of the Cuprite data set are depicted in Fig. 6. Although all methods provide abundance maps which generally agree with previous knowledge about their distribution in this image [32], the MUA-SV results show abundances for all endmembers in Fig. 6 that are more homogeneous and clearly delineated in the regions where the materials are present. Moreover, these results show significantly smaller contributions due to outliers in the background regions of the abundance maps.

The reconstruction errors for all algorithms and both data sets are shown in Table III. For the Houston data, the ELMM and MUA-SV results are very close and significantly smaller than those of the other methods, what agrees with their better representation of the abundance maps. For the Cuprite data, the errors are small and comparable for all algorithms, except for a slightly larger PLMM error. This goes in line with the fact that the abundance maps generally agree with the known distribution of these materials in the scene. However, reconstruction error results should be taken with proper care, as observed in the examples using synthetic data. Their correlation with the quality of abundance estimation is far from straightforward.

The execution times for all methods, shown in Table II, illustrate again the significantly smaller computational load of MUA-SV when compared to other methods addressing spectral variability, as it performed, on average, 5.3 times faster than ELMM and 64.3 times faster than PLMM.

VI. CONCLUSIONS

In this paper we proposed a new data-dependent multiscale model for spectral unmixing accounting for spectral variability of the endmembers. Using a multiscale transformation based on the superpixel decomposition, spatial contextual information was incorporated into the unmixing problem through the decomposition of the observation model into two models in different domains, one capturing coarse image structures and another representing fine scale details. This facilitated the characterization of spatial regularity. Under mild assumptions, the proposed method yields a fast iterative algorithm, in which the abundance estimation problem is solved only once in each scale. Simulation results with both synthetic and real data show that the proposed MUA-SV algorithm outperforms other methods addressing spectral variability, both in accuracy of the reconstructed abundance maps and in computational complexity.

APPENDIX A

DERIVATION OF THE APPROXIMATED MIXING MODEL

Given the coarse pixel model in (20) can be approximated using hypothesis A2 as

\[
y_c \approx \sum_{l=1}^{N} \frac{1}{\text{supp}(W_{l,i})} M_{l} \sum_{j=1}^{N} W_{j,i} a_i + e_{c_i}
\]

\[
= \sum_{l=1}^{N} \frac{1}{\text{supp}(W_{l,i})} M_{l} a_{c_i} + e_{c_i}
\]

\[
= M_{c_i} a_{c_i} + e_{c_i}
\]

(42)

where \( a_{c_i} = \sum_{j=1}^{N} W_{j,i} a_i \). The detail model in (21) can be approximated as

\[
y_{D_i} = M_{f} a_i - \sum_{j=1}^{S} \sum_{l=1}^{N} W_{j,i}^* W_{l,i} M_{l} a_l + e_{D_i}
\]

\[
\approx M_{f} a_i - \left( \sum_{n=1}^{S} \sum_{m=1}^{N} \frac{1}{\text{supp}(W_{n,i}^* W_{m,i})} M_{m} \right) \sum_{j=1}^{S} \sum_{l=1}^{N} W_{j,i}^* W_{l,i} a_l + e_{D_i}
\]
\[ y_{D_i} \approx M_i a_i - M_{C_i} \sum_{j=1}^{S} \sum_{l=1}^{N} W_{j,l}^* W_{l,j} a_i + \epsilon_{D_i} \]

and straightforward computations leads to

\[ y_{D_i} = M_i (a_{D_i} + \sum_{j=1}^{S} W_{j,i}^* \alpha_{C_j}) - M_{C_i} \sum_{j=1}^{S} W_{j,i}^* \alpha_{C_j} + \epsilon_{D_i} \]

\[ = M_i a_{D_i} + \left( M_i - M_{C_i} \right) \sum_{j=1}^{S} W_{j,i}^* \alpha_{C_j} + \epsilon_{D_i} \]

\[ = M_i a_{D_i} + M_{D_i} \left[ \mathbf{A} \mathbf{W}^* \right]_{i} + \epsilon_{D_i} \]

where \( a_{D_i} = a_i - \sum_{j=1}^{S} W_{j,i}^* \alpha_{C_j} \).

**APPENDIX B**

**SLIC SUPERPIXELS FOR HIs**

The SLIC superpixel decomposition consists of an extension of the k-means algorithm, with properly initialized cluster centers and a suitable distance function, defined as [29]

\[ D_{SLIC} = \sqrt{d_{\text{spectral}}^2 + \gamma^2 d_{\text{spatial}}^2 S/N} \]

where \( d_{\text{spatial}} \) and \( d_{\text{spectral}} \) are the spatial and spectral distances, respectively. Although the SLIC algorithm was initially designed to work with color (3 bands) images, it can be extended to HIs straightforwardly by considering \( d_{\text{spectral}} \) to be the Euclidean distance between reflectance vectors (HI pixels) and adjusting the normalization factor \( \gamma \) accordingly.

The superpixel transform requires the number of clusters \( S \) and their regularity \( \gamma \) as parameters to compute the transformation \( \mathbf{Y} \mathbf{W} \). Nevertheless, we found that it is often easier to design the transform using the parameter \( \sqrt{N/S} \) instead of \( S \) since it is invariant to the image size and corresponds to the average sampling interval in the irregular domain. This quantity changes only slightly on a relatively short interval between the different simulations.

**APPENDIX C**

**PARAMETER SELECTION**

For the synthetic data, the parameters were selected by exhaustive search within the range of values used by the respective authors in the original papers, aiming at achieving the minimum MSE for the reconstructed abundances. They are depicted in Table IV for all data cubes and all SNRs. For the real data, the parameters for the MUA-SV were selected in order to produce coherent abundance maps. For the other methods the parameters were extracted from [9]. All parameters used with real data simulations are displayed in Table V.

**APPENDIX D**

**SENSITIVITY ANALYSIS**

The simulations discussed in Section V.B in the manuscript are replicated here for all datasets DC1, DC2 and DC3, and all SNR values 20, 30 and 40dB. Figures 7, 8 and 9 present the sensitivity for the DC1 data cube, Figures 10, 11 and 12 for the DC2 data cube, and Figures 13, 14 and 15 for the DC3 data cube, for SNRs of 20, 30 and 40dB respectively. The results corroborate the discussion presented in Section V.B of the manuscript.
Table IV: Parameters of the algorithms used for unmixing data cubes DC1, DC2 and DC3 (selected in order to provide the best abundance estimation performance).

| SNR | Method   | Parameters | Parameters |
|-----|----------|------------|------------|
| 20dB| FCLS     | ×          |            |
|     | SCLS     | ×          |            |
|     | PLMM     | α = 0.01, β = 1000, γ = 1.5 | λ_M = 5, λ_Ψ = 0.0005, λ_A = 0.5 |
|     | MUA-SV   | λ_M = 0.5, λ_Ψ = 10, λ_A = 1, ρS^2/N^2 = 0.5, √N/S = 5, γ = 0.005 | |
| 30dB| FCLS     | ×          |            |
|     | SCLS     | ×          |            |
|     | PLMM     | α = 0.01, β = 10^4, γ = 1.5 | λ_M = 1, λ_Ψ = 0.5, λ_A = 0.01 |
|     | MUA-SV   | λ_M = 0.5, λ_Ψ = 0.5, λ_A = 1, ρS^2/N^2 = 0.1, √N/S = 3, γ = 0.001 | |
| 40dB| FCLS     | ×          |            |
|     | SCLS     | ×          |            |
|     | PLMM     | α = 0.01, β = 1000, γ = 1.5 | λ_M = 0.1, λ_Ψ = 0.5, λ_A = 0.005 |
|     | MUA-SV   | λ_M = 0.1, λ_Ψ = 1, λ_A = 0.5, ρS^2/N^2 = 0.1, √N/S = 3, γ = 0.01 | |

DC2 data cube

| SNR | Method   | Parameters | Parameters |
|-----|----------|------------|------------|
| 20dB| FCLS     | ×          |            |
|     | SCLS     | ×          |            |
|     | PLMM     | α = 0.000005, β = 1000, γ = 1.5 | λ_M = 50, λ_Ψ = 0.5, λ_A = 0.5 |
|     | MUA-SV   | λ_M = 1, λ_Ψ = 50, λ_A = 100, ρS^2/N^2 = 0.005, √N/S = 9, γ = 0.01 | |
| 30dB| FCLS     | ×          |            |
|     | SCLS     | ×          |            |
|     | PLMM     | α = 0.01, β = 0.05, γ = 1.5 | λ_M = 0.005, λ_Ψ = 0.5, λ_A = 0.05 |
|     | MUA-SV   | λ_M = 0.5, λ_Ψ = 50, λ_A = 50, ρS^2/N^2 = 0.005, √N/S = 9, γ = 0.01 | |
| 40dB| FCLS     | ×          |            |
|     | SCLS     | ×          |            |
|     | PLMM     | α = 0.00005, β = 10, γ = 1.5 | λ_M = 0.005, λ_Ψ = 0.5, λ_A = 0.05 |
|     | MUA-SV   | λ_M = 1, λ_Ψ = 50, λ_A = 100, ρS^2/N^2 = 0.005, √N/S = 9, γ = 0.01 | |

DC3 data cube

| SNR | Method   | Parameters | Parameters |
|-----|----------|------------|------------|
| 20dB| FCLS     | ×          |            |
|     | SCLS     | ×          |            |
|     | PLMM     | α = 0.01, β = 50, γ = 1.5 | λ_M = 0.005, λ_Ψ = 0.0005, λ_A = 0.01 |
|     | MUA-SV   | λ_M = 0.01, λ_Ψ = 0.05, λ_A = 0.005, ρS^2/N^2 = 0.001, √N/S = 4, γ = 0.05 | |
| 30dB| FCLS     | ×          |            |
|     | SCLS     | ×          |            |
|     | PLMM     | α = 0.01, β = 50, γ = 1 | λ_M = 0.005, λ_Ψ = 0.01, λ_A = 0.01 |
|     | MUA-SV   | λ_M = 5, λ_Ψ = 0.05, λ_A = 0.01, ρS^2/N^2 = 0.001, √N/S = 4, γ = 0.01 | |
| 40dB| FCLS     | ×          |            |
|     | SCLS     | ×          |            |
|     | PLMM     | α = 0.01, β = 50, γ = 1 | λ_M = 0.1, λ_Ψ = 0.5, λ_A = 0.0005 |
|     | MUA-SV   | λ_M = 5, λ_Ψ = 0.01, λ_A = 0.01, ρ/√N/S = 0.001, √N/S = 2, γ = 0.0005 | |

REFERENCES
[1] N. Keshava and J. F. Mustard, “Spectral unmixing,” IEEE Signal Processing Magazine, vol. 19, no. 1, pp. 44–57, 2002.
[2] J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. Nasrabadi, and J. Chanussot, “Hyperspectral remote sensing data analysis and future challenges,” IEEE Geoscience and Remote Sensing Magazine, vol. 1, no. 2, pp. 6–36, 2013.
[3] N. Dobigeon, J.-Y. Tourneret, C. Richard, J. C. M. Bermudez, S. McLaughlin, and A. O. Hero, “Nonlinear unmixing of hyperspectral images: Models and algorithms,” IEEE Signal Processing Magazine,
Table V: Parameters of the algorithms used for unmixing Houston and Cuprite data cubes.

| Dataset | Method  | Parameters                                                                 |
|---------|---------|-----------------------------------------------------------------------------|
| Houston | FCLS    | ×                                                                           |
|         | SCLS    | ×                                                                           |
|         | PLMM    | α = 0.0014, β = 500, γ = 1, λM = 0.4, λΨ = 0.001, λA = 0.005               |
|         | ELM  |                                                                                     |
|         | MUA-SV  | λM = 0.5, λΨ = 0.001, λA = 0.001, ρS^2/N^2 = 0.35, √N/S = 5, γ = 0.001  |
| Cuprite | FCLS    | ×                                                                           |
|         | SCLS    | ×                                                                           |
|         | PLMM    | α = 0.00031, β = 500, γ = 1, λM = 0.4, λΨ = 0.005, λA = 0.005             |
|         | ELM  |                                                                                     |
|         | MUA-SV  | λM = 5, λΨ = 0.01, λA = 0.05, ρS^2/N^2 = 0.01, √N/S = 6, γ = 0.001  |

Figure 10: MSE variation due to relative changes in each parameter value about its optimal value (left) and MSE as a function of SLIC parameters √N/S and γ (right) for data cube DC2 with an SNR of 20dB.

Figure 11: MSE variation due to relative changes in each parameter value about its optimal value (left) and MSE as a function of SLIC parameters √N/S and γ (right) for data cube DC2 with an SNR of 30dB.

Figure 12: MSE variation due to relative changes in each parameter value about its optimal value (left) and MSE as a function of SLIC parameters √N/S and γ (right) for data cube DC2 with an SNR of 40dB.

Figure 13: MSE variation due to relative changes in each parameter value about its optimal value (left) and MSE as a function of SLIC parameters √N/S and γ (right) for data cube DC3 with an SNR of 20dB.

Figure 14: MSE variation due to relative changes in each parameter value about its optimal value (left) and MSE as a function of SLIC parameters √N/S and γ (right) for data cube DC3 with an SNR of 30dB.

Figure 15: MSE variation due to relative changes in each parameter value about its optimal value (left) and MSE as a function of SLIC parameters √N/S and γ (right) for data cube DC3 with an SNR of 40dB.
T. Imbiriba, R. A. Borsoi, and J.-Y. Tournéret, “Generalized linear estimation in hyperspectral imaging,” *IEEE Transactions on Image Processing*, vol. 25, no. 3, pp. 1136–1151, March 2016.

J. Chen, C. Richard, and P. Honeine, “Nonlinear estimation of material properties in hyperspectral images,” *IEEE Transactions on Image Processing*, vol. 26, no. 5, pp. 2179–2191, May 2017.

T. Imbiriba, C. J. M. Bermudez, and C. Richard, “Band selection for nonlinear unmixing of hyperspectral images as a maximal clique problem,” *IEEE Transactions on Image Processing*, vol. 26, no. 5, pp. 2179–2191, May 2017.

L. Drumetz, J. Chauussot, and C. Jutten, “Variability of the endmembers in spectral unmixing: recent advances,” in *8th IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, Los Angeles, USA, 2016.

P.-A. Thouvenin, N. Dobigeon, and J.-Y. Tournéret, “Hyperspectral unmixing with spectral variability using a perturbed linear mixing model,” *IEEE Trans. Signal Processing*, vol. 64, no. 2, pp. 525–538, Feb. 2016.

L. Drumetz, M.-A. Veganzones, S. Henrot, R. Phlypo, C. Jutten, and C. Richard, “Blind hyperspectral unmixing using an extended linear mixing model to address spectral variability,” *IEEE Transactions on Image Processing*, vol. 25, no. 8, pp. 3890–3905, 2016.

T. Imbiriba, R. A. Borsoi, and J. C. M. Bermudez, “Generalized linear mixing model accounting for endmember variability,” in *Acoustics, Speech and Signal Processing (ICASSP), 2018 IEEE International Conference on*. IEEE, 2018, pp. 1862–1866.

O. Eches, N. Dobigeon, and J.-Y. Tournéret, “Enhancing hyperspectral image unmixing with spatial correlations,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 11, pp. 4239–4247, 2011.

M. Zorotić and A. Plaza, “Spatial preprocessing for endmember extraction,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 8, pp. 2679–2693, 2009.

M. C. Torres-Madronero and M. Velez-Reyes, “Integrating spatial information in unsupervised unmixing of hyperspectral imagery using multiscale representation,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 6, pp. 1985–1993, 2014.

A. Zymnis, S.-J. Kim, J. Skaf, M. Parente, and S. Boyd, “Hyperspectral image unmixing via alternating projected subgradients,” in *Signals, Systems and Computers, 2007. ACCSC 2007. Conference Record of the Forty-First Asilomar Conference on*. IEEE, 2007, pp. 1164–1168.

J. Chen, C. Richard, and P. Honeine, “Nonlinear estimation of material abundances in hyperspectral images with $\ell_1$-norm spatial regularization,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 5, pp. 2654–2665, 2014.

P. Chen, J. D. Nelson, and J.-Y. Tournéret, “Toward a sparse bayesian markov random field approach to hyperspectral unmixing and classification,” *IEEE Transactions on Image Processing*, vol. 26, no. 1, pp. 426–438, 2017.

Y. Altmann, S. McLaughlin, and A. Hero, “Robust linear spectral unmixing using anomaly detection,” *IEEE Transactions on Computational Imaging*, vol. 1, no. 2, pp. 74–85, 2015.

M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, “Total variation spatial regularization for sparse hyperspectral unmixing,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 11, pp. 4484–4502, 2012.

C. Shi and L. Wang, “Incorporating spatial information in spectral unmixing: A review,” *Remote Sensing of Environment*, vol. 149, pp. 70–87, 2014.

M. V. Afonso, J. M. Bioucas-Dias, and M. A. Figueiredo, “An augmented lagrangian approach to the constrained optimization formulation of imaging inverse problems,” *IEEE Transactions on Image Processing*, vol. 20, no. 3, pp. 681–695, 2011.

Y. Ma, H. Wu, L. Wang, B. Huang, R. Ranjan, A. Zomaya, and W. Jie, “Remote sensing big data computing: Challenges and opportunities,” *Future Generation Computer Systems*, vol. 51, pp. 47–60, 2015.

M. Chi, A. Plaza, J. A. Benediktsson, Z. Sun, J. Shen, and Y. Zhu, “Big data for remote sensing: challenges and opportunities,” *Proceedings of the IEEE*, vol. 104, no. 11, pp. 2207–2219, 2016.

T. Imbiriba, R. A. Borsoi, and J. C. M. Bermudez, “A low-rank tensor regularization strategy for hyperspectral unmixing,” in *2018 IEEE Statistical Signal Processing Workshop (SSP)*, pp. 373–377.

R. A. Borsoi, T. Imbiriba, J. C. Moreira Bermudez, and C. Richard, “Tech Report: A Fast Multiscale Spatial Regularization for Sparse Hyperspectral Unmixing,” *ArXiv e-prints*, Dec. 2017.