FAST HYPERSPECTRAL UNMIXING IN PRESENCE OF SPARSE MULTIPLE SCATTERING NONLINEARITIES

Abderrahim Halimi[1]*, Jose Bioucas-Dias*[2]*, Nicolas Dobigeon*[3]*, Gerald S. Buller*[1]*, Steve McLaughlin*[1]*,*

[(1)] School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh U.K.
[(2)] Instituto de Telecomunicações and Instituto Superior Técnico, Universidade de Lisboa, Portugal
[(3)] University of Toulouse, 2 rue Charles Camichel, BP 7122, 31071 Toulouse cedex 7, France

ABSTRACT

This paper presents a novel nonlinear hyperspectral mixture model and its associated supervised unmixing algorithm. The model assumes a linear mixing model corrupted by an additive term which accounts for multiple scattering nonlinearities (NL). The proposed model generalizes bilinear models by taking into account higher order interaction terms. The inference of the abundances and nonlinearity coefficients of this model is formulated as a convex optimization problem suitable for fast estimation algorithms. This formulation accounts for constraints such as the sum-to-one and non-negativity of the abundances, the non-negativity of the nonlinearity coefficients, and the spatial sparseness of the residuals. The resulting convex problem is solved using the alternating direction method of multipliers (ADMM) whose convergence is ensured theoretically. The proposed mixture model and its unmixing algorithm are validated on both synthetic and real images showing competitive results regarding the quality of the inference and the computational complexity when compared to the state-of-the-art algorithms.

***Index Terms*—** Hyperspectral imagery, collaborative sparse regression, ADMM, nonlinear unmixing, convex optimization

1. INTRODUCTION

As a result of its simplicity, the linear mixing model (LMM) is used by many of the hyperspectral unmixing algorithms presented in the literature [1]. This model assumes that each hyperspectral pixel spectrum is a mixture of several pure materials (endmembers), whose proportions are known as abundances. The supervised unmixing scenario consists then in estimating the abundances while assuming a priori known endmembers (extracted using an endmember extraction algorithm (EEA) such as vertex component analysis (VCA) [2]). The LMM is generally justified when considering flat scenes without component interactions. However, an inherent limitation of the LMM occurs in presence of volumetric scattering, terrain relief, or intimate mixtures of materials which require the definition of new sophisticated models, to take these effects into account. Nonlinear mixture models are an alternative to better account for those effects [3, 4]. We distinguish between signal processing based models and physical based models, which include the intimate mixture models [5] and those accounting for bilinear interactions [6–10]. This paper considers a physical based nonlinearity that generalizes the bilinear formulation in [11, 12] to account for multiple scattering effects.

This work was supported by the EPSRC Grants EP/J015180/1, EP/N003446/1, EP/K015338/1, and by the Portuguese Science and Technology, Project UID/EEA/50008/2013

The first contribution of this paper is the introduction of a mixture model to deal with NL due to multiple scattering. The model assumes a linear mixing model corrupted by an additive residual term [13]. The residual term is assumed to be a linear combination of high order interaction spectra. The number of possible interactions is large. However, in a given pixel, only a few are active. This parsimonious number of interactions is accounted for by assuming that the non-negative nonlinearity coefficients are sparse, so that only a few interactions are active for each pixel. In addition, the corrupted pixels are assumed spatially sparse meaning that only a small number of nonlinear pixels are present, as previously suggested in [14, 15]. This property has been introduced by considering the well known collaborative sparse regression strategy [14, 16–19] as a way of promoting group-sparsity over the residual terms while using the information of the residuals in all the pixels. A clear motivation for this new formulation is the simplification it introduces in the unmixing problem thanks to the linear expression for both the LMM term and the residual term. Note finally that the resulting formulation is general, and covers many NL models [6–10, 15].

The second contribution of this paper is the introduction of a convex formulation for unmixing the proposed observation model. The convexity is obtained thanks to the linearity of the observation model with respect to the unknown parameters, as well as to the considered regularization terms. Indeed, the formulation accounts for the known physical constraints on the estimated parameters, such as the sum-to-one and non-negativity of the abundances, the non-negativity of the nonlinearity coefficients, and the spatial sparseness of the residuals. The resulting convex problem is solved using the alternating direction method of multipliers (ADMM) whose convergence is ensured theoretically. More precisely, we propose an algorithm denoted as NUSAL-K for Nonlinear Unmixing by variable Splitting and Augmented Lagrangian with order K. Note that the ADMM algorithms are well adapted for large scale problems, i.e., with a large number of parameters to be estimated [20, 21]. Moreover, this method offers good performance at a reduced computational cost, as already shown in many hyperspectral unmixing works [18, 22]. The proposed mixture model and estimation algorithm are validated using synthetic and real hyperspectral images. The results obtained are very promising and show the potential of the proposed mixture model and associated inference algorithm with respect to the estimation quality and the computational cost.

The paper is structured as follows. Section 2 presents the proposed NL mixture model. Section 3 introduces the convex unmixing formulation and the ADMM-based optimization algorithm denoted by NUSAL-K. Section 4 analyzes the performance of the proposed algorithm when applied to a synthetic image with known ground truth. Results on a real hyperspectral image are presented in Sec-

tion 5 and conclusions and future work are reported in Section 6.

2. NONLINEAR MIXTURE MODEL

As a result of its simplicity, the LMM has been widely used in hyperspectral image analysis. However, the LMM has some limitations in the presence of multiple scattering effect. This paper deals with this issue by generalizing the observation model proposed in [11], itself inspired from the residual component analysis model described in [13]. The proposed model considers a sum of a linear model and a residual term that accounts for the multiple scattering effect. The general observation model for the $(L \times 1)$ pixel spectrum $\boldsymbol{y}_n$, where L is the number of spectral bands, is given by

$$\boldsymbol{y}_n = \boldsymbol{M}\boldsymbol{a}_n + \boldsymbol{\phi}_n^{\text{NL-}K}\left(\boldsymbol{M}, \boldsymbol{x}_n\right) + \boldsymbol{e}_n \tag{1}$$

where $\boldsymbol{a}_n = \left(a_{1,n}, \cdots, a_{R,n}\right)^T$ is an $(R \times 1)$ vector of abundances associated with the nth pixel, $\boldsymbol{M} = \left(\boldsymbol{m}_1, \cdots, \boldsymbol{m}_R\right)$ is an $(L \times R)$ fixed and assumed known endmembers matrix (e.g., extracted using an EEA), $\boldsymbol{x}_n = \left(x_n^{(1)}, \cdots, x_n^{(D_K)}\right)^T, \forall n$ is a $(D_K \times 1)$ vector of non-negative nonlinearity coefficients associated with the nth pixel, R is the number of endmembers, and $\boldsymbol{e}_n \sim \mathcal{N}\left(\boldsymbol{0}, \boldsymbol{\Sigma}\right)$ is a centered Gaussian noise. Due to physical constraints, the abundance vector $\boldsymbol{a}_n$ satisfies the abundance non-negativity constraint (ANC): $a_{r,n} \geq 0, \forall r \in \{1, \ldots, R\}, \forall n$, and abundance sum-to-one constraint (ASC): $\sum_{r=1}^{R} a_{r,n} = 1, \forall n$.

In this paper, the residual component $\boldsymbol{\phi}_n^{\text{NL-}K}$ accounts for the multiple scattering of order lower or equal to K as

$$\boldsymbol{\phi}_n^{\text{NL-}K}\left(\boldsymbol{M}, \boldsymbol{x}_n\right) = \boldsymbol{Q}^{(K)}(\boldsymbol{M})\boldsymbol{x}_n, \tag{2}$$

where $\boldsymbol{Q}^{(K)}$ is the $(L \times D_K)$ matrix gathering the interaction spectra of the form $\boldsymbol{m}_i \odot \boldsymbol{m}_j \odot \cdots \odot \boldsymbol{m}_l$, ($\odot$ denotes the Hadamard term-wise product), $D_K = \sum_{i=2}^{K} \frac{(R+i-1)!}{i!(R-1)!}$ is the number of coefficients associated with the interaction terms that have an order lower or equal to K and $x!$ denotes the factorial of x. For instance, considering only second order interaction terms (i.e., $K = 2$) leads to $D_2 = \frac{R(R+1)}{2}$, $\boldsymbol{x}_n(2) = \left(x_n^{(1,2)}, \cdots, x_n^{(R-1,R)}, x_n^{(1,1)}, \cdots, x_n^{(R,R)}\right)^T$, $\forall n$, $\boldsymbol{Q}^{(2)}(\boldsymbol{M}) = \left(\sqrt{2}\boldsymbol{m}_{1,2}, \cdots, \sqrt{2}\boldsymbol{m}_{R-1,R}, \boldsymbol{m}_{1,1}, \cdots, \boldsymbol{m}_{R,R}\right)$, and a residual term similar to [11, 12] as follows

$$\begin{aligned}\boldsymbol{\phi}_n^{\text{NL-2}}\left(\boldsymbol{M}, \boldsymbol{x}_n\right) &= \boldsymbol{Q}^{(2)}(\boldsymbol{M})\boldsymbol{x}_n(2) = \sum_{r=1}^{R} x_n^{(r,r)}\boldsymbol{m}_{r,r} \\ &+ \sum_{r=1}^{R-1}\sum_{r'=r+1}^{R} x_n^{(r,r')}\sqrt{2}\boldsymbol{m}_{r,r'}\end{aligned} \tag{3}$$

where $\boldsymbol{m}_{i,j} = \boldsymbol{m}_i \odot \boldsymbol{m}_j$, and the interaction terms are weighted by the coefficient $\sqrt{2}$ obtained by comparison with a homogeneous polynomial kernel of the 2nd degree (see [23] for more details regarding these coefficients and the construction of $\boldsymbol{Q}^{(K)}$). In what follows, and for brevity, we drop the order index (K) for general statements (related to all interaction orders) and only include it when dealing with specific orders. The model proposed in (1) reduces to the LMM for $\boldsymbol{x}_n = 0, \forall n$ and has many links to state-of-the-art models. Indeed, model (1) with $K = 2$ is similar to [11,12] and has a close relation to the RCA model [15] (as shown in [12]). Moreover, it generalizes the GBM model [7,24] by accounting for self-interaction between the endmembers, and also generalizes the PPNMM [6] by considering different weights for the bilinear terms. Overall, model (1) is of a similar polynomial form as the bilinear models (GBM [7], PPNMM [6], Nascimento [8], Fan [9], and Meganem [10] models) with the main difference due to the introduction of higher order interaction terms, and the non-negativity and sum-to-one constraints associated with each model. In contrast with the model described in [25], which accounts for all the interactions by using only one parameter, the model (1) includes a different coefficient for each interaction term, which enables analysis of the interaction between any specific physical components (i.e., availability of interaction maps).

Note that the nonlinear behavior generally affects some pixels of the image as already exploited in [14, 15], which suggest a spatial sparsity of the nonlinear pixels. Moreover, it makes sense to assume that the elements of the nonlinear vector $\boldsymbol{x}_n$ will not be active at the same time, meaning that the vector is sparse. This can be explained since the lowest order of interactions have often a higher effect [7–9] and all the interactions between endmembers are not likely to be active at the same time. These sparsity properties are of great importance and will be exploited when designing the unmixing algorithm associated with model (1) in Section 3.

3. THE UNMIXING ALGORITHM: NUSAL-K

This section introduces the unmixing algorithm used to estimate the abundances and the residual coefficients of the proposed model. To this end, we adopt an optimization approach that minimizes a regularized data fidelity cost function. More precisely, considering an independent and identically distributed (i.i.d.) Gaussian noise ($\boldsymbol{\Sigma}$ proportional to the identity matrix) in model (1) leads to the following data fidelity term

$$\mathcal{L}_{\boldsymbol{Q}}\left(\boldsymbol{Z}\right) = \frac{1}{2}||\boldsymbol{Y} - [\boldsymbol{M}, \boldsymbol{Q}]\boldsymbol{Z}||_F^2 \tag{4}$$

where $\boldsymbol{Y} = [\boldsymbol{y}_1, \cdots, \boldsymbol{y}_N]$, N is the total number of pixels, $\boldsymbol{Z} = \left[\boldsymbol{A}^\top, \boldsymbol{X}^\top\right]^\top$ is the $(R + D) \times N$ matrix gathering the $(R \times N)$ abundance matrix $\boldsymbol{A}$ and the $(D \times N)$ residual coefficients $\boldsymbol{X}$ and $||\boldsymbol{Y}||_F = \sqrt{\text{trace}\left(\boldsymbol{Y}\boldsymbol{Y}^\top\right)}$ denotes the Frobenius norm. Estimating the abundances and the residual coefficients is an ill-posed inverse problem that requires the introduction of prior knowledge (or regularization terms) about those parameters of interest. In this paper we consider two assumptions (i) the nonlinearity appears in some pixels of the image, (ii) in a nonlinear pixel, only a few interactions are active. Under these considerations, we propose to solve the following optimization problem

$$\begin{aligned}\mathcal{C}_{\text{NUSAL-}K}\left(\boldsymbol{Z}\right) = \mathcal{L}_{\boldsymbol{Q}}\left(\boldsymbol{Z}\right) + i_{\mathbb{R}_+}\left(\boldsymbol{A}\right) + i_{\left\{\boldsymbol{1}_{(1,R)}\right\}}\left(\boldsymbol{1}_{(1,R)}\boldsymbol{A}\right) \\ + \tau_1||\boldsymbol{X}||_1 + \tau_2||\boldsymbol{X}||_{2,1} + i_{\mathbb{R}_+}\left(\boldsymbol{X}\right)\end{aligned} \tag{5}$$

where $\tau_1 > 0, \tau_2 > 0$ are two regularization parameters, $i_{\mathbb{R}_+}\left(\boldsymbol{A}\right) = \sum_{n=1}^{N} i_{\mathbb{R}_+}\left(\boldsymbol{a}_n\right)$ is the indicator function that imposes the ANC ($i_{\mathbb{R}_+}\left(\boldsymbol{a}_n\right) = 0$ if $\boldsymbol{a}_n$ belongs to the non-negative orthant and $+\infty$ otherwise), $i_{\left\{\boldsymbol{1}_{(1,R)}\right\}}\left(\boldsymbol{1}_{(1,R)}\boldsymbol{A}\right) = \sum_{n=1}^{N} i_{\{1\}}\left(\boldsymbol{1}_{(1,R)}\boldsymbol{a}_n\right)$ is the indicator function that imposes the ASC to each abundance vector $\boldsymbol{a}_n$, $\boldsymbol{1}_{(i,j)}$ denotes the $i \times j$ vector of 1s. The first line of (5) is a sum of the quadratic data fidelity term associated with the Gaussian noise statistics and two convex terms imposing the abundance constraints. The second line of (5) accounts for the sparsity behavior of the residual coefficients. The first convex term $||\boldsymbol{X}||_1 = \sum_{n=1}^{N} ||\boldsymbol{x}_n||_1$ is an ℓ_1 norm that promotes element-wise sparsity on the $D \times N$ matrix $\boldsymbol{X}$. This term imposes a point-wise

sparse repartition of the active elements of $\boldsymbol{X}$. The second convex term $||\boldsymbol{X}||_{2,1} = \sum_{n=1}^{N} ||\boldsymbol{x}_n||_2 = \sum_{n=1}^{N} \sqrt{\boldsymbol{x}_n^T \boldsymbol{x}_n}$ is the ℓ_{21} mixed norm of $\boldsymbol{X}$ which promotes sparsity among the columns of $\boldsymbol{X}$, i.e., it promotes solutions of (5) with a small number of nonlinear pixels. This regularization term has received increasing interest in recent years [14, 16–19] and is known as a collaborative regularization since it uses information about the residuals in all the pixels to promote group-sparsity over the columns of $\boldsymbol{X}$. Equation (5) includes a combination of the ℓ_1 norm and the ℓ_{21} mixed norm which leads to a slightly different effect, i.e., it allows for sparse element inside the active columns of $\boldsymbol{X}$. Finally the cost function (5) is a sum of convex functions that is solved using the ADMM algorithm proposed in [20, 26] and described in the next section.

3.1. The ADMM algorithm

Consider the optimization problem

$$\underset{\boldsymbol{Z}}{\operatorname{argmin}}\, \mathcal{C}(\boldsymbol{Z}) = \underset{\boldsymbol{Z}}{\operatorname{argmin}} \sum_{j=1}^{J} g_j(\boldsymbol{H}_j \boldsymbol{Z}) \tag{6}$$

where $\boldsymbol{Z} \in \mathbb{R}^{(R+D)\times N}$, $g_j : \mathbb{R}^{p_j \times N} \to \mathbb{R}$ are closed, proper, convex functions, and $\boldsymbol{H}_j \in \mathbb{R}^{p_j \times (R+D)}$ are arbitrary matrices. After denoting $\boldsymbol{U}_j = \boldsymbol{H}_j \boldsymbol{Z} \in \mathbb{R}^{p_j \times N}$ and introducing the auxiliary variable $\boldsymbol{F}_j \in \mathbb{R}^{p_j \times N}$, the authors in [20, 26] introduced the ADMM variant summarized in Algo. 1 to solve (6) using a variable splitting and an augmented Lagrangian algorithm. This algorithm is designed to solve any sum of an ℓ_2 norm with convex functions. Moreover, [27, Theorem 1] states that Algo. 1 converges when the matrix $\boldsymbol{G} = \left[\sum_{j=1}^{J} (\boldsymbol{H}_j)^\top \boldsymbol{H}_j\right]$ has full rank, and the functions g_j are closed, proper, and convex. Under these conditions, the same theorem states that, for any $\mu > 0$, if (6) has a non-empty set of solutions, then the generated sequence $\boldsymbol{Z}^{(k)}$ converges to a solution. If (6) does not have a solution, then at least one of the sequences $\boldsymbol{U}^{(k)}$ or $\boldsymbol{F}^{(k)}$ diverges. Note that the main steps of Algo. 1, in each iteration, are the solution of a linear system of equations (line 8), the computation of the Moreau proximity operators (MPOs) [28] (line 12), and the updating of the Lagrange multipliers (line 16). Another important point to note is that the setting of μ has a strong impact on the convergence speed of the algorithm. In this paper, μ is updated using the adaptive procedure described in [18, 21], whose objective is to keep the ratio between the ADMM primal and dual residual norms within a given positive interval, as they both converge to zero. Note finally that the algorithm is stopped if the primal or dual residual norms are lower than a given threshold [21]. We refer the reader to [18, 20, 21, 26] for more details regarding the ADMM algorithm.

3.2. The NUSAL-K algorithm

This section presents the optimization problem considered for estimating the parameters of the NL model (1). Using the same notation as in (6), problem (5) can be expressed as the sum of $J = 5$ convex terms given by

$$\begin{aligned}
g_1(\boldsymbol{U}_1) &= \mathcal{L}_{\boldsymbol{Q}}(\boldsymbol{U}_1), & \boldsymbol{H}_1 &= \mathbb{I}_{(R+D_K)} \\
g_2(\boldsymbol{U}_2) &= i_{\mathbb{R}_+}(\boldsymbol{U}_2), & \boldsymbol{H}_2 &= \mathbb{I}_{(R+D_K)} \\
g_3(\boldsymbol{U}_3) &= i_{\{\boldsymbol{1}^\top\}}\left(\boldsymbol{1}^\top \boldsymbol{U}_3\right), & \boldsymbol{H}_3 &= \left[\mathbb{I}_R, \boldsymbol{0}_{(R,D_K)}\right] \\
g_4(\boldsymbol{U}_4) &= \tau_1 ||\boldsymbol{U}_4||_1, & \boldsymbol{H}_4 &= \left[\boldsymbol{0}_{(D_K,R)}, \mathbb{I}_{D_K}\right] \\
g_5(\boldsymbol{U}_5) &= \tau_2 ||\boldsymbol{U}_5||_{2,1}, & \boldsymbol{H}_5 &= \left[\boldsymbol{0}_{(D_K,R)}, \mathbb{I}_{D_K}\right]
\end{aligned} \tag{7}$$

where $\mathbb{I}_n$ denotes the $n \times n$ identity matrix and $\boldsymbol{0}_{(i,j)}$ denotes the $i \times j$ matrix of zeros. For this problem, the matrix $\boldsymbol{G}$ is given by $\boldsymbol{G} = \operatorname{diag}\left\{[3\boldsymbol{1}_{(1,R)}, 4\boldsymbol{1}_{(1,D_K)}]\right\}$ which is full rank. This matrix and the properties of $g_i, i \in \{1, \cdots, J\}$ ensure the algorithm convergence.

Algorithm 1 ADMM variant for (6)

1: Initialization
2: Initialize $\boldsymbol{U}_j^{(0)}, \boldsymbol{F}_j^{(0)}, \forall j, \mu > 0$. Set $k \leftarrow 0$, conv$\leftarrow 0$
3: **while** conv$= 0$ **do**
4: **for** j=1:J **do**
5: $\xi_j^{(k)} \leftarrow \boldsymbol{U}_j^{(k)} + \boldsymbol{F}_j^{(k)}$,
6: **end for**
7: Linear system of equations
8: $\boldsymbol{Z}^{(k+1)} \leftarrow \boldsymbol{G}^{-1} \sum_{j=1}^{J} (\boldsymbol{H}_j)^\top \xi_j^{(k)}$,
9: Moreau proximity operators
10: **for** j=1:J **do**
11: $\boldsymbol{V}_j^{(k)} \leftarrow \boldsymbol{H}_j \boldsymbol{Z}^{(k+1)} - \boldsymbol{F}_j^{(k)}$,
12: $\boldsymbol{U}_j^{(k+1)} \leftarrow \underset{\boldsymbol{U}_j}{\operatorname{argmin}} \frac{\mu}{2} ||\boldsymbol{U}_j - \boldsymbol{V}_j^{(k)}||^2 + g_j(\boldsymbol{U}_j)$,
13: **end for**
14: Update Lagrange multipliers
15: **for** j=1:J **do**
16: $\boldsymbol{F}_j^{(k+1)} \leftarrow \boldsymbol{U}_j^{(k+1)} - \boldsymbol{V}_j^{(k)}$,
17: **end for**
18: $k = k + 1$
19: **end while**

The optimization problems shown in line 12 of algo. 1 admit analytical solutions that are not presented here for brevity (see [23]). The computational complexity of Algo. 1 per iteration is $\mathcal{O}\left((R+D)^2 N\right)$, which is related to the most expensive step introduced by solving a linear system to obtain $\boldsymbol{U}_1$. Finally, it is interesting to note that the matrices to inverse involve low complexity, for instance, the matrix $\boldsymbol{G}$ in line 8 is diagonal and easy to inverse.

4. SIMULATION RESULTS ON SYNTHETIC DATA

This section evaluates the performance of the proposed NUSAL-K algorithm when considering a synthetic image with a known ground truth. The synthetic image has 100×100 pixels, $L = 207$ spectral bands, $R = 3$ endmembers extracted from the ENVI software library [29] and it has been corrupted by an i.i.d. Gaussian noise with SNR$= 25$ dB. The image has been partitioned into 4 spatial classes associated with the LMM, NL-3 model (1) (where $\boldsymbol{x}_n \sim \mathcal{N}_{(\mathbb{R}+)^D}(\boldsymbol{0}_{D,1}, 0.1\mathbb{I}_D)$), GBM (with random nonlinear coefficients in $[0.8, 1]$) and PPNMM (with $b = 0.5$), respectively. Note that the generated nonlinear coefficients $\boldsymbol{x}_n$ are not sparse, which is a challenging scenario for the NUSAL-K algorithm. The abundances have been generated uniformly in the simplex of ANC and ASC. The performance of the algorithms has been assessed in terms of abundance root mean square error $\text{RMSE}(\boldsymbol{A}) = \sqrt{\frac{1}{NR} \sum_{n=1}^{N} \|\boldsymbol{a}_n - \hat{\boldsymbol{a}}_n\|_2^2}$ and spectral angle mapper $\text{SAM} = \frac{1}{N} \sum_{n=1}^{N} \arccos\left(\frac{\hat{\boldsymbol{y}}_n^T \boldsymbol{y}_n}{\|\boldsymbol{y}_n\|_2 \|\hat{\boldsymbol{y}}_n\|_2}\right)$, where $\arccos(\cdot)$ is the inverse cosine operator and $\boldsymbol{y}_n$, $\hat{\boldsymbol{y}}_n$ denote the $\#n$th measured and estimated pixel spectra. All simulations have been implemented using MATLAB R2015a on a computer with Intel(R) Core(TM) i7-4790 CPU@3.60GHz and 32GB RAM.

The two variants NUSAL-2 and NUSAL-3 are compared with the linear unmixing SUNSAL algorithm [22], and the NL algorithms CDA-NL [11], SKhype [30], and RNMF [14]. For a fair comparison, the endmembers of these algorithms have been fixed to the actual spectra used to generate the data and the CDA-NL algorithm

Table 1. Results on the LMM-NL based synthetic image.

	RMSE				RMSE	SAM	Time
	$\mathcal{C}_1$ LMM	$\mathcal{C}_2$ NL-3	$\mathcal{C}_3$ GBM	$\mathcal{C}_4$ PPNMM			
SUNSAL	1.4	20.3	5.8	11.9	10.8	7.6	0.1
SKhype	2.2	11.7	3.0	3.9	6.0	–	466
CDA-NL	1.4	4.5	2.1	4.2	2.9	5.8	182
RNMF	1.5	12.8	2.5	5.2	6.4	6.8	110
NUSAL-2	1.4	3.9	2.0	5.0	2.8	5.8	7
NUSAL-3	1.4	2.9	2.0	4.9	2.6	5.7	19

has been used while fixing the illumination coefficient to the value #1. The regularization parameters of RNMF, and NUSAL-K have been selected to provide the best performance (in terms of abundance RMSE) when testing the following values: λ of RNMF varies in $\{0.01\lambda_0, 0.1\lambda_0, \lambda_0\}$ (where λ_0 has been suggested in [14]), and for NUSAL-K: τ_1 and τ_2 vary in $\{0.01, 0.05, 0.1\}$.

Table 1 reports the obtained results. The proposed NUSAL-2 and NUSAL-3 algorithms provide the best RMSE performance for the LMM, RCA-NL-3 and the GBM pixels. For PPNMM, the best RMSE is obtained with SKhype that is well adapted to this polynomial nonlinearity. The best overall RMSE is obtained by the NUSAL-2 and NUSAL-3 algorithms with a slightly better values for NUSAL-3 since it estimates more parameters than NUSAL-2. Except for the LMM-based algorithms, the data are well fitted by the algorithms as indicated by the values of SAM. Moreover, it is important to mention the reduced computational time of the proposed NUSAL-K algorithms. Indeed, Table 1 clearly shows that the NUSAL-2 and NUSAL-3 algorithms are faster than the NL state-of-the-art algorithms, i.e., CDA-NL, RNMF and SKhype. The table also highlights the effect of accounting for the third order nonlinear interaction terms that improve the unmixing at a price of a higher computational time. Note that the estimated nonlinearity coefficients are visually evaluated in the next section when considering a real image. Note also that additional experiments have been conducted in [23] and are not provided here for brevity.

5. RESULTS ON REAL DATA

This section illustrates the performance of the proposed algorithm when applied to a real hyperspectral image that has received much attention in the remote sensing community [7, 31]. The image was acquired over Moffett Field, CA, in 1997 by AVIRIS. It contains 100×100 pixels, $L = 152$ spectral bands (after removing water absorption bands) acquired in the interval $0.4 - 2.5\mu$m, has a spatial resolution of 100m and is mainly composed of water, soil, and vegetation (see Fig. 1 (top-left)). This image is interesting since it is known to include bilinear scattering effects [7, 11, 14] which makes it suitable for the assessment of the NUSAL-K algorithm.

Processing this image with the studied algorithms shows a better fit for the NL algorithms than the LMM-based ones (the SAM values are: 12.7 for SUNSAL, 10.7 for CDA-NL, 8.3 for RNMF, 11 for NUSAL-2, and 10.4 for NUSAL-3). Among the NL algorithms, the proposed NUSAL-2, and NUSAL-3 provided the best performance for the computational cost (177 s for SKhype, 317 s for CDA-NL, 278 s for RNMF, 13 s for NUSAL-2, and 29 s for NUSAL-3). All the algorithms generated similar abundance maps that are not shown here for space limitation. Figs. 1 (top-right) and (bottom) present the residual maps associated with the NL algorithms (NUSAL-3 provided similar results than NUSAL-2). These

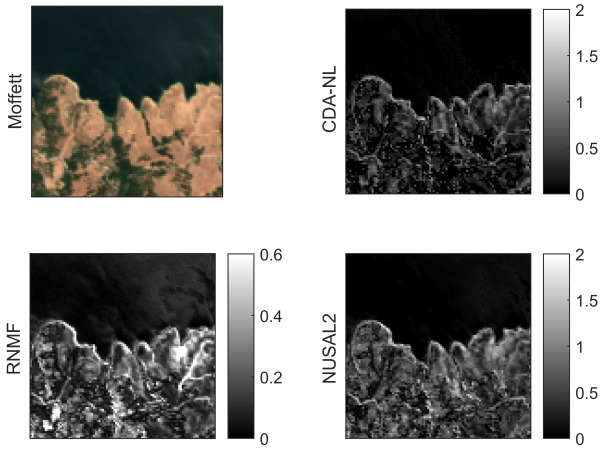

Fig. 1. (top-left) Real hyperspectral Moffett images. (top-right) and (bottom) Residual maps for the Moffett image obtained with $||\hat{\boldsymbol{y}}_n - \boldsymbol{M}\hat{\boldsymbol{a}}_n||$.

figures highlight a good agreement between the NL algorithms that detect nonlinearity in the coastal region and in presence of vegetation (as in [7]). In addition to these regions, RNMF detects other mismodelling effects probably due to endmember variability as already reported in [11]. Note that the NUSAL algorithm estimates the nonlinear coefficients associated with high order interactions. For this real image, the sparse coefficients are mainly due to the second order interactions. Indeed, the sum of the estimated second order coefficients of NUSAL-3 represents 95% of the total sum of these coefficients (including both second and third order terms). To summarize, the obtained results highlight the benefit of NUSAL-K that not only provides information regarding the high order interactions but also estimates abundances and residual maps which are in good agreement with state-of-the-art algorithms, at a lower computational cost.

6. CONCLUSIONS

This paper introduced a nonlinear model and its supervised unmixing algorithm. Nonlinearity was modeled by considering a residual term in addition to the linear mixture of endmembers. The residual term was expressed as a sparse linear combination of the interaction spectra, thus, the proposed model reduced to a linear combination with respect to the abundances and the residual coefficients. The unknown parameters associated with this model were estimated using an optimization approach that included convex regularization terms. More precisely, the non-negativity and sum-to-one constraints were imposed on the abundances and the nonlinear coefficients were assumed to be spatially sparse by considering a collaborative sparse regression approach. The resulting convex problem was solved using an alternating direction method of multipliers whose convergence was theoretically ensured. The proposed algorithm showed good performance when processing synthetic data. Results on real data confirmed the good performance of the proposed algorithm and showed its ability to extract different features in the observed scene, with a reduced computational cost. Future work includes the introduction of a model/algorithm to jointly deal with the endmember variability and the nonlinearity effects.

7. REFERENCES

[1] J. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observat. Remote Sens.*, vol. 5, no. 2, pp. 354–379, April 2012.

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

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

[4] N. Dobigeon, J.-Y. Tourneret, C. Richard, J. Bermudez, S. McLaughlin, and A. Hero, "Nonlinear unmixing of hyperspectral images: Models and algorithms," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 82–94, Jan 2014.

[5] B. W. Hapke, "Bidirectional reflectance spectroscopy. I. Theory," *J. Geophys. Res.*, vol. 86, pp. 3039–Ű3054, 1981.

[6] Y. Altmann, A. Halimi, N. Dobigeon, and J.-Y. Tourneret, "Supervised nonlinear spectral unmixing using a postnonlinear mixing model for hyperspectral imagery," *IEEE Trans. Image Process.*, vol. 21, no. 6, pp. 3017–3025, June 2012.

[7] A. Halimi, Y. Altmann, N. Dobigeon, and J.-Y. Tourneret, "Nonlinear unmixing of hyperspectral images using a generalized bilinear model," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4153–4162, 2011.

[8] J. Bioucas-Dias and J. Nascimento, "Nonlinear mixture model for hyperspectral unmixing," in *Proc. SPIE Image Signal Process. Remote Sens. XV*, L. Bruzzone, C. Notarnicola, and F. Posa, Eds., vol. 7477, no. 1. SPIE, 2009, p. 74770I.

[9] W. Fan, B. Hu, J. Miller, and M. Li, "Comparative study between a new nonlinear model and common linear model for analysing laboratory simulated-forest hyperspectral data," *Int. Journal Remote Sens.*, vol. 30, no. 11, pp. 2951–2962, June 2009.

[10] I. Meganem, P. Deliot, X. Briottet, Y. Deville, and S. Hosseini, "Linear-quadratic mixing model for reflectances in urban environments," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 544–558, Jan 2014.

[11] A. Halimi, P. Honeine, and J. Bioucas-Dias, "Hyperspectral unmixing in presence of endmember variability, nonlinearity or mismodelling effects," *IEEE Trans. Image Process.*, vol. 25, no. 10, pp. 4565–4579, Oct 2016.

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

[13] A. A. Kalaitzis and N. D. Lawrence, "Residual components analysis," in *Proc. ICML*, 2012, pp. 1–3.

[14] C. Févotte and N. Dobigeon, "Nonlinear hyperspectral unmixing with robust nonnegative matrix factorization," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 4810–4819, June 2015.

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

[16] P. Sprechmann, I. Ramirez, G. Sapiro, and Y. C. Eldar, "C-Hilasso: A collaborative hierarchical sparse modeling framework," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4183–4198, Sept 2011.

[17] H. Aggarwal and A. Majumdar, "Hyperspectral unmixing in the presence of mixed noise using joint-sparsity and total variation," *IEEE J. Sel. Topics Appl. Earth Observat. Remote Sens.*, 2016, to appear.

[18] M.-D. Iordache, J. Bioucas-Dias, and A. Plaza, "Collaborative sparse regression for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 341–354, Jan 2014.

[19] M.-D. Iordache, J. Bioucas-Dias, A. Plaza, and B. Somers, "MUSIC-CSR: Hyperspectral unmixing via multiple signal classification and collaborative sparse regression," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 7, pp. 4364–4382, July 2014.

[20] M. Afonso, J. Bioucas-Dias, and M. Figueiredo, "An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems," *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 681–695, March 2011.

[21] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jan 2011.

[22] J. Bioucas-Dias and M. Figueiredo, "Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing," in *Proc. IEEE GRSS Workshop on Hyperspectral Image and SIgnal Processing: Evolution in Remote Sensing (WHISPERS)*, June 2010, pp. 1–4.

[23] A. Halimi, J. Bioucas-Dias, N. Dobigeon, G. Buller, and S. McLaughlin, "Fast hyperspectral unmixing in presence of nonlinearity or mismodelling effects," in *ArXiv e-prints*, Jul. 2016.

[24] A. Halimi, Y. Altmann, N. Dobigeon, and J.-Y. Tourneret, "Unmixing hyperspectral images using the generalized bilinear model," in *Proc. IEEE Int. Conf. Geosci. Remote Sens. (IGARSS)*, July 2011, pp. 1886–1889.

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

[26] M. Figueiredo and J. Bioucas-Dias, "Restoration of Poissonian images using alternating direction optimization," *IEEE Trans. Image Process.*, vol. 19, no. 12, pp. 3133–3145, Dec 2010.

[27] J. Eckstein and D. P. Bertsekas, "On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators," *Math. Programm.*, vol. 55, no. 1, pp. 293–318, 1992.

[28] P. L. Combettes and J.-C. Pesquet, *Proximal Splitting Methods in Signal Processing*. New York, NY: Springer New York, 2011, pp. 185–212.

[29] RSI (Research Systems Inc.), *ENVI User's guide Version 4.0*, Boulder, CO 80301 USA, Sept. 2003.

[30] J. Chen, C. Richard, and P. Honeine, "Nonlinear unmixing of hyperspectral data based on a linear-mixture/nonlinear-fluctuation model," *IEEE Trans. Signal Process.*, vol. 61, no. 2, pp. 480–492, Jan 2013.

[31] N. Dobigeon, J.-Y. Tourneret, and C.-I Chang, "Semi-supervised linear spectral unmixing using a hierarchical Bayesian model for hyperspectral imagery," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 2684–2695, July 2008.