HYPERSPECTRAL IMAGE UNMIXING VIA SIMULTANEOUSLY SPARSE AND LOW RANK ABUNDANCE MATRIX ESTIMATION

Paris V. Giampouras, Konstantinos E. Themelis, Athanasios A. Rontogiannis, Konstantinos D. Koutroumbas

IAASARS, National Observatory of Athens, 152 36, P. Penteli, Greece
{parisg,themelis,tronto,koutroum}@noa.gr

ABSTRACT

This paper proposes a semi-supervised unmixing method, which simultaneously exploits the abundance sparsity and spatial correlation that are inherent in hyperspectral images. To capture spatial information, a sliding window containing neighboring image pixels is first defined. Then, a regularized optimization problem is defined that incorporates a mixed penalty consisting of a sparsity inducing ℓ_1 norm and a matrix rank-penalizing weighted trace norm. To solve the optimization problem, an efficient alternating direction method of multipliers based algorithm is developed. Experimental results conducted on both synthetic and real data illustrate the effectiveness of the proposed algorithm.

Index Terms— sparse and low rank matrix estimation, abundance estimation, ADMM, spectral unmixing, spatial correlation

1. INTRODUCTION

During the last decade spectral unmixing has rapidly caught the interest of the signal processing community, [1]. Spectral unmixing is the problem of decomposing the measured spectra of the pixels in a hyperspectral image to a collection of pure materials' spectra, called *endmembers*, accompanied by their corresponding proportional percentages, called *abundances*. Actually, the steps of endmember extraction and abundance estimation are considered as separate tasks in the unmixing literature, each treated by a plethora of diverse methods.

In the following we focus on the abundance estimation problem. We assume that a set of endmembers is known *a priori* and that the pixels' spectra are *linear* combinations of these endmembers. Under these assumptions, abundance estimation can be expressed as an inverse problem, subject to the physical constraint for *abundance non-negativity*, [1, 2]. Interestingly, recent developments in compressive sensing paved the way for the development of *sparse* abundance estimation methods, e.g., [2]. The rationale behind sparse methods is that only a subset of the available endmembers will contribute to the spectrum of a single pixel. Hence, the abundance coefficient vectors shall be sparse, in the sense that they shall contain only a few non-zero values. In a different research direction, abundance estimation techniques have also benefited from the exploitation of the spatial information of hyperspectral images, e.g., [3]. The underlying assumption here is that adjacent pixels shall probably share the same set of endmembers. By considering blocks of adjacent pixels, spatial correlation is reflected on the abundance fractions that form a matrix, which is likely to be joint-sparse or *low-rank*.

In this work, we aspire to unite the aforementioned directions and propose the estimation of the abundance matrix by simultaneously imposing the *sparsity* and *low-rankness* constraints on it. To this end, we formulate a regularized least squares problem, where a sparsity inducing ℓ_1-norm term is combined with a trace norm term, which is known to penalize the matrix rank, [4]. Specifically, we propose the utilization of weighted ℓ_1 and trace norms, which are known to provide more robust estimates. To solve the resulting optimization problem, we develop an alternating direction method of multipliers (ADMM), [5]. The proposed algorithm converges very fast and performs better than state-of-the-art algorithms. Its estimation performance is illustrated via simulations in both synthetic and real hyperspectral data.

2. PROBLEM FORMULATION

Let us consider a hyperspectral image consisting of L spectral bands and a square sliding window in the spatial dimension of the image that contains K neighboring pixels. The spectral measurements of the window pixels are L-dimensional vectors, denoted by $\mathbf{y}_k, k = 1, \ldots, K$. Moreover, we assume that a total of N endmembers are present in the image. Let $\boldsymbol{\phi}_i \in \mathbb{R}_+^L$ denote the spectral signature of the ith endmember, while $\boldsymbol{\Phi} = [\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \ldots, \boldsymbol{\phi}_N]$ and $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_K]$ are the matrices having as their columns the N endmembers and the K spectral measurements of the window pixels, respectively. Assuming the linear mixing model, the measured spectra of the pixels within the window are generated as

$$\mathbf{Y} = \boldsymbol{\Phi}\mathbf{W} + \mathbf{E}, \tag{1}$$

where $\mathbf{W} \in \mathbb{R}_+^{N\times K}$ is the abundance matrix that has the N-dimensional abundance vectors of the K pixels as its columns, and $\mathbf{E} \in \mathbb{R}^{L\times K}$ is an i.i.d., zero-mean Gaussian noise matrix. To comply with the physical restrictions of our mixing model, we explicitly impose a non-negativity constraint on the abundance coefficients, $\mathbf{W} \geq 0$, where the inequality $\geq$ is to be understood element-wise. Assuming that $\boldsymbol{\Phi}$ is known, we wish to estimate the abundance matrix $\mathbf{W}$, subject to the non-negativity constraint. This inverse problem is well known in the spectral unmixing literature and has been previously addressed by many methods, [1]. In this paper, we impose two prominent requirements on the abundance matrix $\mathbf{W}$, namely a) *low-rankness* and b) *sparsity*.

The low-rankness of $\mathbf{W}$ is linked with the wish to exploit the spatial information of the hyperspectral image. Indeed, a usual assumption in image processing is that there is a high degree of correlation between adjacent image pixels. Taking into account the mixing model in (1), correlation is reflected in the abundance matrix $\mathbf{W}$, in

This research has been co-financed by the European Union (European Social Fund - ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research Funding Program: ARISTEIA- HSI-MARS-1413.

the sense that adjacent columns of $\mathbf{W}$ are expected to share similar abundance values. Hence, $\mathbf{W}$ is expected to be of low rank.

The sparsity requirement, that has first been introduced in [6], relies on the assumption that only a few of the available endmembers will contribute to the spectrum of a single pixel. This assumption has permitted sparsity-driven unmixing techniques to proliferate, e.g. [2]. Sparsity is actually attributed to the abundance vectors, that may have many zero or close to zero entries. Thus, it is natural to expect that $\mathbf{W}$ will also be a sparse matrix.

At this point, we should notice that, to our knowledge, the *concurrent* imposition of sparsity and low-rankness has not been previously proposed for spectral unmixing. In [3], the low-rankness property has been enforced by penalizing the *nuclear* norm $||\mathbf{W}||_*$ of $\mathbf{W}$, which is defined as the sum of the singular values $\sigma_i(\mathbf{W})$ of $\mathbf{W}$. Moreover, a widely used approach for imposing sparsity on the abundance coefficients is to penalize the ℓ_1 norm of $\mathbf{W}$, defined as the sum of the absolute values of its elements. That said, and according to [4], we define the following optimization problem

$$\underset{\mathbf{W}\in\mathbb{R}_+^{N\times K}}{\operatorname{argmin}} \left\{ \frac{1}{2}\|\mathbf{Y}-\mathbf{\Phi W}\|_F^2 + \tau\|\mathbf{W}\|_1 + \gamma\|\mathbf{W}\|_* \right\} \quad (2)$$

where $\gamma, \tau \geq 0$ are parameters that control the trade-off among the rank and sparsity regularization terms and the data fidelity term. The optimization function in (2) is flexible enough to impose either one of the two properties of $\mathbf{W}$, depending on the values of the trade-off parameters. We further propose a generalization of (2) that includes the *weighted* nuclear and ℓ_1 norms, expressed as

$$\|\mathbf{A}\odot\mathbf{W}\|_1 = \sum_{i=1}^{N}\sum_{j=1}^{K} a_{ij}|w_{ij}|, \quad (3)$$

$$\|\mathbf{W}\|_{\mathbf{b},*} = \sum_{i=1}^{\operatorname{rank}(\mathbf{W})} b_i\sigma_i(\mathbf{W}), \quad (4)$$

where $a_{ij}, b_i \geq 0$ are weighting coefficients. Utilizing (3) and (4), the optimization problem becomes

$$\underset{\mathbf{W}\in\mathbb{R}_+^{N\times K}}{\operatorname{argmin}} \left\{ \frac{1}{2}\|\mathbf{Y}-\mathbf{\Phi W}\|_F^2 + \tau\|\mathbf{A}\odot\mathbf{W}\|_1 + \gamma\|\mathbf{W}\|_{\mathbf{b},*} \right\}. \quad (5)$$

The proposed optimization problem in (5) is convex but non-differentiable, due to the additive norm terms. To minimize this function we employ an alternating direction method of multipliers, [5], which is described in the next section.

3. THE PROPOSED ADMM ALGORITHM

In this section we develop an instance of the alternating direction method of multipliers that solves the abundance matrix estimation problem in (5). First, we consider the augmented Lagrangian

$$\begin{aligned}\mathcal{L}_2(\mathbf{W},\mathbf{\Omega}_1,\mathbf{\Omega}_2,\mathbf{\Omega}_3,\mathbf{\Omega}_4) &= \|\mathbf{\Omega}_1-\mathbf{Y}\|_F^2 + \tau\|\mathbf{A}\odot\mathbf{\Omega}_2\|_1 \\ &+ \gamma\|\mathbf{\Omega}_3\|_{\mathbf{b},*} + 1_{\mathbb{R}_+}(\mathbf{\Omega}_4) + \operatorname{tr}\left[\mathbf{D}_1^\top(\mathbf{\Phi W}-\mathbf{\Omega}_1)\right] \\ &+ \operatorname{tr}\left[\mathbf{D}_2^\top(\mathbf{W}-\mathbf{\Omega}_2)\right] + \operatorname{tr}\left[\mathbf{D}_3^\top(\mathbf{W}-\mathbf{\Omega}_3)\right] \\ &+ \operatorname{tr}\left[\mathbf{D}_4^\top(\mathbf{W}-\mathbf{\Omega}_4)\right] + \frac{\mu}{2}\left(\|\mathbf{\Phi W}-\mathbf{\Omega}_1\|_F^2 + \|\mathbf{W}-\mathbf{\Omega}_2\|_F^2\right. \\ &\left.+\|\mathbf{W}-\mathbf{\Omega}_3\|_F^2 + \|\mathbf{W}-\mathbf{\Omega}_4\|_F^2\right) \quad (6)\end{aligned}$$

where $\mathbf{D}_1,\mathbf{D}_2,\mathbf{D}_3,\mathbf{D}_4$ are matrices of Lagrange multipliers, $1_{\mathbb{R}_+}(x) = 0(+\infty)$ if $x \geq 0 (x < 0)$, and $\mu > 0$ is a penalty parameter. Let $\mathbf{\Omega} = (\mathbf{\Omega}_1,\ \mathbf{\Omega}_2,\ \mathbf{\Omega}_3,\ \mathbf{\Omega}_4)$,

$$\mathbf{G} = \begin{bmatrix} \mathbf{\Phi} \\ \mathbf{I}_N \\ \mathbf{I}_N \\ \mathbf{I}_N \end{bmatrix}, \text{ and } \mathbf{B} = \begin{bmatrix} -\mathbf{I}_L & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{I}_L & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & -\mathbf{I}_L & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & -\mathbf{I}_L \end{bmatrix}. \quad (7)$$

Then (6) can be written in a more compact form as

$$\begin{aligned}\mathcal{L}_3(\mathbf{W},\mathbf{\Omega},\mathbf{\Lambda}) &= \|\mathbf{\Omega}_1-\mathbf{Y}\|_F^2 + \tau\|\mathbf{A}\odot\mathbf{\Omega}_2\|_1 + \gamma\|\mathbf{\Omega}_3\|_{\mathbf{b},*} \\ &+ 1_{\mathbb{R}_+}(\mathbf{\Omega}_4) + \frac{\mu}{2}\|\mathbf{GW}+\mathbf{B\Omega}-\mathbf{\Lambda}\|_F^2, \quad (8)\end{aligned}$$

where $\mathbf{\Lambda} = [\mathbf{\Lambda}_1\ \mathbf{\Lambda}_2\ \mathbf{\Lambda}_3\ \mathbf{\Lambda}_4], \mathbf{\Lambda}_i = (1/\mu)\mathbf{D}_i, i = 1,\ldots,4$ are the scaled Lagrange multipliers. The optimization of $\mathcal{L}_3(\mathbf{W},\mathbf{\Omega},\mathbf{\Lambda})$ is now performed in alternating steps. In each step, $\mathcal{L}_3(\mathbf{W},\mathbf{\Omega},\mathbf{\Lambda})$ is minimized with respect to a single variable, while keeping the remaining variables fixed to their latest values. The Lagrange multiplies are also updated at the end of each updating cycle.

The subproblems that arise in the iterative procedure are easier to handle. Specifically, the optimization with respect to $\mathbf{W}$ gives

$$\begin{aligned}\mathbf{W}^{t+1} &= \underset{\mathbf{W}}{\operatorname{argmin}}\,\mathcal{L}_3(\mathbf{W},\mathbf{\Omega}^t,\mathbf{\Lambda}^t) = \left(\mathbf{\Phi}^\top\mathbf{\Phi}+3\mathbf{I}\right)^{-1} \\ &[\mathbf{\Phi}^\top(\mathbf{\Omega}_1^t+\mathbf{\Lambda}_1^t)+\mathbf{\Omega}_2^t+\mathbf{\Lambda}_2^t+\mathbf{\Omega}_3^t+\mathbf{\Lambda}_3^t+\mathbf{\Omega}_4^t+\mathbf{\Lambda}_4^t]. \quad (9)\end{aligned}$$

Next, minimization with respect to $\mathbf{\Omega}_1$ is performed as

$$\begin{aligned}\mathbf{\Omega}_1^{t+1} &= \underset{\mathbf{\Omega}_1}{\operatorname{argmin}}\,\mathcal{L}_3(\mathbf{W}^{t+1},\mathbf{\Omega},\mathbf{\Lambda}^t) \\ &= \frac{1}{1+\mu}\left(\mathbf{Y}+\mu\left(\mathbf{\Phi W}^{t+1}-\mathbf{\Lambda}_1^t\right)\right), \quad (10)\end{aligned}$$

while minimizing (8) with respect to $\mathbf{\Omega}_2$ we get

$$\begin{aligned}\mathbf{\Omega}_2^{t+1} &= \underset{\mathbf{\Omega}_2}{\operatorname{argmin}}\,\mathcal{L}_3(\mathbf{W}^{t+1},\mathbf{\Omega},\mathbf{\Lambda}^t) \\ &= \operatorname{SHR}_{\tau\mathbf{A}^t}(\mathbf{W}^{t+1}-\mathbf{\Lambda}_2^t), \quad (11)\end{aligned}$$

where the soft-thresholding operator is defined as $\operatorname{SHR}_\delta(x) = \operatorname{sign}(x)\max(0,|x|-\delta)$, and δ is a threshold parameter. $\mathbf{\Omega}_3$ is also computed by a shrinkage operation on the singular values of $\mathbf{W}^{t+1}-\mathbf{\Lambda}_3^t$, i.e.,

$$\mathbf{\Omega}_3^{t+1} = \underset{\mathbf{\Omega}_3}{\operatorname{argmin}}\,\mathcal{L}_3(\mathbf{W}^{t+1},\mathbf{\Omega},\mathbf{\Lambda}^t) = \operatorname{SVT}_{\gamma\mathbf{b}^t}(\mathbf{W}^{t+1}-\mathbf{\Lambda}_3^t), \quad (12)$$

where $\operatorname{SVT}_{\boldsymbol{\delta}}(\mathbf{X}) = \mathbf{U}\operatorname{diag}\left(\{\max\left(0,\sigma_i(\mathbf{X})-\delta_i\right)\}\right)\mathbf{V}^\top$, and $\mathbf{X} = \mathbf{U}\operatorname{diag}(\{\sigma_i(\mathbf{X})\})\mathbf{V}^\top$ is the singular value decomposition (SVD) of $\mathbf{X}$. Next, for the auxiliary variable $\mathbf{\Omega}_4$, a projection onto the nonnegative orthant is required,

$$\mathbf{\Omega}_4^{t+1} = \underset{\mathbf{\Omega}_4}{\operatorname{argmin}}\,\mathcal{L}_3(\mathbf{W}^{t+1},\mathbf{\Omega},\mathbf{\Lambda}^t) = \Pi_{\mathbb{R}_+}(\mathbf{W}^{t+1}-\mathbf{\Lambda}_4^t), \quad (13)$$

which is performed element-wise as $\Pi_{\mathbb{R}_+}(x_{ij}) = \max(0,x_{ij})$. Finally, the scaled Lagrange multipliers in $\mathbf{\Lambda}$ are sequentially updated as follows

$$\begin{aligned}\mathbf{\Lambda}_1^{t+1} &= \mathbf{\Lambda}_1^t - \mathbf{\Phi W}^{\mathbf{t+1}} + \mathbf{\Omega}_1^{t+1} \\ \mathbf{\Lambda}_i^{t+1} &= \mathbf{\Lambda}_i^t - \mathbf{W}^{t+1} + \mathbf{\Omega}_i^{t+1},\ i = 2,3,4 \quad (14)\end{aligned}$$

Algorithm 1 The alternating direction sparse and low-rank unmixing algorithm

Inputs : $\mathbf{Y}$, $\mathbf{\Phi}$
Set parameters : μ, τ, γ
Initialize : $\mathbf{W}^0, \mathbf{\Omega}^0, \mathbf{\Lambda}^0$
repeat

$$\mathbf{W}^{t+1} = \left(\mathbf{\Phi}^\top\mathbf{\Phi} + 3\mathbf{I}\right)^{-1}\left[\mathbf{\Phi}^\top(\mathbf{\Omega}_1^t + \mathbf{\Lambda}_1^t) + \mathbf{\Omega}_2^t + \mathbf{\Lambda}_2^t + \mathbf{\Omega}_3^t + \mathbf{\Lambda}_3^t + \mathbf{\Omega}_4^t + \mathbf{\Lambda}_4^t\right]$$
$$\mathbf{\Omega}_1^{t+1} = 1/(1+\mu)\left(\mathbf{Y} + \mu\left(\mathbf{\Phi}\mathbf{W}^{t+1} - \mathbf{\Lambda}_1^t\right)\right)$$
$$\mathbf{\Omega}_2^{t+1} = \mathrm{SHR}_{\tau\mathbf{A}^t}(\mathbf{W}^{t+1} - \mathbf{\Lambda}_2^t)$$
$$\mathbf{\Omega}_3^{t+1} = \mathrm{SVT}_{\gamma\mathbf{b}^t}(\mathbf{W}^{t+1} - \mathbf{\Lambda}_3^t)$$
$$\mathbf{\Omega}_4^{t+1} = \Pi_{\mathbb{R}_+}(\mathbf{W}^{t+1} - \mathbf{\Lambda}_4^t)$$
$$\mathbf{\Lambda}_1^{t+1} = \mathbf{\Lambda}_1^t - \mathbf{\Phi}\mathbf{W}^{\mathbf{t+1}} + \mathbf{\Omega}_1^{t+1}$$
$$\mathbf{\Lambda}_i^{t+1} = \mathbf{\Lambda}_i^t - \mathbf{W}^{t+1} + \mathbf{\Omega}_i^{t+1},\ i = 2, 3, 4$$

until convergence
Output : Abundance matrix $\hat{\mathbf{W}} = \mathbf{W}^{t+1}$

The proposed alternating direction sparse and low-rank unmixing algorithm (ADSpLRU) is detailed in Algorithm 1. An iteration of ADSpLRU consists of the update steps given in (9), (10), (11), (12), (13), and (14). The weights used in the norms of (11) and (12) are based on the estimates of the previous iteration, i.e.,

$$a_{ij}^t = \left(\frac{1}{w_{ij}^t - \lambda_{2,ij}^t + \varepsilon}\right) \text{ and, } b_i^t = \left(\frac{1}{\sigma_i(\mathbf{W}^t - \mathbf{\Lambda}_3^t) + \varepsilon}\right), \tag{15}$$

where $\varepsilon = 10^{-16}$. The ADSpLRU algorithm meets the general convergence prerequisites discussed in [7]. In practice, it converges within a few iterations. The stopping criterion is met when the norms of the primal and dual residuals, $\mathbf{R}^t = \mathbf{G}\mathbf{W}^t + \mathbf{B}\mathbf{\Omega}^t$, and, $\mathbf{D}^t = \mu\mathbf{G}^\top\mathbf{B}\left(\mathbf{\Omega}^t - \mathbf{\Omega}^{t-1}\right)$, respectively, become sufficiently small, i.e.,

$$\|\mathbf{R}^t\|_F \leq \epsilon \text{ and } \|\mathbf{D}^t\|_F \leq \epsilon, \tag{16}$$

where $\epsilon = 0.5 \times 10^{-4}\sqrt{(3N + L)K}$, c.f. [5]. The penalty parameter $\mu > 0$ is also adjusted in each iteration, so as to keep the primal and dual residual norms within a factor of ten of one another. The scheme used is

$$\mu^{t+1} = \begin{cases} 2\mu^t, & \text{if } \|\mathbf{R}^t\|_F > 10\|\mathbf{D}^t\|_F \\ \mu^t/2, & \text{if } \|\mathbf{R}^t\|_F < 10\|\mathbf{D}^t\|_F \\ \mu^t, & \text{otherwise.} \end{cases} \tag{17}$$

The computation complexity of the ADSpLRU is $\mathcal{O}(N^2 + K^3)$ per iteration. The most demanding part of the computation involves the singular value thresholding operation in (12).

4. EXPERIMENTAL RESULTS

In this section, the estimation performance of the proposed ADSpLRU algorithm is explored in a series of experiments conducted on both simulated and real data. In all experiments, we compare the proposed ADSpLRU algorithm with two state-of-the-art, ADMM based unmixing schemes, namely, the non-negative constrained sparse unmixing by variable splitting and augmented Lagrangian algorithm (CSUnSAl+), [8], and the recently introduced non-negative constrained joint-sparse MMV-ADMM, [3].

To quantify the estimation performance of the considered algorithms, we employ two metrics. First, we define the mean square error (MSE) as

$$\mathrm{MSE} = \frac{1}{NT}\sum_{i=1}^{T}\|\hat{\mathbf{w}}_i - \mathbf{w}_i\|_2^2, \tag{18}$$

where $\hat{\mathbf{w}}_i$ and $\mathbf{w}_i$ represent the estimated and the actual abundance vectors of the i-th pixel, and T is the number of the pixels in the image. The second metric is the signal-to-reconstruction error (SRE), which reflects the ratio between the power of signal and the error, and is given by the formula

$$\mathrm{SRE} = 10\log_{10}\left(\frac{\frac{1}{T}\sum_{i=1}^{T}\|\hat{\mathbf{w}}_i\|_2^2}{\frac{1}{T}\sum_{i=1}^{T}\|\hat{\mathbf{w}}_i - \mathbf{w}_i\|_2^2}\right). \tag{19}$$

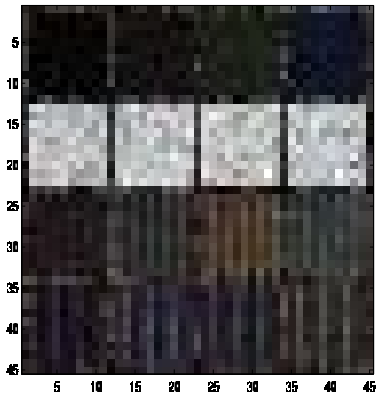
4.1. Experiment on Synthetic Image

This experiment highlights the efficacy of the proposed method in estimating abundance matrices that are in accordance with our assumptions, i.e., they either have a sparse or a low-rank structure or both. To this end, we generate a hyperspectral image using the linear mixing model in (1). The endmember matrix $\mathbf{\Phi}$ comprises $N = 25$ endmembers, that are randomly selected from the USGS library, [9]. The reflectance values are measured in 224 spectral bands, uniformly distributed in the interval $0.4 - 2.5\mu m$. As shown in Fig. 1a, the simulated hyperspectral image consists of 16 blocks of size 10×10 each. Each one of the 4 'block' rows of the image is generated using an abundance matrix of specific structure. Specifically, the first such row is generated by joint-sparse and low-rank $\mathbf{W}$'s, the second row is built using solely low-rank $\mathbf{W}$'s, and finally, rows 3 and 4 are produced by simultaneously sparse and low-rank abundance matrices. In addition, in each 'block' row, every 10×10 block is obtained using abundance matrices that have a different proportion of non-zero values and different rank. The detailed description of the parameters used to generate each abundance matrix is depicted in Table 1b. Finally, the image is also corrupted with white i.i.d Gaussian noise at a SNR level of 30dB.

Table 1c contains the MSE and SRE results obtained by the proposed ADSpLRU, as well as CSUnSAl+ and MMV-ADMM algorithms. The sparsity-imposing parameter τ of all three algorithms and the low-rankness control parameter γ of ADSpLRU were fine-tuned with respect to MSE. Moreover, the parameter μ of ADSpLRU and MMV-ADMM was set to 10^{-2}. A quick comparison of the results reveals that ADSpLRU outperforms both CSUnSAl+ and MMV-ADMM algorithms in all rows of the synthetic image, both in terms of MSE and SRE. Hence, we conclude that the proposed method achieves better performance not only when both sparsity and low-rankness are present, but also in the case where either one of them prevails.

4.2. Experiment on Real data

In this section, the considered algorithms are applied on a real hyperspectral image, namely the AVIRIS Cuprite dataset, [2]. The examined scene consists of 14 endmembers that were extracted utilizing the minimum volume simplex analysis method (MVSA), [10]. The abundance maps of two endmembers estimated by ADSpLRU, CSUnSAl+ and MMV-ADMM are shown in Fig. 2a, 2b and 2c, respectively. A quantitative examination of the abundance maps reveals that all three algorithms provide comparable results. Exploiting ground truth information on the Cuprite image, the endmembers shown can be identified as muscovite and alunite in the left and right columns in Fig. 2, respectively.



(a) Synthetic Image, 4×4 blocks of 10×10 pixels

row	column			
	1^{st}	2^{nd}	3^{rd}	4^{th}
joint sparse - 1^{st}	$(4, 1)$	$(8, 1)$	$(12, 1)$	$(16, 1)$
low-rank - 2^{nd}	$(100, 1)$	$(100, 2)$	$(100, 3)$	$(100, 4)$
sparse & low-rank - 3^{rd}	$(4, 2)$	$(8, 2)$	$(12, 2)$	$(16, 2)$
sparse & low-rank - 4^{th}	$(4, 3)$	$(8, 3)$	$(12, 3)$	$(16, 3)$

(b) Structure of $\mathbf{W}$ in each block of the synthetic image, each cell contains the pair : $\left(\frac{|\text{supp}(\mathbf{W})|}{|\mathbf{W}|}\%, \text{rank}(\mathbf{W})\right)$

Algorithm	row (1^{st})		row (2^{nd})		row (3^{rd})		row (4^{th})	
	MSE	SRE	MSE	SRE	MSE	SRE	MSE	SRE
CSunSal+	0.09	17.99	1.53	12.22	0.34	14.35	0.29	15.16
MMV-ADMM	0.07	18.70	1.78	11.69	0.41	12.80	0.29	14.39
ADSpLRU	0.01	39.98	0.75	17.12	0.11	25.40	0.08	23.85

(c) MSE(10^{-2}) and SRE(dB) results on synthetic image for each row.

Fig. 1: Structure of Synthetic Image and Results

5. REFERENCES

[1] W.-K. Ma, J.M. Bioucas-Dias, Tsung-Han C., et al., "A signal processing perspective on hyperspectral unmixing: Insights from remote sensing," *Signal Processing Magazine, IEEE*, vol. 31, no. 1, pp. 67–81, Jan 2014.

[2] K. E. Themelis, A. A. Rontogiannis, and K. D. Koutroumbas, "A novel hierarchical Bayesian approach for sparse semisupervised hyperspectral unmixing," *Signal Processing, IEEE Transactions on*, vol. 60, no. 2, pp. 585–599, Feb 2012.

[3] Q. Qu, N.M. Nasrabadi, and T.D. Tran, "Abundance estimation for bilinear mixture models via joint sparse and low-rank representation," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 52, no. 7, pp. 4404–4423, July 2014.

[4] E. Richard, P.-A. Savalle, and N. Vayatis, "Estimation of simultaneously sparse and low rank matrices," in *ICML*, 2012.

[5] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

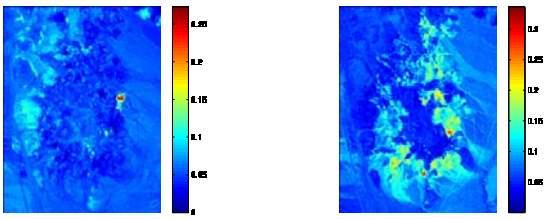
[6] K. E. Themelis, A. A. Rontogiannis, and K. D. Koutroumbas, "Semi-supervised hyperspectral unmixing via the weighted lasso," in *ICASSP*, 2010, pp. 1194–1197.

[7] J. Eckstein and D. Bertsekas, "On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators," *Mathematical Programming*, vol. 55, no. 1-3, pp. 293–318, 1992.

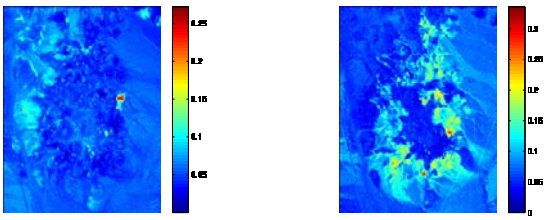
[8] J. M. Bioucas-Dias and M. A.T. Figueiredo, "Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing," in *Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, 2010, pp. 1–4.

[9] R. N. Clark, G. A. Swayze, et al., "USGS digital spectral library," 2007, `http://speclab.cr.usgs.gov/spectral.lib06/ds231/datatable.html`.

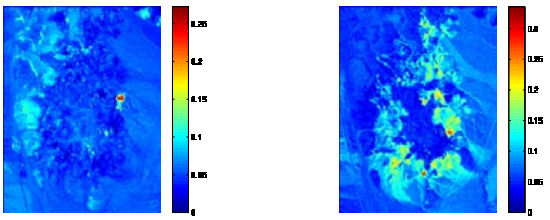
[10] J. Li and J. M. Bioucas-Dias, "Minimum volume simplex analysis: A fast algorithm to unmix hyperspectral data," in *Geoscience and Remote Sensing Symposium, IGARSS*. IEEE, 2008, vol. 3, pp. III–250.



(a) ADSpLRU



(b) CSUnSAl+



(c) MMV-ADMM

Fig. 2: Estimated abundance maps for the Aviris Cuprite data set for the materials muscovite and alunite, using the proposed ADSpLRU, CSUnSAl+ and MMV-ADMM algorithms.