Distributed Blind Hyperspectral Unmixing via Joint Sparsity and Low-Rank Constrained Non-Negative Matrix Factorization

Christos G. Tsinos, *Member, IEEE*, Athanasios A. Rontogiannis, *Member, IEEE*, and Kostas Berberidis, *Senior Member, IEEE*

***Abstract*—Hyperspectral unmixing is a crucial processing step in remote sensing image analysis. Its aim is the decomposition of each pixel in a hyperspectral image into a number of materials, the so-called endmembers, and their corresponding abundance fractions. Among the various unmixing approaches that have been suggested in the literature, we are interested here in unsupervised techniques that rely on some form of non-negative Matrix factorization (NMF). NMF-based techniques provide an easy way to simultaneously estimate the endmembers and their corresponding abundances, though they suffer from mediocre performance and high computational complexity due to the nonconvexity of the involved cost function. Improvements in performance have been recently achieved by imposing additional constraints to the NMF optimization problem related to the sparsity of the abundances. Another feature of hyperspectral images that can be exploited is their high spatial correlation, which is translated into the low rank of the involved abundance matrices. Motivated by this, in this paper we propose a novel unmixing method that is based on a simultaneously sparse and low-rank constrained NMF. In addition, prompted by the rapid evolution of multicore processors and graphics processing units, we devise a distributed unmixing scheme that processes in parallel different parts of the image. The proposed distributed unmixing algorithm achieves improved performance and faster convergence than existing state-of-the-art techniques as it is verified by extensive simulations on synthetic and real hyperspectral data.**

***Index Terms*—Hyperspectral unmixing, L_1-regularization, non-negative matrix mactorization (NMF), nuclear norm, sparse estimation.**

Manuscript received September 29, 2016; revised February 4, 2017 and April 4, 2017; accepted April 4, 2017. Date of publication April 12, 2017; date of current version May 8, 2017. This work was supported in part by the National Observatory of Athens (in the framework of the project ARISTEIA HSI-MARS) and in part by the University of Patras. The guest editor coordinating the review of this manuscript and approving it for publication was Dr. Shuchin Aeron. *(Corresponding Author: Christos G. Tsinos.)*

C. G. Tsinos is with the Interdisciplinary Centre for Security, Reliability and Trust, University of Luxembourg, Luxembourg 2721, Europe (e-mail: christos.tsinos@uni.lu).

A. A. Rontogiannis is with the Institute for Astronomy, Astrophysics, Space Applications and Remote Sensing, National Observatory of Athens, Penteli GR-15236, Greece (e-mail: tronto@noa.gr).

K. Berberidis is with the Department of Computer Engineering and Informatics and CTI/RU8, University of Patras, Patras 26500, Greece (e-mail: berberid@ceid.upatras.gr).

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

Digital Object Identifier 10.1109/TCI.2017.2693967

I. Introduction

In RECENT years, hyperspectral images have been extensively used for earth observation and other remote sensing applications (see [1]–[4] and references therein). In such applications usually the employed sensors have low spatial resolution resulting in images with highly mixed spectra, i.e., the spectrum of each pixel is composed of the spectrums of a number of different substances. Thus, the development of hyperspectral unmixing methods is important in order to decompose the mixed pixels into their endmembers and the corresponding fractional abundances. In the literature so far, most of the existing approaches treat the unmixing problem under the assumption of a Linear Mixture Model (LMM). The LMM fits well in several cases [1]–[3] and due to its simplicity, it is commonly adopted in hyperspectral unmixing. Under this model, a pixel is viewed as a linear combination of the endmembers' spectral signatures by their corresponding abundances. Moreover, the abundances usually satisfy a non-negativity and a sum-to-one constraint.

Among the different techniques suggested so far are those based on a proper exploitation of the geometric properties of the constrained LMM, e.g., the premise that the observations lie in a simplex whose vertices are the endmembers of the scene. Representative approaches include the N-FINDR algorithm [5], the Vertex Component Analysis [6], the Pixel Purity Index [7] and the Simplex Growing Algorithm (SGA) [8]. The previous schemes provide only estimates of the endmembers' spectral signatures, while the corresponding abundances can be estimated via a Fully Constrained Least Squares (FCLS) method [9]. Moreover, these approaches rely on the existence of pure pixels in the image, an assumption which is in general not true due to the low spatial resolution of the latter.

Hyperspectral unmixing algorithms that do not require the pure pixel assumption have also been developed in the literature with the view of improving the performance on highly mixed data. Such approaches that are based on the geometric properties of the constrained LMM include the Iterated Constrained Endmembers (ICE) algorithm [10], the Minimum Volume Simplex Analysis (MVSA) [11] and the Minimum-Volume Enclosing Simplex (MVES) algorithm [12].

In a statistical framework, unsupervised hyperspectral unmixing can be treated as a Blind Source Separation (BSS) problem. Among the first approaches in this framework was the

Independent Component Analysis (ICA) method of [13]. Unfortunately, ICA is based on the assumption that the mixed sources (endmembers) are mutually independent, which usually does not hold true for hyperspectral data. Departing from ICA, a large number of unmixing techniques were proposed within the context of the so-called Non-Negative Matrix Factorization (NMF) [14]. NMF-based techniques decompose the data into two non-negative matrix factors, the endmembers and the abundance matrices, by solving an optimization problem subject to a sum-to-one constraint on the pixel abundances. Therefore, they provide a compact and simple way to estimate simultaneously the latter matrices. Considering also that they have no requirement for existence of pure pixels, NMF-based techniques seem to be very attractive for unmixing hyperspectral data. Nevertheless, direct application of the sum-to-one constrained NMF to the unmixing problem usually suffers from poor estimation performance due to the non-convexity of the involved cost function. A typical solution that improves the performance is to solve the NMF optimization problem subject to a number of additional constraints related to the underlying structure of the hyperspectral data.

Such relevant approaches impose constraints related to the geometric properties of the assumed constrained LMM [15], the smoothness [16] and the sparsity [17], [18] of the abundances, endmembers' dissimilarity [19] and the manifold structure [20]. Among the aforementioned approaches, the sparsity constrained NMF gained a lot of interest lately. Since only a subset of the scene's endmembers contribute to each of the pixels, the abundance matrix exhibits a sparse structure [21]. The sparsity of the abundance matrix can be imposed by incorporating a l_0 regularizer in the cost function of the NMF optimization problem. This results in a NP-hard optimization problem and thus, alternative sparsity promoting reguralizers have been sought. To that end, the performance of the l_1 and $l_{1/2}$ reguralizers was studied for the NMF hyperspectral unmixing problem in [17], [18]. In [18], the $l_{1/2}$ regularized version of NMF, denoted as $L_{1/2}$-NMF, is shown to lead in general to sparser solutions than its l_1 counterpart [17]. Nevertheless, both these sparse approaches are in general unstable and their performance depends heavily on the values of the regularization parameters. Moreover, they rely on projective gradient steps or on the well-known multiplicative update rule [22], with both of them exhibiting very slow convergence. In a different framework, a multitask sparse NMF for hyperspectral imagery denoising was developed in [23] that exhibits improved convergence speed and improved denoising performance by linking the per band denoising across the spectral domain in order to exploit the inherent correlations.

Aside from the sparsity of the abundance matrix, a key characteristic of hyperspectral images is the high spatial correlation that is mainly observed *locally* in the image. That is, the abundance vectors of adjacent pixels are linearly dependent, which results in pixel neighborhoods that correspond to *low rank* abundance sub-matrices. Rank constrained optimization problems are NP-hard to solve, so a typical approach is to relax the problem by replacing the rank constraint with a nuclear norm (l_*) regularizer in the cost function [24]. In the last years, rank minimization techniques have been an active research field in the signal processing literature [25], [26]. These works consider the case of cost functions that involve sparse plus low rank decompositions (sum of a low rank and a sparse matrix). Such an approach was already considered in the context of NMF for classification and clustering problems [27].

In the hyperspectral image literature, rank constrained techniques that capture the local spatial correlation were developed in the context of semi-supervised unmixing [28]–[31]. In semi-supervised unmixing the hyperspectral image pixels are expressed as linear combinations of a number of pure spectral signatures that are known in advance. Therefore, the unmixing procedure boils down to determining which endmembers, from an available library, represent better the scene and what the proportion of their contribution is, that is their corresponding abundances. To cope with the NP hardness of this problem, [28], [29] employed linear sparse regression techniques as a means to obtain a solution more efficiently. According to [30], [31] the imposition of a rank regularizer in the problem can improve significantly the unmixing performance. Especially, in [31] a novel semi-supervised unmixing technique is developed that imposes simultaneously a low-rank and a sparsity constraint on the abundance matrix. It is noteworthy that there are only few approaches in existing signal processing literature that examine optimization problems that pose simultaneous constraints on the structure of the involved matrices, e.g., [32]–[34].

While the local spatial correlation of the abundance matrix has already been considered in unsupervised unmixing methods from a Bayesian perspective [35], it has yet to be addressed as a rank-constrained *unsupervised (blind)* unmixing optimization problem in the literature. To that end, the major contribution of the present work is the development of a NMF-based unmixing scheme that imposes simultaneously low-rank and sparsity constraints on the abundance matrix. The proposed solution is based on the so-called Alternating Direction Method of Multipliers (ADMM) [36] which enables the estimation of both the endmembers and abundance matrices via an iterative procedure of alternating optimization steps that converges faster and achieves better performance than the existing approaches.

Recent advances in the hyperspectral imaging field include the development of parallel unmixing techniques that exploit the available multi-cores in modern computational units (CPUs-GPUs) [37], [38] to distribute the computational overhead among the available units and thus to speed up the unmixing task. Motivated by this, the proposed algorithm is extended so that distributed processing is enabled. The resulting technique provides estimates of both the endmember and the abundance matrices in a completely parallel manner and distributes the computational overhead among the available computational units improving further the proposed method's speed of convergence.

The rest of the manuscript is organized as follows. In Section II the system model is defined. In Section III the concept of the NMF based hyperspectral unmixing problem is explained. In Section IV, the development of the proposed algorithm is described. In Section V, a discussion regarding the convergence of the proposed technique is given. In Section VI, the distributed version of the proposed technique is developed. Section VII presents a number of simulations on both synthetic and real

hyperspectral data that corroborate the effectiveness of the proposed technique. Finally, Section VIII concludes this work.

II. System Model

Let us consider a $M \times Q \times L$ hyperspectral image where L is the number of spectral bands and $N = MQ$ is the number of image pixels. Under the assumption that a LMM describes adequately the underlying mixing of the observed scene, the ith pixel of the image $\mathbf{x}_i \in \mathbb{R}_+^L$ is expressed as

$$\mathbf{x}_i = \mathbf{A}\mathbf{s}_i + \mathbf{e}_i, \tag{1}$$

where $\mathbf{s}_i \in \mathbb{R}_+^P$ is the vector of abundance fractions corresponding to the P endmembers in the scene, $\mathbf{e}_i \in \mathbb{R}^L$ is the additive noise vector and $\mathbf{A} \in \mathbb{R}_+^{L\times P}$ is the matrix whose columns are the spectral signatures of the P endmembers. It is often assumed that the entries of a pixel's abundance vector add to one (sum-to-one constraint), that is

$$\sum_{p=1}^{P} s_{pi} = 1, \tag{2}$$

where s_{pi} denotes the pth element of vector $\mathbf{s}_i$. By stacking the N pixels of the image in a $L \times N$ matrix $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_N]$ the LMM of (1) can be written in matrix form as

$$\mathbf{X} = \mathbf{A}\mathbf{S} + \mathbf{E}, \tag{3}$$

where matrices $\mathbf{S} = [\mathbf{s}_1, \ldots, \mathbf{s}_N]$ and $\mathbf{E} = [\mathbf{e}_1, \ldots, \mathbf{e}_N]$, are the horizontal concatenation of the abundances ($\mathbf{s}_i$) and noise vectors ($\mathbf{e}_i$), respectively.

III. NMF Based Hyperspectral Unmixing

In NMF based unmixing we seek two non-negative matrices $\mathbf{A} \in \mathbb{R}_+^{L\times P}$ and $\mathbf{S} \in \mathbb{R}_+^{P\times N}$ whose product approximates matrix $\mathbf{X} \in \mathbb{R}_+^{L\times N}$ in the least squares sense by solving the following optimization problem

$$(\mathcal{P}_1): \quad \min_{\mathbf{A},\mathbf{S}} \quad \frac{1}{2}\|\mathbf{X} - \mathbf{A}\mathbf{S}\|_F^2 \quad \text{s.t.} \quad \mathbf{A} \geq 0,\ \ \mathbf{S} \geq 0,\ \ \mathbf{1}_P^T\mathbf{S} = \mathbf{1}_N, \tag{4}$$

where $\mathbf{1}_P$ is the $P \times 1$ vector of ones.

The solution of $(\mathcal{P}_1)$ can be obtained by projective gradient techniques or by the well-known multiplicative update rule [22]. Unfortunately, due to the non-convex nature of the cost function, the aforementioned approaches are usually trapped in local minima resulting to poor unmixing performance.

More recently, in [17], [18] further constraints have been imposed to the NMF optimization problem $(\mathcal{P}_1)$ in order to reduce the set of feasible optimal solutions. A common idea is to exploit the sparse structure of the abundance matrix $\mathbf{S}$ by adding a sparsity regularizer to the cost function of $(\mathcal{P}_1)$. That is,

$$(\mathcal{P}_2): \quad \min_{\mathbf{A},\mathbf{S}} \quad \frac{1}{2}\|\mathbf{X} - \mathbf{A}\mathbf{S}\|_F^2 + \lambda\|\mathbf{S}\|_q, \quad \text{s.t.} \quad \mathbf{A} \geq 0,\ \ \mathbf{S} \geq 0,\ \ \mathbf{1}_P^T\mathbf{S} = \mathbf{1}_N, \tag{5}$$

where $0 < q \leq 1$, λ is a regularization parameter and

$$\|\mathbf{S}\|_q = \sum_{p=1}^{P}\sum_{i=1}^{N} s_{pi}^{1/q}. \tag{6}$$

It has been observed [18] that the selection $q = 1/2$ gives the best results. The sparse structure of the abundance matrix is due to the fact that a pixel is usually a mixture of a small subset of the endmembers existing in the scene.

Nevertheless, the solution of a sparsity constrained NMF relies heavily on the initial selection of $\mathbf{A}$ and $\mathbf{S}$, and its performance depends critically on the values of the regularization parameters, hence in several cases the unmixing results are still poor, as it was also stated in Section I. That is, the sparsity constraint cannot improve significantly the performance of simple NMF based unmixing.

IV. Sparsity and Low-Rank Constrained NMF

According to the description of the previous sections, NMF-based techniques can be used for unmixing a hyperspectral image in an unsupervised manner, since they can estimate simultaneously the abundance and endmembers matrices. While the latter is a highly desired feature (the method does not require a dictionary with the endmembers signatures, as in semi-supervised unmixing), it comes with poor estimation results. Following the discussion of Section III, it is clear that the performance can be improved significantly if additional structural constraints are imposed on the NMF optimization problem $(\mathcal{P}_1)$. As the sparsity constraint alone offers only slight improvement in performance, we propose here the incorporation of additional structural constraints in the NMF optimization problem, so as to reduce the feasible set of optimal solutions. Such a constraint is the *low rank* of the abundance matrices corresponding to adjacent pixels, which is fully justified by the high spatial correlation of pixels belonging to the same neighborhood [31]. In view of this, we present in this section a new NMF-based algorithm, which is subject to sparsity and low-rank constraints at the same time. In the present section, the centralized version of the proposed approach is presented. In the next section, the algorithm is extended in a distributed form, enabling the use of parallel processing units so as to speed up convergence.

Let us assume that the input image is divided into K *non-overlapping* sub-images of size $r \times r$. Let us further assume that the abundance matrix of the kth sub-image is given by $\mathbf{S}_k = [\mathbf{s}_{(k-1)r^2}, \ldots, \mathbf{s}_{kr^2-1}]$ with $1 \leq k \leq K$. The columns of $\mathbf{S}_k$ correspond to the sub-image's pixels stored in vector form. It is also straightforward to see that $\mathbf{S} = [\mathbf{S}_1, \ldots, \mathbf{S}_K]$. The low rank structure is usually enforced by adding a nuclear norm (l_*) related reguralizer [24] to avoid the NP-hardness of the original rank-constrained problem. Moreover, the sparse structure of the abundance matrices is also exploited by adding a sparsity enforcing l_1 regularizer [39] that relaxes in its turn the original l_0 constrained problem (also NP-hard).

Thus, in order to enforce simultaneously the *local* low-rank and the sparse structure to the abundance matrix $\mathbf{S}_k$, we may

minimize the following cost function

$$f(\mathbf{A},\mathbf{S}_1,\ldots,\mathbf{S}_K)=\frac{1}{2}\sum_{k=1}^{K}\|\mathbf{X}_k-\mathbf{A}\mathbf{S}_k\|_F^2+\lambda\sum_{k=1}^{K}\|\mathbf{S}_k\|_1+\gamma\sum_{k=1}^{K}\|\mathbf{S}_k\|_*, \tag{7}$$

where $\mathbf{X}_k=[\mathbf{x}_{(k-1)r^2},\ldots,\mathbf{x}_{kr^2-1}]$, λ and γ are regularization parameters,

$$\|\mathbf{S}_k\|_*=\sum_{j=1}^{\min\{P,r^2\}}\sigma_j\left(\mathbf{S}_k\right), \tag{8}$$

is the nuclear norm of matrix $\mathbf{S}_k$, σ_j, $1\leq j\leq\min\{P,r^2\}$ are its corresponding (ordered) singular values and we used the equation $\|\mathbf{X}-\mathbf{A}\mathbf{S}\|_F^2=\sum_{k=1}^{K}\|\mathbf{X}_k-\mathbf{A}\mathbf{S}_k\|_F^2$. From (7) and the initial optimization problem $(\mathcal{P}_1)$, the unmixing problem is re-defined as

$$(\mathcal{P}_3):\quad\min_{\mathbf{A},\mathbf{S}_1,\ldots,\mathbf{S}_K}\quad f(\mathbf{A},\mathbf{S}_1,\ldots,\mathbf{S}_K)$$
$$\text{s.t.}\qquad\mathbf{A}\geq 0,\ \ \mathbf{S}_k\geq 0,\ \ \mathbf{1}_P^T\mathbf{S}_k=\mathbf{1}_{r^2}. \tag{9}$$

Let us now describe how to solve $(\mathcal{P}_3)$ in (9). By introducing the auxiliary variables $\mathbf{C}$ and $\mathbf{D}_k$, $1\leq k\leq K$, the optimization problem $(\mathcal{P}_3)$ can be written as

$$(\mathcal{P}_4):\min_{\mathbf{A},\mathbf{C},\mathbf{S}_k,\mathbf{D}_k}\ \frac{1}{2}\sum_{k=1}^{K}\|\mathbf{X}_k-\mathbf{A}\mathbf{S}_k\|_F^2+\lambda\sum_{k=1}^{K}\|\mathbf{S}_k\|_1+\gamma\sum_{k=1}^{K}\|\mathbf{D}_k\|_*+\sum_{k=1}^{K}\mathbb{1}_{\mathbb{R}^+}\{\mathbf{D}_k\}+\mathbb{1}_{\mathbb{R}^+}\{\mathbf{C}\}$$
$$\text{s.t.}\qquad\mathbf{A}-\mathbf{C}=\mathbf{0},\ \ \mathbf{S}_k-\mathbf{D}_k=\mathbf{0},\forall k\in[1,K] \tag{10}$$

where the indicator functional $\mathbb{1}_{\mathbb{R}^+}\{\cdot\}$ is 0 when the operand matrix is positive definite and $+\infty$, otherwise. Note that in $(\mathcal{P}_4)$, we have omitted the sum-to-one constraint since it contradicts the l_1 regularizer that appears in the cost function of the optimization problem. Note that, by imposing the sum-to-one constraint results in an l_1 norm that is always equal to one and thus, the impact of the corresponding regularizer on the $(\mathcal{P}_4)$ is canceled. Therefore, one out of the two constraints has to be taken away from the problem. Moreover, note that the sum-to-one constraint has been prone to strong criticisms in the relevant literature [28], [40], [41] and thus, several sparse unmixing approaches [28], [31], [41]–[43] choose to drop this constraint, as it is also done here.

To solve $(\mathcal{P}_4)$ we employ the so-called Alternating Direction Method of Multipliers (ADMM) [36]. This method is an extension of the augmented Lagrangian multiplier method [44]. ADMM first performs an alternating optimization with respect to each one of the involved variables individually and then, it updates the corresponding Lagrange multipliers via a gradient based rule. For $(\mathcal{P}_4)$, the augmented Lagrangian function is given by

$$\begin{aligned}\mathcal{L}(\mathbf{A},\mathbf{S}_1,\ldots,\mathbf{S}_K,\mathbf{C},\mathbf{D}_1,\ldots,\mathbf{D}_K,\mathbf{\Lambda},\mathbf{\Pi}_1,\ldots,\mathbf{\Pi}_K)&\\=\frac{1}{2}\sum_{k=1}^{K}\|\mathbf{X}_k-\mathbf{A}\mathbf{S}_k\|_F^2+\lambda\sum_{k=1}^{K}\|\mathbf{S}_k\|_1+\gamma\sum_{k=1}^{K}\|\mathbf{D}_k\|_*&\\+\mathbb{1}_{\mathbb{R}^+}\{\mathbf{C}\}+\sum_{k=1}^{K}\mathbb{1}_{\mathbb{R}^+}\{\mathbf{D}_k\}+\langle\mathbf{\Lambda},\mathbf{A}-\mathbf{C}\rangle&\\+\sum_{k=1}^{K}\langle\mathbf{\Pi}_k,\mathbf{S}_k-\mathbf{D}_k\rangle+\frac{\alpha}{2}\|\mathbf{A}-\mathbf{C}\|_F^2+\sum_{k=1}^{K}\frac{\beta}{2}\|\mathbf{S}_k-\mathbf{D}_k\|_F^2,&\end{aligned} \tag{11}$$

where for two matrices $\mathbf{W}$ and $\mathbf{G}$, $\langle\mathbf{W},\mathbf{G}\rangle=\sum_{i,j}w_{ij}g_{ij}$, $\mathbf{\Lambda}$ and $\mathbf{\Pi}_k$ are the $L\times P$ and $P\times r^2$ Lagrange multiplier matrices respectively and α,β are scalar parameters. From now on, we will consider $\beta=\alpha$ without loss of generality. Let us now define $\bar{\mathbf{S}}=\{\mathbf{S}_1,\ldots,\mathbf{S}_K\}$, $\bar{\mathbf{D}}=\{\mathbf{D}_1,\ldots,\mathbf{D}_K\}$ and $\bar{\mathbf{\Pi}}=\{\mathbf{\Pi}_1,\ldots,\mathbf{\Pi}_K\}$ for presentation purposes. According to the ADMM approach, the solution to the optimization problems $(\mathcal{P}_4)$ is given by the following alternating minimization steps

$$\mathbf{A}_n=\arg\min_{\mathbf{A}}\mathcal{L}(\mathbf{A},\bar{\mathbf{S}}_{n-1},\mathbf{C}_{n-1},\bar{\mathbf{D}}_{n-1},\mathbf{\Lambda}_{n-1},\bar{\mathbf{\Pi}}_{n-1}) \tag{12}$$

$$\mathbf{S}_{k,n}=\arg\min_{\mathbf{S}_k}\mathcal{L}(\mathbf{A}_n,\bar{\mathbf{S}},\mathbf{C}_{n-1},\bar{\mathbf{D}}_{n-1},\mathbf{\Lambda}_{n-1},\bar{\mathbf{\Pi}}_{n-1}) \tag{13}$$

$$\mathbf{C}_n=\arg\min_{\mathbf{C}}\mathcal{L}(\mathbf{A}_n,\bar{\mathbf{S}}_n,\mathbf{C},\bar{\mathbf{D}}_{n-1},\mathbf{\Lambda}_{n-1},\bar{\mathbf{\Pi}}_{n-1}) \tag{14}$$

$$\mathbf{D}_{k,n}=\arg\min_{\mathbf{D}_k}\mathcal{L}(\mathbf{A}_n,\bar{\mathbf{S}}_n,\mathbf{C}_n,\bar{\mathbf{D}},\mathbf{\Lambda}_{n-1},\bar{\mathbf{\Pi}}_{n-1}) \tag{15}$$

$$\mathbf{\Lambda}_n=\mathbf{\Lambda}_{n-1}+\alpha(\mathbf{A}_n-\mathbf{C}_n) \tag{16}$$

$$\mathbf{\Pi}_{k,n}=\mathbf{\Pi}_{k,n-1}+\alpha\left(\mathbf{S}_{k,n}-\mathbf{D}_{k,n}\right). \tag{17}$$

where n is the iteration index. Let us first consider the minimization problem given in (12) which can be shown to be equivalent to

$$\mathbf{A}_n=\arg\min_{\mathbf{A}}\left\{\frac{1}{2}\|\mathbf{X}-\mathbf{A}\mathbf{S}_{n-1}\|_F^2+\langle\mathbf{\Lambda}_{n-1},\mathbf{A}-\mathbf{C}_{n-1}\rangle+\frac{\alpha}{2}\|\mathbf{A}-\mathbf{C}_{n-1}\|_F^2\right\}, \tag{18}$$

where $\mathbf{S}_n=[\mathbf{S}_{1,n},\ldots,\mathbf{S}_{K,n}]$. The solution to eq. (18) is obtained by taking the gradient of the cost function with respect to $\mathbf{A}$ and then equating to zero, yielding

$$\mathbf{A}_n=\left(\mathbf{X}\mathbf{S}_{n-1}^T-\mathbf{\Lambda}_{n-1}+\alpha\mathbf{C}_{n-1}\right)\left(\mathbf{S}_{n-1}\mathbf{S}_{n-1}^T+\alpha\mathbf{I}_L\right)^{-1}. \tag{19}$$

We move now to the derivation of the solution to (14). It can be shown, that the optimization problem under consideration is equivalent to

$$\mathbf{C}_n=\arg\min_{\mathbf{C}}\left\{\mathbb{1}_{\mathbb{R}^+}\{\mathbf{C}\}+\langle\mathbf{\Lambda}_{n-1},\mathbf{A}_n-\mathbf{C}\rangle+\frac{\alpha}{2}\|\mathbf{A}_n-\mathbf{C}\|_F^2\right\}, \tag{20}$$

the solution of which corresponds to the projection of the matrix $\mathbf{A}_n + \alpha^{-1}\mathbf{\Lambda}_{n-1}$ onto the non-negative orthant. It can be shown that the aforementioned solution of (20) admits the following closed form [39]

$$\mathbf{C}_n = \Pi_{\mathbb{R}_+}\left\{\mathbf{A}_n + \frac{1}{\alpha}\mathbf{\Lambda}_{n-1}\right\}, \tag{21}$$

where $\Pi_{\mathbb{R}_+}\{\cdot\}$ is the operator that projects the operand matrix to the corresponding non-negative orthant.

Before proceeding further, let us first define some operators. For a matrix $\mathbf{B}$ with Singular Value Decomposition (SVD) $\mathbf{B} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H \in \mathbb{R}^{D1\times D2}$, we define

$$SHR_\theta\{\mathbf{B}\} = sgn(\mathbf{B}) \odot \max\{|\mathbf{B}| - \theta, \mathbf{0}\} \tag{22}$$

and

$$SVT_\theta\{\mathbf{B}\} = \mathbf{U}\max\{\mathbf{\Sigma} - \theta, \mathbf{0}\}\mathbf{V}^H, \tag{23}$$

where $sgn(\cdot)$ is the element-wise sign operator, $\odot$ is the Hadamard product of two matrices and $\max\{\cdot,\cdot\}$ is the element-wise maximum of two matrices. We may now proceed to the derivation of the solution of (13) and (15). The variables that are updated in these equations are related to the *local* sub-image abundance matrices $\mathbf{S}_k$, $\forall k \in [1, K]$. Let us start with the optimization problem in (13) which can be written as follows,

$$\mathbf{S}_{k,n} = \arg\min_{\mathbf{S}_k}\left\{\frac{1}{2}\|\mathbf{X}_k - \mathbf{A}_n\mathbf{S}_k\|_F^2 + \lambda\|\mathbf{S}_k\|_1 + \langle\mathbf{\Pi}_{k,n-1}, \mathbf{S}_k - \mathbf{D}_{k,n-1}\rangle + \frac{\alpha}{2}\|\mathbf{S}_k - \mathbf{D}_{k,n-1}\|_F^2\right\}, \tag{24}$$

The optimization problem in (24) can be solved using the results in [39] leading to the following expression,

$$\mathbf{S}_{k,n} = SHR_{\frac{\lambda}{\alpha}}\left\{\left(\mathbf{A}_n^T\mathbf{A}_n + \alpha\mathbf{I}_P\right)^{-1} \times \left(\mathbf{A}_n^T\mathbf{X}_k - \mathbf{\Pi}_{k,n-1} + \alpha\mathbf{D}_{k,n-1}\right)\right\}. \tag{25}$$

We move now to problem (15) which is equivalent to the one given by,

$$\mathbf{D}_{k,n} = \arg\min_{\mathbf{D}_k}\left\{\gamma\|\mathbf{D}_k\|_* + \mathbb{1}_{\mathbb{R}^+}\{\mathbf{D}_k\} \langle\mathbf{\Pi}_{k,n-1}, \mathbf{S}_{k,n} - \mathbf{D}_k\rangle + \frac{\alpha}{2}\|\mathbf{S}_{k,n} - \mathbf{D}_k\|_F^2\right\}. \tag{26}$$

From [45], it can be shown that (26) has the following closed form solution

$$\mathbf{D}_{k,n} = \Pi_{\mathbb{R}_+}\left\{SVT_{\frac{\gamma}{\alpha}}\left\{\mathbf{S}_{k,n} + \frac{1}{\alpha}\mathbf{\Pi}_{k,n-1}\right\}\right\}. \tag{27}$$

From the above description, ADMM requires the application of the alternating minimization steps of (12)-(17) in a cyclic iterative manner until a convergence criterion is satisfied. Here, the following composite criterion is used,

Algorithm 1: Sparsity and Low-Rank Constrained NMF.

Input: $\mathbf{X}$, p, K, α, γ, λ, tol, N_{max}

1: Initialize $\mathbf{A}$ and $\mathbf{S}$ by the VCA-FCLS techique to $\mathbf{A}_0$, $\mathbf{S}_0$
2: Set $\mathbf{\Lambda}_0$ and $\mathbf{D}_{k,0}$ as zero entries matrices
3: **while** The criteria in (28) are not met & $n \leq N_{max}$ **do**
4: $\quad\mathbf{A}_n = \left(\mathbf{X}\mathbf{S}_{n-1}^T - \mathbf{\Lambda}_{n-1} + \alpha\mathbf{C}_{n-1}\right)\times$
5: $\qquad\left(\mathbf{S}_{n-1}\mathbf{S}_{n-1}^T + \alpha\mathbf{I}_L\right)^{-1}$
6: $\quad\mathbf{C}_n = \Pi_{\mathbb{R}_+}\left\{\mathbf{A}_n + \frac{1}{\alpha}\mathbf{\Lambda}_{n-1}\right\}$
7: $\quad\mathbf{\Lambda}_n = \mathbf{\Lambda}_{n-1} + \alpha(\mathbf{A}_n - \mathbf{C}_n)$
8: $\quad$**for** $1 \leq k \leq K$ **do**
9: $\qquad\mathbf{S}_{k,n} = SHR_{\frac{\lambda}{\alpha}}\left\{\left(\mathbf{A}_n^T\mathbf{A}_n + \alpha\mathbf{I}_P\right)^{-1}\times\right.$
10: $\qquad\qquad\left.\left(\mathbf{A}_n^T\mathbf{X}_k - \mathbf{\Pi}_{k,n-1} + \alpha\mathbf{D}_{k,n-1}\right)\right\}$
11: $\qquad\mathbf{D}_{k,n} = \Pi_{\mathbb{R}_+}\left\{SVT_{\frac{\gamma}{\alpha}}\left\{\mathbf{S}_{k,n} + \frac{1}{\alpha}\mathbf{\Pi}_{k,n-1}\right\}\right\}$
12: $\qquad\mathbf{\Pi}_{k,n} = \mathbf{\Pi}_{k,n-1} + \alpha\left(\mathbf{S}_{k,n} - \mathbf{D}_{k,n}\right)$
13: $\quad$**end for**
14: **end while**

Return: $\mathbf{A}$, $\mathbf{S}$

$$\frac{|f_n - f_{n-1}|}{|f_{n-1}|} \leq \epsilon_1 \;\&\; \|\mathbf{A}_n - \mathbf{C}_n\|_F^2 \leq \epsilon_2 \;\&\; \sum_{k=1}^{K}\|\mathbf{S}_{k,n} - \mathbf{D}_{k,n}\|_F^2 \leq \epsilon_3 \tag{28}$$

where $f_n = \|\mathbf{X} - \mathbf{A}_n\mathbf{S}_n\|_F^2$ and ϵ_1, ϵ_2 and ϵ_3 are pre-defined tolerances. The first condition checks if the algorithm has converged to a point. The second and the third ones check if the KKT conditions are satisfied with respect to the corresponding tolerances. Note that due to the non-convex nature of the studied problem, it is also useful to add a constraint on the maximum number of iterations N_{max} as a stopping criterion of the iterative procedure. A final issue to discuss is the initialization of the technique which plays also an important role due to the non-convexity of the involved cost function. The most common approaches followed in literature are the use of random initialization or VCA-FCLS initialization [6], [9]. For the simulations presented here we adopted the latter approach since it provides approximations close to the optimal points that usually improve the performance of the NMF-based methods.

The proposed blind (centralized) unmixing algorithm is summarized in Algorithm 1.

V. Distributed Sparsity and Low-Rank Constrained NMF

In this section, the distributed/parallel version of the proposed method is developed. The description is tailored for a multi-core system architecture, albeit it can be easily adjusted to a distributed one (where each processing unit is on a different node), as it is also shown later in this section. Let us start with the estimation of the abundance matrix $\mathbf{S}$. It is noteworthy, that given that the endmember matrix $\mathbf{A}_n$ is estimated, each

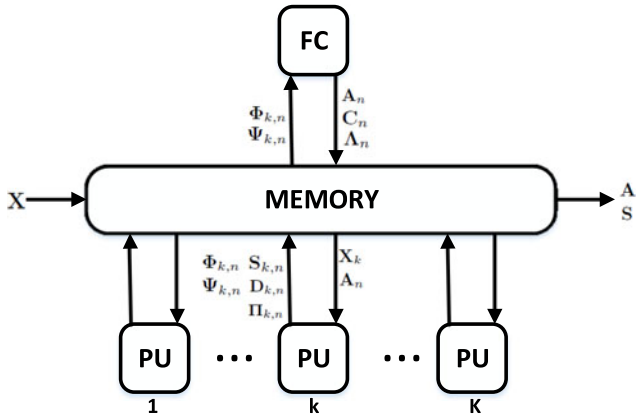


Fig. 1. A parallel architecture with K Processing Units (PU) sharing a common memory with a Fusion Center (FC).

abundance sub-matrix, $\mathbf{S}_{k,n}$ is updated via the kth optimization problem in (24). These K optimization problems are entirely decoupled and thus, parallel estimation of $\mathbf{S}_k$'s in a modern multi-core computational system (CPU and/or GPU) is quite straightforward. In more detail, Fig. 1 shows the diagram of a hypothetical parallel/distributed network. This network consists of at least K processing units that share a common memory in which the complete hyperspectral image $\mathbf{X}$ is stored. The K processing units are able to access in parallel the non-overlapping parts of the image $\mathbf{X}_k$ and the matrices $\mathbf{A}_n$, $\mathbf{D}_{k,n-1}$ and $\mathbf{\Pi}_{k,n-1}$ that are necessary for the update of $\mathbf{S}_{k,n}$ through (25). The description that will follow is also applicable for the case where $K' < K$ processing units are available. If this is the case, multiple sub-images are handled by each processing unit, though the gains from the parallel processing are less, since now $\lceil K/K' \rceil$[1] processing steps are required for the calculation of the corresponding quantities, instead of one. Note that matrices $\mathbf{D}_{k,n}$ and $\mathbf{\Pi}_{k,n}$ can also be updated in a parallel manner via (27) and (17), respectively by using the processing network since, their corresponding optimization problems are again decoupled with respect to the optimizing variables. On the contrary, the distributed/parallel estimation of $\mathbf{A}$ is not a trivial task since each sub-image involves the whole matrix and thus, its estimation should be done jointly.

By inspecting (19), the coupling is identified in the terms $\mathbf{\Psi}_n \equiv \mathbf{X}\mathbf{S}_n^T$ and $\mathbf{\Phi}_n \equiv \mathbf{S}_n\mathbf{S}_n^T$ that require the complete image and abundance matrices for their computation. It is instructive, to remind that each of the processing units has access to matrices $\mathbf{X}_k$ and $\mathbf{S}_{k,n}$ and the aim is to distribute the computational cost of the terms $\mathbf{\Psi}_n$ and $\mathbf{\Phi}_n$ among them. To that end, the following equations are used

$$\mathbf{\Psi}_n = \sum_{k=1}^{K} \mathbf{X}_k\mathbf{S}_{k,n}^T \text{ and } \mathbf{\Phi}_n = \sum_{k=1}^{K} \mathbf{S}_{k,n}\mathbf{S}_{k,n}^T. \qquad (29)$$

From (29), the key point now is the computation of $\mathbf{\Phi}_n$ and $\mathbf{\Psi}_n$ in a distributed/parallel manner. Each processing unit is capable of computing in parallel the quantities $\mathbf{S}_{k,n}^T\mathbf{S}_{k,n}$ and $\mathbf{S}_{k,n}^T\mathbf{X}_k$ and then, the sums in (29) can be computed in a processing unit that plays the role of the fusion center (which can be an additional processing unit or one out of the K existing ones, as well), as it is shown in Fig. 1. In the system of Fig. 1, the fusion center then uses (19), (21) and (16) to update the corresponding variables. Note that the fusion node is also responsible for checking the convergence criteria of the procedure. The distributed version of the proposed blind unmixing method is shown in Algorithm 2.

[1] $\lceil \cdot \rceil$ is the smallest integer greater than or equal to the given operand.

Algorithm 2: Distributed Sparsity and Low-Rank Constrained NMF.

Input: $\mathbf{X}$, p, K, α, γ, λ, tol, N_{max}

Initialize $\mathbf{A}$ and $\mathbf{S}$ by the VCA-FCLS techique to $\mathbf{A}_0$, $\mathbf{S}_0$

Set $\mathbf{\Lambda}_0$ and $\mathbf{D}_{k,0}$ as zero entries matrices

while The criteria in (28) are not met & $n \leq N_{max}$ **do**

  parfor $1 \leq k \leq K$ **do**

$$\mathbf{S}_{k,n} = SHR_{\frac{\lambda}{\alpha}}\left\{\left(\mathbf{A}_{n-1}^T\mathbf{A}_{n-1} + \alpha\mathbf{I}_P\right)^{-1} \times \left(\mathbf{A}_{n-1}^T\mathbf{X}_k - \mathbf{\Pi}_{k,n-1} + \alpha\mathbf{D}_{k,n-1}\right)\right\}$$

$$\mathbf{D}_{k,n} = \Pi_{\mathbb{R}_+}\left\{SVT_{\frac{\gamma}{\alpha}}\left\{\mathbf{S}_{k,n} + \tfrac{1}{\alpha}\mathbf{\Pi}_{k,n-1}\right\}\right\}$$

$$\mathbf{\Pi}_{k,n} = \mathbf{\Pi}_{k,n-1} + \alpha\left(\mathbf{S}_{k,n} - \mathbf{D}_{k,n}\right)$$

$$\mathbf{\Phi}_{k,n} = \mathbf{S}_{k,n}\mathbf{S}_{k,n}^T$$

$$\mathbf{\Psi}_{k,n} = \mathbf{X}_k\mathbf{S}_{k,n}^T$$

  end parfor

$$\mathbf{\Phi}_n = \sum_{k=1}^{K} \mathbf{\Phi}_{k,n}$$

$$\mathbf{\Psi}_n = \sum_{k=1}^{K} \mathbf{\Psi}_{k,n}$$

$$\mathbf{A}_n = \left(\mathbf{\Psi}_n - \mathbf{\Lambda}_{n-1} + \alpha\mathbf{C}_{n-1}\right)\left(\mathbf{\Phi}_n + \alpha\mathbf{I}_L\right)^{-1}$$

$$\mathbf{C}_n = \Pi_{\mathbb{R}_+}\left\{\mathbf{A}_n + \tfrac{1}{\alpha}\mathbf{\Lambda}_{n-1}\right\}$$

$$\mathbf{\Lambda}_n = \mathbf{\Lambda}_{n-1} + \alpha(\mathbf{A}_n - \mathbf{C}_n)$$

end while

Return $\mathbf{A}$, $\mathbf{S}$

At the final part of this section, we describe how the proposed method can be applied in distributed sensing / processing nodes across a network [46]–[48]. Let us assume that K nodes have sensing and processing capabilities, so each one of them acquires an image of same dimensions that is stored in matrix $\mathbf{X}_k$, as described in Sections III-IV. If a fusion center exists, it is straightforward to see that given that the matrix $\mathbf{A}_{n-1}$ is available at each one of the nodes, matrices $\mathbf{S}_{k,n}$, $\mathbf{D}_{k,n}$ and $\mathbf{\Pi}_{k,n}$ may be updated at each node in parallel, since all the required information is locally known. Then, the quantities $\mathbf{X}_k\mathbf{S}_{k,n}^T$ and $\mathbf{S}_{k,n}\mathbf{S}_{k,n}^T$ are locally computed at each node, again and the result is forwarded to the fusion center which via (29) computes $\mathbf{\Psi}_n$ and $\mathbf{\Phi}_n$ that are later used in the global estimation of $\mathbf{A}_n$. Finally, the new version of $\mathbf{A}_n$ is returned to the nodes that continue the calculations upon convergence. In the absence of a fusion center, distributed averaging methods can be employed for estimating $\mathbf{\Psi}_n$ and $\mathbf{\Phi}_n$ [46]. Now at every iteration of the method, each node computes local estimations of $\mathbf{\Psi}_n$ and $\mathbf{\Phi}_n$ which are then used to estimate also locally, the matrices $\mathbf{A}_n$, $\mathbf{C}_n$ and $\mathbf{\Lambda}_n$ via (19), (21) and (16), respectively. In general,

an implementation method would depend on the parallel architecture of a multi-core system or the network infrastructure of the distributed nodes, which are both beyond the scopes of the present paper. In the experiments presented in Section VI, for the sake of simplicity, we considered the multi-core structure of Fig. 1.

VI. Convergence Of the Algorithm

In this section, we present very briefly a result concerning the convergence of the proposed algorithm. Note that in literature so far, strong convergence results for the ADMM have been derived for convex problems that involve only two sets of variables. Moreover, strong convergence results, for non-convex problems when the ADMM sequence is applied, are in general unknown and the proof of convergence for such schemes remains an open research problem. Here, the optimization problem under consideration involves four blocks of variables and on top of that the cost function is non-convex. Thus, obtaining strong convergence results is an intractable task, if not impossible, and certainly beyond the scope of the present paper.

The following lemma gives some conditions under of which the sequence generated by the alternating minimization steps of eqs. (12)-(17) converges to a limit point that satisfies the Karush-Kuhn-Tucker (KKT) conditions. The proof of this Lemma can be derived by following the analysis presented in [49].

Lemma 1: Let $\mathbf{Z}_n = \{\mathbf{A}_n, \bar{\mathbf{S}}_n, \bar{\mathbf{D}}_n, \mathbf{\Lambda}_n, \mathbf{\Pi}_n\}$ with $\bar{\mathbf{S}}_n = \{\mathbf{S}_{1,n}, \ldots, \mathbf{S}_{K,n}\}$ and $\bar{\mathbf{D}}_n = \{\mathbf{D}_{1,n}, \ldots, \mathbf{D}_{K,n}\}$ a sequence generated by the alternating minimization steps of eqs. (12)-(17). If the multiplier sequence $\{\mathbf{\Lambda}_n, \mathbf{\Pi}_{k,n}\}$ is bounded and satisfies

$$\sum_{n=0}^{\infty}\left(\|\mathbf{\Lambda}_n - \mathbf{\Lambda}_{n-1}\|_F^2 + \sum_{k=1}^{K}\|\mathbf{\Pi}_{k,n} - \mathbf{\Pi}_{k,n-1}\|_F^2\right) < \infty, \tag{30}$$

then $\mathbf{Z}_n - \mathbf{Z}_{n-1} \to \mathbf{0}$, where $\mathbf{0}$ denotes a matrix with zeros as entries and thus, the sequence $\{\mathbf{Z}_n\}$ converges to a limit point. Furthermore, these limit points satisfy the KKT conditions (first order optimality conditions) of $(\mathcal{P}_4)$.

In general Lemma 1 states that if the multiplier sequence $\{\mathbf{\Lambda}_n, \mathbf{\Pi}_{k,n}\}$ converges to a limit point (30), then the ADMM sequence $\{\mathbf{Z}_n\}$ converges also to a limit point that satisfies the KKT conditions of $(\mathcal{P}_4)$.

VII. Experiments

In this section the performance of the proposed approach (termed "DSPLR-NMF" in the relevant figures) is evaluated in experiments on synthetic and real hyperspectral data and compared to the ones of the classical Projected Gradient NMF ("PG-NMF"), the $L_{1/2}$-NMF [18], the Hyperspectral Unmixing Turbo-Approximate Message Passing[2] ("HUT-AMP") [35] and the VCA [6] algorithms. Note that the experiments were based only on the distributed version of the proposed algorithm (Sec. V), since, as it is evident from the description given in Sections IV–V, both the centralized and the distributed version compute at each step the same quantities and update the corresponding variables of the optimization problem via identical expressions and thus, they will eventually exhibit identical performance. To that end, the two approaches are only compared with respect to the required mean processing time (see Fig. 5), so as to highlight the gains from a distributed/parallel approach. All the experiments were executed on a cluster consisted of eight core systems based on the Intel Xeon E5-268@2.7GHz processor.

The evaluation/comparison is based on the Spectral Angle Distance (SAD) and Root Mean Square Error (RMSE) metrics defined as

$$\mathrm{SAD}_p = \arccos\left(\frac{\mathbf{A}_p^T\hat{\mathbf{A}}_p}{\|\mathbf{A}_p\|\|\hat{\mathbf{A}}_p\|}\right) \tag{31}$$

and

$$\mathrm{RMSE}_p = \left(\frac{1}{N}|\mathbf{S}_p - \hat{\mathbf{S}}_p|^2\right)^{1/2}, \tag{32}$$

respectively, where $\mathbf{A}_p$ is the ground truth spectral signature of the pth endmember and $\hat{\mathbf{A}}_p$ its corresponding estimate and $\mathbf{S}_p$ is the ground truth abundance vector of all the pixels for the pth endmember and $\hat{\mathbf{S}}_p$ its corresponding estimated vector.

A. Synthetic Data

The synthetic data experiments are based on spectral signatures of endmembers chosen from the United States Geological Survey (USGS) digital spectral library [50]. The ground truth abundances are generated according to [15]. That is, we divide a $m^2 \times m^2$ image in $m \times m$ sub-images. Each sub-image is initialized with the same ground cover type which is one endmember that is selected randomly each time. Then, a $(m+1) \times (m+1)$ low pass filter is applied on the image to generate mixed pixels and force the abundances to vary smoothly. Finally, a threshold η is used so as to increase the mixture level of the pixels. To that end, if a pixel has an abundance value that is greater than η, it is replaced by a mixture of all the endmembers with equal abundances. Typical values for η is between 0.6 and 1.

The outcome of the previous procedure is a synthetic hyperspectral image. In order to study the impact of noise on the proposed unmixing technique, zero mean white Gaussian noise is added to the synthetic image and the Signal-to-Noise Ratio (SNR) is defined as

$$\mathrm{SNR} = 10\log_{10}\frac{\mathbb{E}\{\mathbf{x}_n^T\mathbf{x}_n\}}{\mathbb{E}\{\mathbf{e}_n^T\mathbf{e}_n\}}, \tag{33}$$

where $\mathbf{x}_n$ and $\mathbf{e}_n$ are defined in (1) and $\mathbb{E}\{\cdot\}$ is the expectation operator. In the following, the results of different experiments are presented in order to evaluate the performance of the proposed approach with respect to the different involved parameters.

In Fig. 2 the two performance metrics of (31)-(32) are depicted for different SNR values. The simulation parameters for generating the hyperspectral data are $\eta = 0.7$, $m = 8$. The parameters of DSPLR-NMF are set to $\gamma = 0.1$, $\lambda = 0.05$,

[2]We used the Matlab code which can be found in http://www2.ece.ohio-state.edu/~schniter/HUTAMP/HUTAMP.html

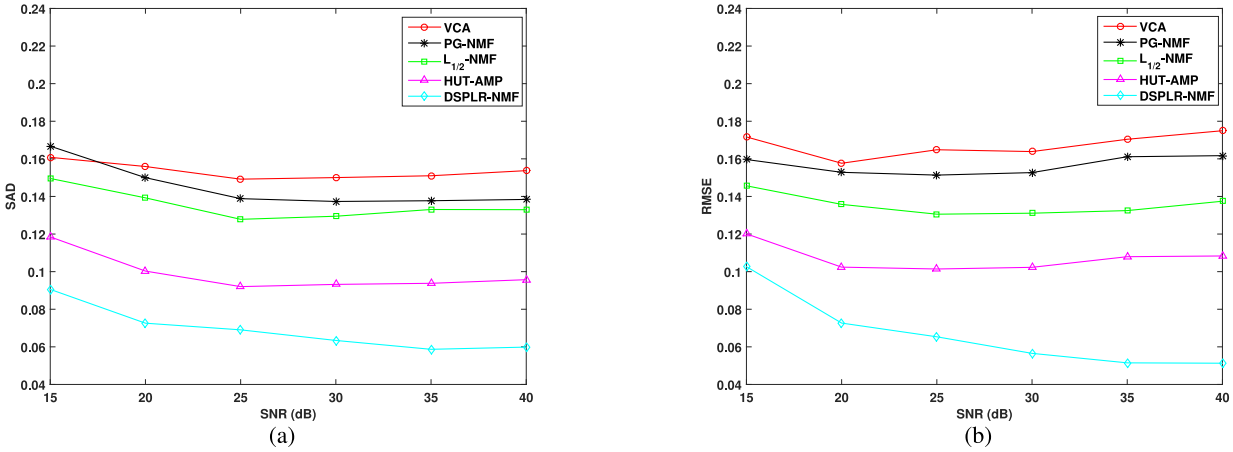


Fig. 2. Performance of the unmixing algorithms with respect to SNR on synthetic data. a) SAD and b) RMSE.

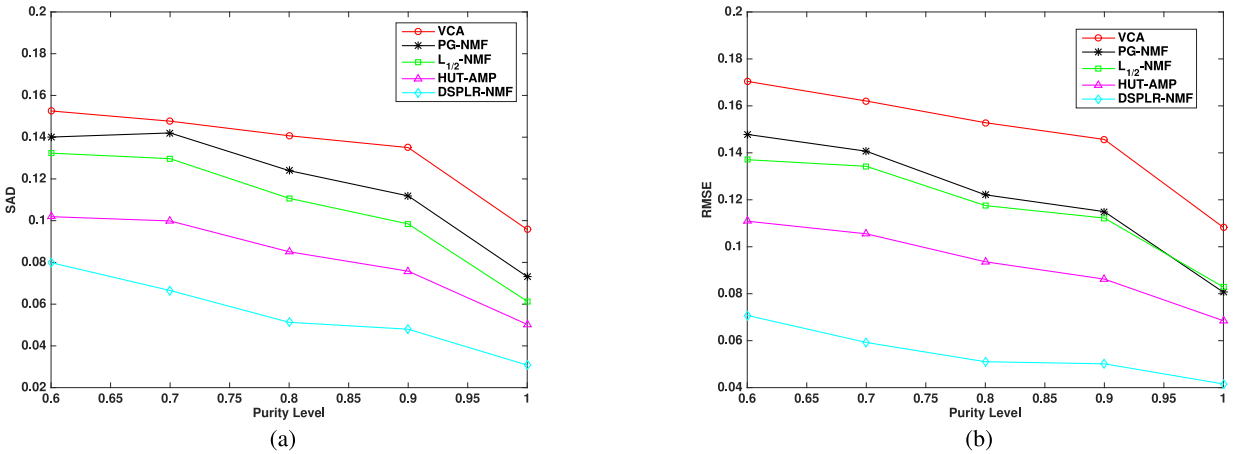


Fig. 3. Performance of the unmixing algorithms with respect to the pixel purity parameter η on synthetic data. a) SAD and b) RMSE.

$\alpha = \beta = 100$ and $r = 8$, while the sparsity controlling parameter of $L_{1/2}$-NMF is set according to the descriptions of [18]. Finally, all the tested methods are initialized by the outcome of the VCA, the maximum allowed number of iterations is set to 3000, all of the tolerances are set to 10^{-6} and the results are based on the average of 100 Monte Carlo simulations. As it is evident, the performance of the proposed technique is clearly superior compared to that of existing algorithms in terms of both the SAD and the RMSE metrics.

Let us now examine the impact of the pixel purity parameter on the performance of the tested techniques. To that end, we assume the same experimental setup with the one of Fig. 2 though now the SNR is fixed to 25 dB and parameter η varies between 0.6 and 1. The parameters of the algorithms are selected as in the previous experiment. In Fig. 3 SAD and RMSE are plotted with respect to η. As it was expected, all the methods present improved performance as the pixel purity parameter increases, with the proposed algorithm achieving the best performance in terms of both performance metrics and for all the examined values of η.

In Fig. 4, SAD and RMSE are plotted as the number of pixels of the hyperspectral image increases. In this experiment, m varies between 8 and 12, SNR = 25 dB and all remaining settings are similar to those of the first experiment. As it is shown in the figure, DSPLR-NMF outperforms again its competing schemes in terms of both SAD and RMSE.

The mean processing time required by each one of the competing algorithms to converge to its solution with respect to the size of the image (number of pixels), is depicted in Fig. 5.(a). Apart from the distributed version of the proposed algorithm which runs on a 8-core system, the performance of the centralized one (denoted by "SPLR-NMF") is also depicted in the same figure. It is apparent that the proposed algorithm (both the centralized and the distributed versions) is much faster than PG-NMF and $L_{1/2}$-NMF especially as the number of pixels increases. The superiority of the proposed approach can be attributed to the following facts 1) DSPLR-NMF stems from the ADMM which appears to have good behavior for the non-convex problem examined here whereas PG-NMF and $L_{1/2}$-NMF are based on the multiplicative update rule [22] which in general exhibits slow convergence and 2) The proposed approach takes simultaneously into account the sparse and low rank structure of the abundance matrix while, on the other hand, in $L_{1/2}$-NMF only the sparse structure is considered and PG-NMF does not assume any structure (apart from the non-negativity) of the latter matrix.

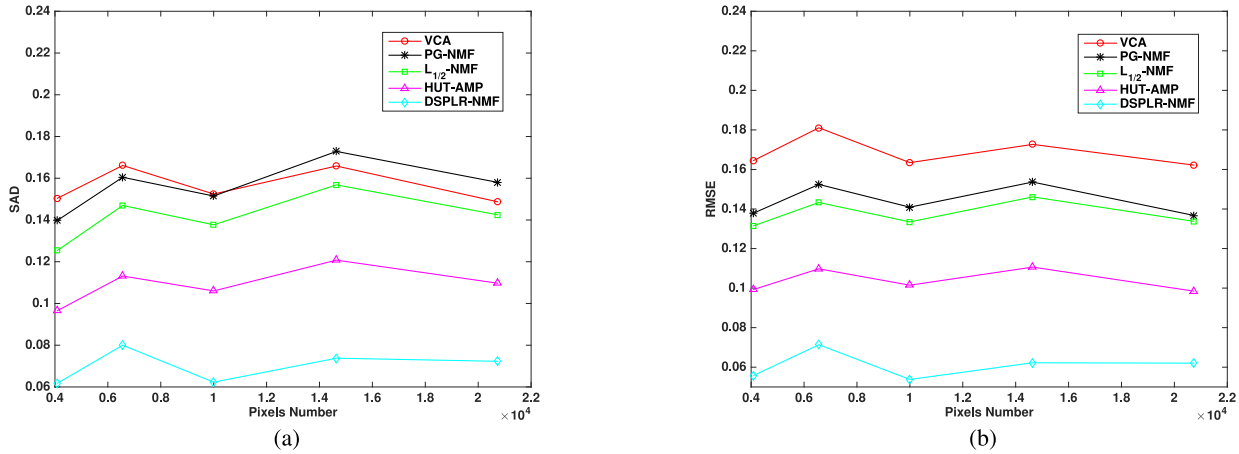


Fig. 4. Performance of the unmixing algorithms with respect to the number of pixels on synthetic data. a) SAD and b) RMSE.

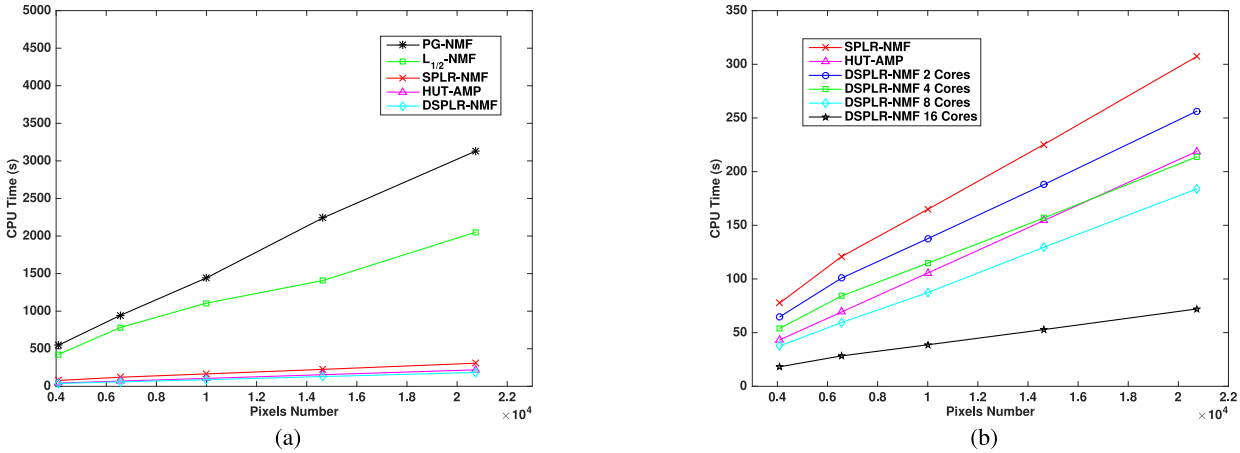


Fig. 5. CPU time of the DSPLR-NMF algorithm with respect to the number of pixels on synthetic data. a) Comparison with the other unmixing approaches b) Performance with respect the number of the employed cores.

The new technique is also compared to the HUT-AMP algorithm [35] which is based on a Bayesian approach and not on the NMF. From Fig. 5.(a), the HUT-AMP approach presents slightly less mean processing time than the centralized version of the proposed approach. On the contrary, the distributed one exploits parallelism and manages to reduce the required mean processing time below that of the HUT-AMP. Obviously, if more processing units were available, the processing time could be further reduced by increasing the level of parallelism. To highlight better the previous, the impact of the number of cores to the CPU time required by the DSPLR-NMF is depicted in Fig. 5.(b). There, the CPU time of the DSPLR-NMF versus the number of pixels is plotted for 2, 4, 8 and 16 cores. In the same figure, the CPU time of the centralized version (single core) and of HUT-AMP are again depicted for comparison purposes. As it is evident, an increase in the number of cores results in a decrease in the required CPU time of DSPLR-NMF. Note, that for only 4 cores, the DSPLR-NMF requires CPU-time very close (and even better for large number of pixels), to the HUT-AMP approach. All in all, the proposed algorithms achieve significantly improved performance compared to the HUT-AMP for comparable if not less processing time while being at the same time much simpler.

Let us now examine the impact of inaccurate estimation of the number of endmembers on the performance of the unmixing techniques under study. As it is evident from the analysis given in Sec. II-IV, a crucial parameter for all unmixing techniques is the number of endmembers present in the scene. This number is, in general, unknown and usually it is estimated prior to unmixing by employing relevant literature methods e.g., [51]-[52]. However, these techniques may provide erroneous estimates of the number of endmembers that may affect the performance of unmixing algorithms. In Fig. 6 SAD and RMSE are depicted under the experimental setup of Fig. 2, with the number of endmembers varying from 6 (which is the correct one) to 9. In the x-axis of Fig. 6 the difference from the actual endmember number is depicted. As it is shown, the performance of all the algorithms degrades, as the difference to the actual endmember number increases. It is noteworthy that the performance of the proposed algorithm is again better under all the examined cases.

Similar conclusions with the previous experiments can be drawn if we compare the new algorithm with the other related

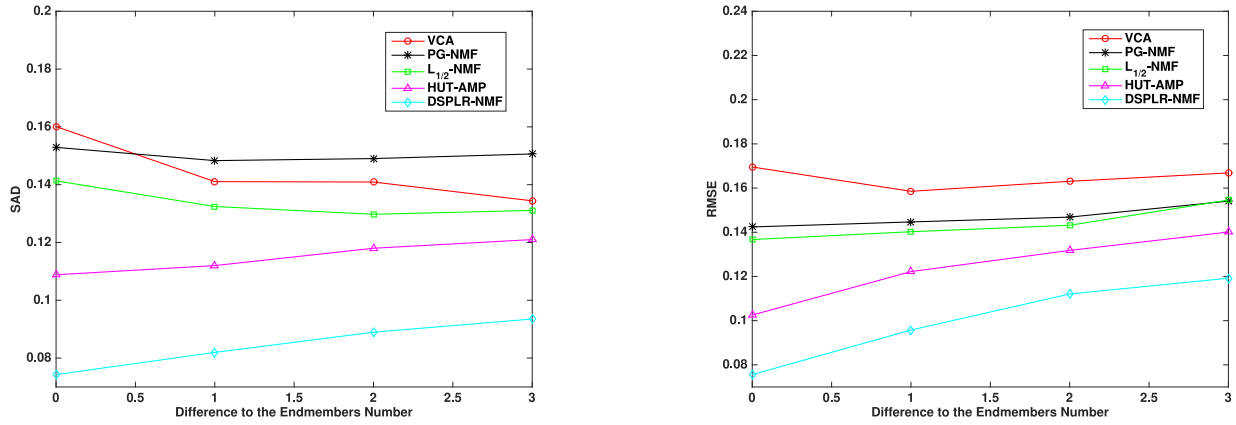


Fig. 6. Performance of the unmixing algorithms with respect to the difference to the actual endmembers number on synthetic data. a) SAD and b) RMSE.

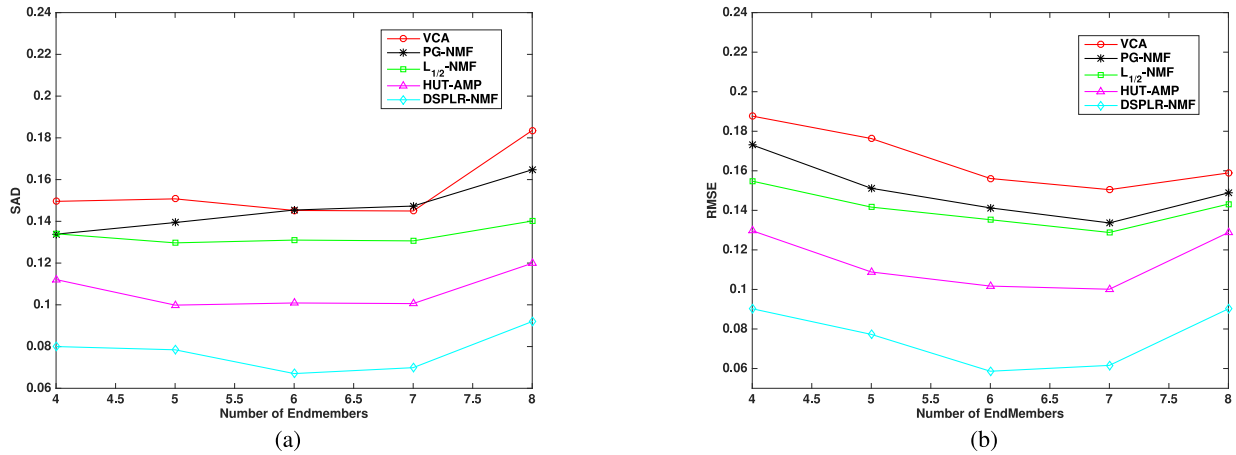


Fig. 7. Performance of the unmixing algorithms with respect to the endmembers number on the synthetic data scene. a) SAD and b) RMSE.

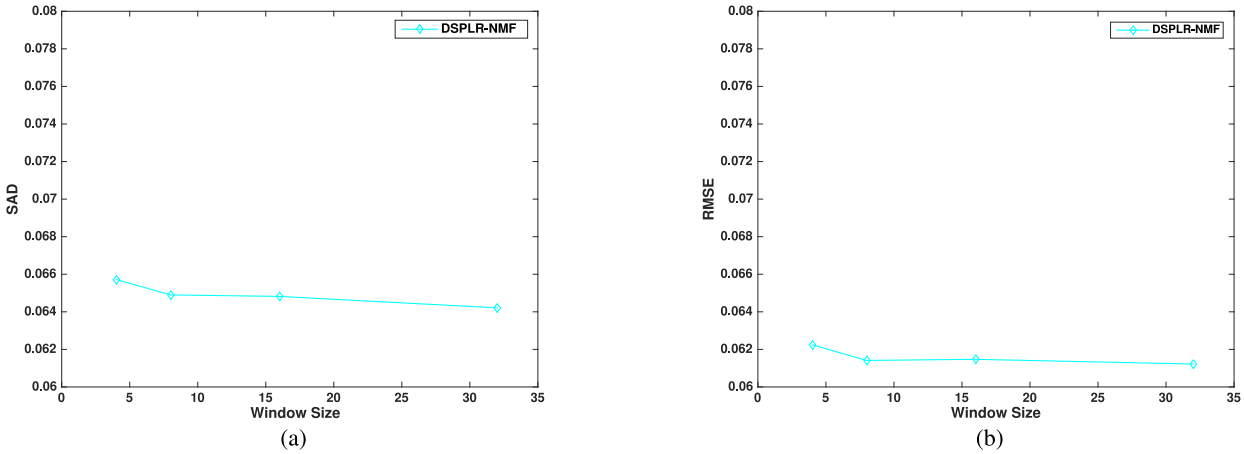
unmixing schemes in case the number of endmembers in the image increases. This is clearly shown in Fig. 7, where again DSPLR-NMF outperforms all other related algorithms in terms of SAD and RMSE.

In Fig. 8 we examine how the performance of the proposed algorithm is affected by the size of the processing window r. To that end, we vary the value of r and plot the SAD and RMSE of DSPLR-NMF. As it is shown, the proposed algorithm appears to be almost insensitive to the size of the window used, and so the value of parameter r can be tuned easily in order to optimize its performance.

We close the study on the synthetic data-set by presenting results regarding the impact of parameters λ, γ, α and β on the performance of the DSPLR-NMF. To that end, the experimental set-up of Fig. 2 is once more considered, for $SNR = 25$ dB and the results are produced by varying the value of the parameter under examination (the rest of the parameters are set to fixed values as in the experiments of Fig. 2). The SAD and RMSE values are shown in Figs 9, 10 and 11 for the parameters λ, γ and $\alpha = \beta$, respectively. It is clear that, all the parameters have a range of values that result in satisfactory unmixing performance. Moreover, the results justify also the selection of the parameters' values that were used in the experiments of Figs. 2–8.

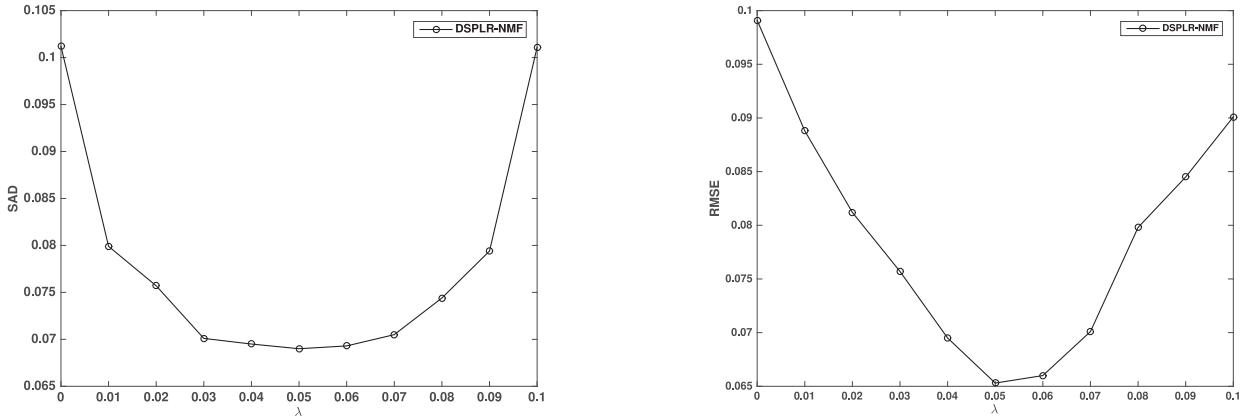
B. Real Hyperspectral Data

In this subsection we evaluate the performance of the proposed approach for two real hyperspectral data sets. The first one is the well-known Indian Pines data set captured by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) in north-western Indiana. The data set consists of an 145×145 hyperspectral image taken in 220 bands that cover the wavelength range $0.4 - 2.5$ μm. The image spectral and spatial resolutions are 10nm and 17m, respectively. Two thirds of the imaged scene are crops and almost one third is forest or other natural vegetation. Moreover, there is a number of man-made areas, including two highways, one railway, some buildings and smaller roads. A false color image of the scene is shown in Fig. 12 by assigning the 70th, 99th and 136th bands to the R, G, and B components, respectively. A ground truth analysis of the image is reported in [53]. The ground area is divided into 17 classes ingoring many small variations within the fields that can be seen in the image. In order to simplify the evaluation of the unmixing techniques, classes with similar spectral characteristics are merged together resulting into the following six typical classes: Corn which includes Corn-notill, Corn-mintill and Corn, Wheat that includes Oat and Wheat, Natural Vegetation in which Alfalfa, Grass/Pasture, Grass/Trees and Woods are



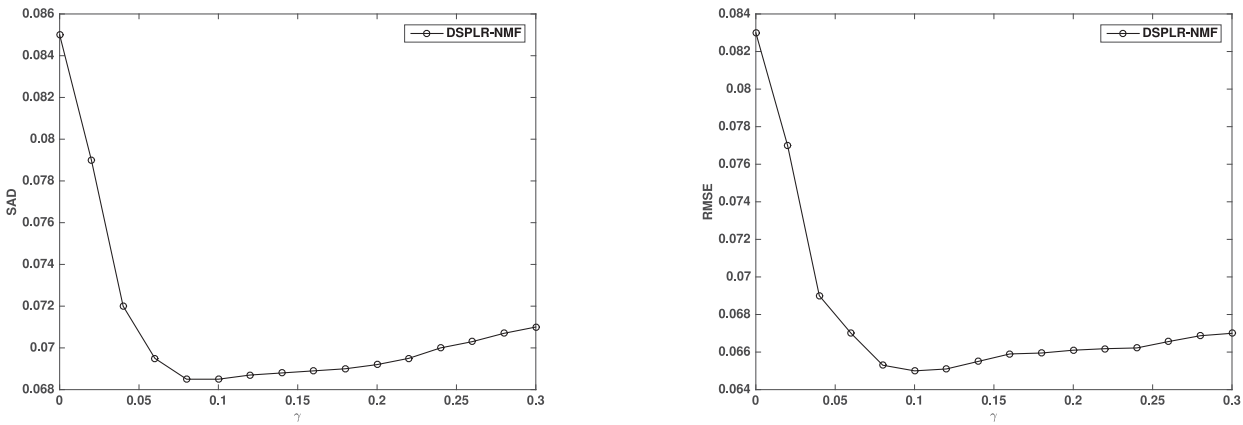
(a) (b)

Fig. 8. Performance of the unmixing algorithms with respect to the processing window size on synthetic data. a) SAD and b) RMSE.



(a) (b)

Fig. 9. Performance of the DSPLR-NMF algorithm with respect to the sparsity controlling parameter λ on synthetic data. a) SAD and b) RMSE.



(a) (b)

Fig. 10. Performance of the DSPLR-NMF algorithm with respect to the rank controlling parameter γ on synthetic data. a) SAD and b) RMSE.

merged, Artificial Structures which includes Buildings-Grass-Trees-Drives and Stone-Steel-Towers, Hay that includes Hay-windrowed and Grass/Pasture-mowed and Soyebean in which Soybin-mintill and Soybean-clean are merged.

Prior to unmixing, the 1st-4th, 78th-82nd, 103rd-115th, 148th-166th, and 211-220th bands are removed as noisy or water-absorption bands. The resulting hyperspectral image has 169 bands left. The endmember reference signatures are extracted artificially by using the ground truth [53]. Thus, 15, 10, 20, 10, 10 and 15 pure or near pure pixels are selected which correspond to corn, wheat, natural vegetation, man-made lands, haystack and soybean, respectively. The reference spectral

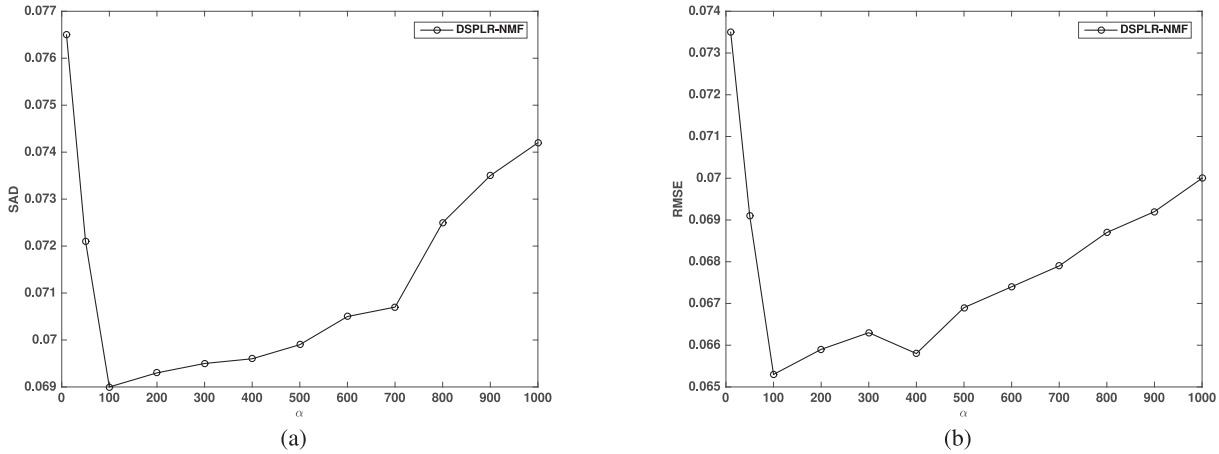


Fig. 11. Performance of the DSPLR-NMF algorithm with respect to the ADMM penalty parameter $\alpha = \beta$ on synthetic data. a) SAD and b) RMSE.

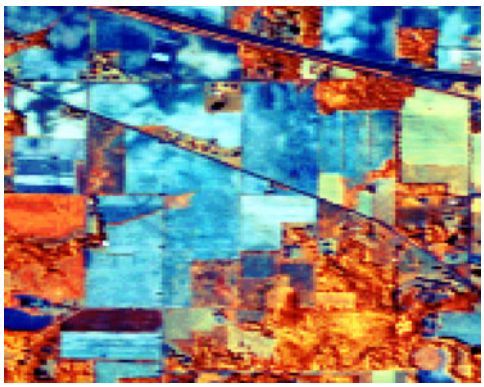


Fig. 12. False color image for the Indian Pines data set.

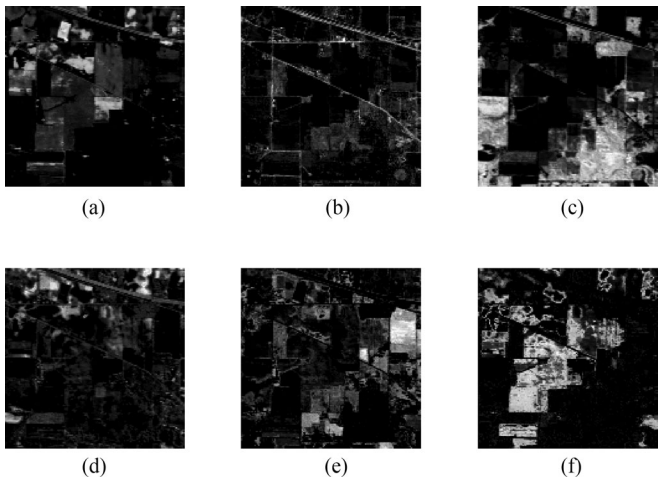


Fig. 13. Abundance maps provided by DSPLR-NMF for the Indian Pines data set. a) Corn, b) Man-made land, c) Natural Vegetation, d) Wheat, e) Haystack and f) Soybean.

signatures of the endmembers corresponding to each class are computed by averaging the aforementioned observations.

The abundance maps of the six previously mentioned endmembers-materials obtained by applying the proposed algorithm are shown in Fig. 13. The parameters of the algorithm have been set as $\gamma = 1$, $\lambda = 0.05$, $\alpha = \beta = 100$ and $r = 8$. It

TABLE I
SAD RESULTS ON THE INDIAN PINES DATA SET.

	DSPLR-NMF	HUT-AMP	$L_{1/2}$-NMF	PG-NMF	VCA
Corn	**0.0574**	**0.1114**	**0.1238**	**0.1235**	**0.0980**
Man-made land	0.0772	0.1457	0.2950	0.3109	0.1987
Natural Vegetation	0.0716	0.0601	0.1224	0.1639	0.1021
Wheat	0.1248	0.0647	0.2258	0.2446	0.1889
Haystack	0.0352	0.0720	0.3934	0.3660	0.2588
Soybean	0.0559	0.1087	0.1026	0.1510	0.1335
Mean	0.0704	0.0938	0.2105	0.2266	0.1633

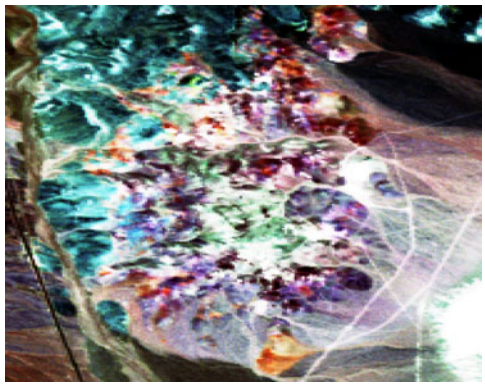


Fig. 14. False color image for the Culprite data set.

can be verified that these results are in close agreement with the ground truth reported in [53]. In addition, in order to quantitatively assess the performance of DSPLR-NMF for this real dataset, we compare it with that of the other related unmixing algorithms with respect to the SAD metric. The results are shown in Table I, where the SAD values per endmember and their mean are depicted for the four competing algorithms. Similar to the previous experiments on synthetic data, the algorithms are initialized using the VCA, the maximum allowed number of iterations is set to 3000 and the tolerances are set to 10^{-6}. As shown in Table I, DSPLR-NMF achieves the smaller SAD values for all endmembers, clearly outperforming its rivals in a real data scenario.

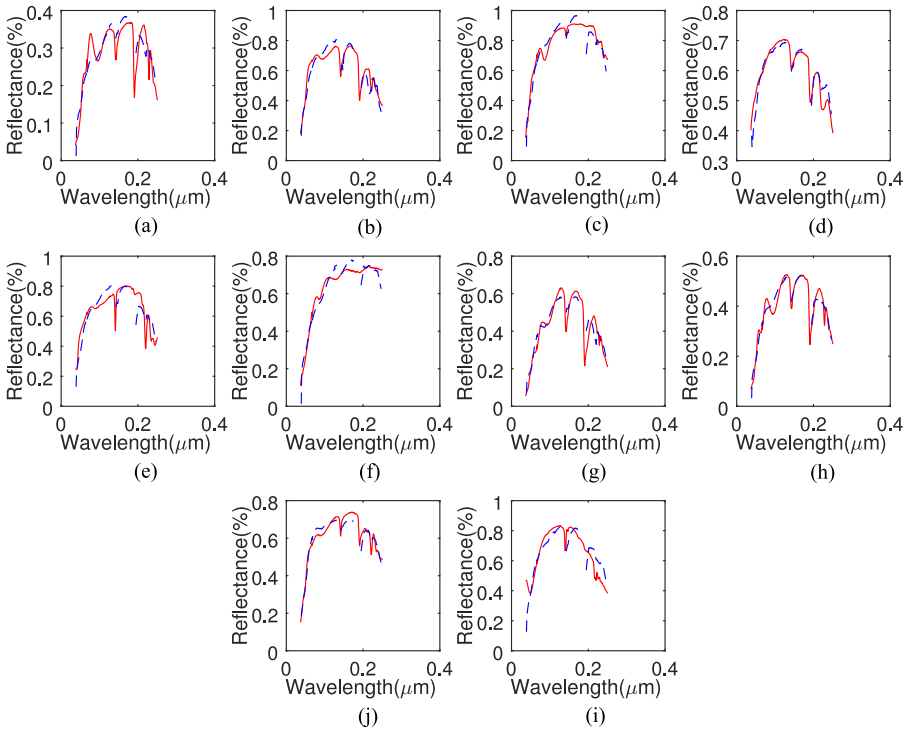


Fig. 15. Endmembers provided by the proposed method for the Cuprite data set: a) Alunite, b) Andradite, c) Buddingtonite, d) Dumotierite, e) Kaolinite, f) Muscovite, g) Montmorillonite, h) Nontronite, i) Pyrope and g) Sphene.

TABLE II
SAD RESULTS ON THE CULPRITE DATA SET.

	DSPLR-NMF	HUT-AMP	$L_{1/2}$-NMF	PG-NMF	VCA
Alunite	0.1077	0.1345	0.1652	0.1260	**0.1014**
Andradite land	0.0744	0.0705	**0.0531**	0.1195	0.0733
Buddingtonite	**0.0650**	0.0884	0.1611	0.0882	0.1254
Dumotierite	**0.0514**	0.1003	0.0899	0.1798	0.0974
Kaolinite	**0.1066**	0.0614	0.1423	0.1853	0.2523
Muscovite	**0.0593**	0.1142	0.1222	0.4252	0.1360
Montmorillonite	0.1001	**0.0651**	0.1302	0.1220	0.1338
Nontronite	**0.0649**	0.0848	0.0821	0.0817	0.0942
Pyrope	**0.0497**	0.0923	0.0588	0.0656	0.1262
Sphene	0.1100	0.1128	0.1012	0.1540	**0.0744**
Mean	**0.0789**	0.0920	0.1106	0.1547	0.1214

The second real-world data set is the Cuprite image acquired by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) sensor over the Cuprite region in South Nevada. This regional scene contains an abundant supply of minerals [54]. The image consists of 224 spectral bands in the range of $0.37 - 2.48$ μm and the spectral resolution is 10 nm. The minerals appearing in the scene are highly mixed and thus this data set is appropriate for evaluating the performance of the unmixing algorithms. For our purposes, a hyperspectral image of dimensions 250×192 is extracted from the original data set. A false color image is shown in Fig. 14 by assigning the 150th, 163th and 177th bands to the R, G, and B components, respectively.

A typical pre-processing procedure [18] was followed where the low SNR and water-vapor absorption bands (1-2, 104-113, 148-167 and 221-224) have been removed. In [6] a number of 14 types of minerals were identified. Nevertheless, here we searched for a number of $p = 10$ endmembers since the variants of the same mineral having slightly different spectra can be viewed as the same endmember. The simulation parameters for all the examined algorithms are the same to the ones used for the experiment on the Indian Pines data set. In Fig. 15, the estimated endmember signatures are compared to the ones of the USGS library spectra. It is clear that the extracted signatures are very close to the ones of the USGS library. Finally, in Table II we compare also the performance of the proposed approach to the one of the other methods with respect to the SAD metric (per endmember and mean value). As it is evident, the proposed method achieves better performance for most of the minerals in the image.

VIII. CONCLUSION

In this paper, a novel unmixing technique has been developed based on a simultaneously low-rank and sparsity constrained NMF. The proposed approach exploits the sparse and low-rank structure of the abundance matrix in order to constrain the set of feasible optimal points of the non-convex NMF problem and improve its performance. Moreover, the new approach was extended in a distributed manner so as to process in parallel different parts of the image and improve the convergence speed in multi-core systems. Indicative simulations on both synthetic and real world data show that the proposed approach achieves improved performance and faster convergence as compared to existing methods. Extension of the proposed work to a distributed hyperspectral sensors scenario is currently under development.

REFERENCES

[1] J. Bioucas-Dias *et al.*, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 354–379, Apr. 2012.

[2] T. M. Lillesand, *Remote Sensing and Image Interpretation*. New York, NY, USA: Wiley, 2006.

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

[4] G. Shaw and D. Manolakis, "Signal processing for hyperspectral image exploitation," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 12–16, Jan. 2002.

[5] M. E. Winter, "N-findr: An algorithm for fast autonomous spectral endmember determination in hyperspectral data," *Proc. SPIE*, vol. 3753, pp. 266–275, 1999.

[6] 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, Apr. 2005.

[7] J. W. Boardman *et al.*, "Automating spectral unmixing of aviris data using convex geometry concepts," in *Proc. Summaries 4th Annu. JPL Airborne Geosci. Workshop*, 1993, vol. 1, pp. 11–14.

[8] C.-I. Chang, C.-C. Wu, W. min Liu, and Y.-C. Ouyang, "A new growing method for simplex-based endmember extraction algorithm," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 10, pp. 2804–2819, Oct. 2006.

[9] D. Heinz and C.-I. Chang, "Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 3, pp. 529–545, Mar. 2001.

[10] M. Berman, H. Kiiveri, R. Lagerstrom, A. Ernst, R. Dunne, and J. Huntington, "Ice: A statistical approach to identifying endmembers in hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 10, pp. 2085–2095, Oct. 2004.

[11] J. Li and J. Bioucas-Dias, "Minimum volume simplex analysis: A fast algorithm to unmix hyperspectral data," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, vol. 3, Jul. 2008, pp. III-250–III-253.

[12] T.-H. Chan, C.-Y. Chi, Y.-M. Huang, and W.-K. Ma, "A convex analysis-based minimum-volume enclosing simplex algorithm for hyperspectral unmixing," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4418–4432, Nov. 2009.

[13] J. Nascimento and J. Bioucas Dias, "Does independent component analysis play a role in unmixing hyperspectral data?" *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 1, pp. 175–187, Jan. 2005.

[14] M. W. Berry, M. Browne, A. N. Langville, V. P. Pauca, and R. J. Plemmons, "Algorithms and applications for approximate nonnegative matrix factorization," *Comput. Statist. Data Anal.*, vol. 52, no. 1, pp. 155–173, 2007.

[15] L. Miao and H. Qi, "Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 3, pp. 765–777, Mar. 2007.

[16] S. Jia and Y. Qian, "Constrained nonnegative matrix factorization for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 1, pp. 161–173, Jan. 2009.

[17] P. O. Hoyer, "Non-negative matrix factorization with sparseness constraints," *J. Mach. Learn. Res.*, vol. 5, pp. 1457–1469, 2004.

[18] Y. Qian, S. Jia, J. Zhou, and A. Robles-Kelly, "Hyperspectral unmixing via $l_{1/2}$ sparsity-constrained nonnegative matrix factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4282–4297, Nov. 2011.

[19] N. Wang, B. Du, and L. Zhang, "An endmember dissimilarity constrained non-negative matrix factorization method for hyperspectral unmixing," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 554–569, Apr. 2013.

[20] X. Lu, H. Wu, Y. Yuan, P. Yan, and X. Li, "Manifold regularized sparse NMF for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 5, pp. 2815–2826, May 2013.

[21] K. E. Themelis, A. A. Rontogiannis, and K. Koutroumbas, "Semi-supervised hyperspectral unmixing via the weighted lasso," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, 2010, pp. 1194–1197.

[22] D. D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization," in *Proc. Adv. Neural Inf. Process. Syst.*, 2001, pp. 556–562.

[23] M. Ye, Y. Qian, and J. Zhou, "Multitask sparse nonnegative matrix factorization for joint spectral-spatial hyperspectral imagery denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2621–2639, May 2015.

[24] M. Fazel, "Matrix rank minimization with applications," Ph.D. dissertation, Stanford University, Stanford, CA, USA, 2002.

[25] M. Mardani, G. Mateos, and G. Giannakis, "Decentralized sparsity-regularized rank minimization: Algorithms and applications," *IEEE Trans. Signal Process.*, vol. 61, no. 21, pp. 5374–5388, Nov. 2013.

[26] M. Mardani, G. Mateos, and G. Giannakis, "Subspace learning and imputation for streaming big data matrices and tensors," *IEEE Trans. Signal Process.*, vol. 63, no. 10, pp. 2663–2677, May 2015.

[27] S. Wu, X. Zhang, N. Guan, D. Tao, X. Huang, and Z. Luo, *Non-negative Low-Rank and Group-Sparse Matrix Factorization*. New York, NY, USA: Springer-Verlag, 2015, pp. 536–547.

[28] M.-D. Iordache, J. Bioucas-Dias, and A. Plaza, "Sparse unmixing of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 6, pp. 2014–2039, Jun. 2011.

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

[30] Q. Qu, N. Nasrabadi, and T. Tran, "Abundance estimation for bilinear mixture models via joint sparse and low-rank representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 7, pp. 4404–4423, Jul. 2014.

[31] P. V. Giampouras, K. E. Themelis, A. A. Rontogiannis, and K. D. Koutroumbas, "Simultaneously sparse and low-rank abundance matrix estimation for hyperspectral image unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 4775–4789, Aug. 2016.

[32] E. Richard, P.-A. Savalle, and N. Vayatis, "Estimation of simultaneously 841 sparse and low rank matrices," in *Proc. 29th Int. Conf. Machine Learn.*, Edinburgh, Scotland, GB, 2012, pp, 1351–1358.

[33] M. Golbabaee and P. Vandergheynst, "Compressed sensing of simultaneous low-rank and joint-sparse matrices," *arXiv:1211.5058*, 2012.

[34] S. Oymak, A. Jalali, M. Fazel, Y. Eldar, and B. Hassibi, "Simultaneously structured models with application to sparse and low-rank matrices," *IEEE Trans. Inf. Theory*, vol. 61, no. 5, pp. 2886–2908, May 2015.

[35] J. Vila, P. Schniter, and J. Meola, "Hyperspectral unmixing via turbo bilinear approximate message passing," *IEEE Trans. Comput. Imag.*, vol. 1, no. 3, pp. 143–158, Sep. 2015.

[36] 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, 2011.

[37] G. Callico, S. Lopez, B. Aguilar, J. Lopez, and R. Sarmiento, "Parallel implementation of the modified vertex component analysis algorithm for hyperspectral unmixing using opencl," *IEEE J. Sel. Topics Appl. Earth Obser. Remote Sens.*, vol. 7, no. 8, pp. 3650–3659, Aug. 2014.

[38] E. Torti, G. Danese, F. Leporati, and A. Plaza, "A hybrid cpu–GPU real-time hyperspectral unmixing chain," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 2, pp. 945–951, Feb. 2016.

[39] P. L. Combettes and J.-C. Pesquet, "Proximal splitting methods in signal processing," in *Fixed-point Algorithms for Inverse Problems in Science and Engineering*. New York, NY, USA: Springer-Verlag, 2011, pp. 185–212.

[40] C. A. Bateson, G. P. Asner, and C. A. Wessman, "Endmember bundles: A new approach to incorporating endmember variability into spectral mixture analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 2, pp. 1083–1094, Mar. 2000.

[41] Z. Guo, T. Wittman, and S. Osher, "L1 unmixing and its application to hyperspectral image enhancement," *Proc. SPIE*, vol. 7334, 2009, Art. no. 73341M.

[42] J. M. Bioucas-Dias and M. A. T. Figueiredo, "Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing," in *Proc. 2nd Workshop Hyperspectral Image Signal Process. Evol. Remote Sens.*, Jun. 2010, pp. 1–4.

[43] M. D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Total variation spatial regularization for sparse hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 11, pp. 4484–4502, Nov. 2012.

[44] D. P. Bertsekas, "Nonlinear programming," Belmont, MA, USA: Athena Scientific, 1999.

[45] J.-F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM J. Optim.*, vol. 20, no. 4, pp. 1956–1982, 2010.

[46] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Syst. Control Lett.*, vol. 53, no. 1, pp. 65–78, 2004.

[47] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1989.

[48] H. Zhu, A. Cano, and G. Giannakis, "Distributed consensus-based demodulation: Algorithms and error analysis," *IEEE Trans. Wireless Commun.*, vol. 9, no. 6, pp. 2044–2054, Jun. 2010.

[49] Y. Xu, W. Yin, Z. Wen, and Y. Zhang, "An alternating direction algorithm for matrix completion with nonnegative factors," *Frontiers Math. China*, vol. 7, no. 2, pp. 365–384, 2012.
[50] R. Clark, G. Swayze, A. Gallagher, T. King, and W. Calvin, "The US geological survey digital spectral library: Version 1 (0.2 to 3.0 μm)," US Geological Surv., Denver, CO, USA, *Open File Rep. 93–592*, 1993.
[51] C.-I. Chang and Q. Du, "Estimation of number of spectrally distinct signal sources in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 3, pp. 608–619, Mar. 2004.
[52] J. Bioucas-Dias and J. Nascimento, "Hyperspectral subspace identification," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 8, pp. 2435–2445, Aug. 2008.
[53] D. Landgrebe, "Multispectral data analysis: A signal theory perspective," *Purdue Univ.*, West Lafayette, IN, USA, 1998.
[54] G. Swayze, R. Clark, F. Kruse, S. Sutley, and A. Gallagher, "Ground-truthing aviris mineral mapping at cuprite, nevada," in *Proc. 3rd 908 Annu. JPL Airborne Geosci. Workshop*, 1992, vol. 1, pp. 47–49.

Christos G. Tsinos (S'08–M'14) received the Diploma degree in computer engineering and informatics, the M.Sc. and the Ph.D. degree in signal processing and communication systems, and the M.Sc. degree in applied mathematics from the University of Patras, Patras, Greece, in 2006, 2008, 2013, and 2014, respectively. From August 2014 to June to 2015, he was a Postdoctoral Researcher at the University of Patras. Since July 2015, he has been as a Research Associate in the Interdisciplinary Centre for Security, Reliability and Trust, University of Luxembourg, Luxembourg, Europe. He is involved or was involved in the past in a number of different R&D projects funded by national and/or EU funds. He is currently the PI of R&D Project ECLECTIC (Energy and CompLexity EffiCienT mIllimeter-wave Large-Array Communications), funded under FNR CORE Framework. His current research interests include signal processing for mmWave, massive MIMO, cognitive radio and satellite communications and hyperspectral image processing, as well. He is a member of the Technical Chamber of Greece.

Athanasios A. Rontogiannis (M'97) received the (5-yr) Diploma degree in electrical engineering from the National Technical University of Athens, Athens, Greece, in 1991, the M.A.Sc. degree in electrical and computer engineering from the University of Victoria, Victoria, BC, Canada, in 1993, and the Ph.D. degree in communications and signal processing from the University of Athens, Athens, Greece, in 1997. From 1998 to 2003, he was a Lecturer at the University of Ioannina, Ioannina, Greece. In 2003, he joined the National Observatory of Athens, where he is currently a Research Director at the Institute for Astronomy, Astrophysics, Space Applications and Remote Sensing. His research interests include the general areas of statistical signal processing and wireless communications with emphasis on adaptive signal processing, hyperspectral image processing, compressive sensing, channel estimation/equalization, and multicarrier communications. He has served at the Editorial Boards of the *EURASIP Journal on Advances in Signal Processing*, Springer since 2008 and the *EURASIP Signal Processing Journal*, Elsevier since 2011. He is a member of EURASIP and the Technical Chamber of Greece.

Kostas Berberidis (S'87–M'90–SM'07) received the Diploma degree in electrical engineering from Democritus University of Thrace, Komotini, Greece, in 1985, and the Ph.D. degree in signal processing and communications from the University of Patras, Patras, Greece, in 1990. During 1991, he worked at the Signal Processing Laboratory of the National Defense Research Center. From 1992 to 1994 and from 1996 to 1997, he was a Researcher at the Computer Technology Institute, Patras, Greece. In period 1994/1995, he was a Postdoctoral Fellow at CCETT/CNET, Rennes, France. Since December 1997, he has been in the Computer Engineering and Informatics Department, University of Patras, where he is currently a Professor, and Head of the Signal Processing and Communications Laboratory. Also, since 2008, he has been Director of the Signal Processing and Communications Research Unit of the Computer Technology Institute and Press "Diophantus." His research interests include adaptive signal processing, distributed processing and learning, signal processing for communications, and wireless sensor networks.

Prof. Berberidis has served or has been serving as a member of scientific and organizing committees of several international conferences, as Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the IEEE SIGNAL PROCESSING LETTERS, as a Guest Editor for the *EURASIP Journal on Advances in Signal Processing,* and as Associate Editor for the *EURASIP Journal on Advances in Signal Processing*. He serves as a member of the Board of Directors of EURASIP, since January 2017. Also, since February 2010, he has been serving as Chair of the Greece Chapter of the IEEE Signal Processing Society. He is a member of the Signal Processing Theory and Methods Technical Committee of the IEEE Signal Processing Society and the Signal Processing for Communications and Electronics Technical Committee of the IEEE Communications Society. Moreover, since August 2015, he is a member of the EURASIP Special Area Teams Theoretical and Methodological Trends for Signal Processing and Signal Processing for Multisensor Systems. He is a member of the Technical Chamber of Greece, a member of EURASIP.