Sparse Unmixing of Hyperspectral Data Using Spectral a Priori Information

Wei Tang, Zhenwei Shi, *Member, IEEE*, Ying Wu, *Senior Member, IEEE*, and Changshui Zhang, *Member, IEEE*

Abstract

Given a spectral library, sparse unmixing aims at finding the optimal subset of endmembers from it to model each pixel in the hyperspectral scene. However, sparse unmixing still remains a challenging task due to the usually high mutual coherence of the spectral library. In this paper, we exploit the spectral a priori information in the hyperspectral image to alleviate this difficulty. It assumes that some materials in the spectral library are known to exist in the scene. Such information can be obtained via field investigation or hyperspectral data analysis. Then, we propose a novel model to incorporate the spectral a priori information into sparse unmixing. Based on the alternating direction method of multipliers (ADMM), we present a new algorithm termed sparse unmixing using spectral a priori information (SUnSPI) to solve the model. Experimental results on both synthetic and real data demonstrate that the spectral a priori information is beneficial to sparse unmixing and SUnSPI can exploit this information effectively to improve the abundance estimation.

Index Terms

Hyperspectral unmixing, sparse unmixing, alternating direction method of multipliers (ADMM), spectral a priori information.

Wei Tang is with Image Processing Center, School of Astronautics, Beihang University, Beijing 100191, PR China (e-mail: tangwei@sa.buaa.edu.cn).

Zhenwei Shi (Corresponding Author) is with Image Processing Center, School of Astronautics, Beihang University, Beijing 100191, PR China, with State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing 100191, PR China and also with Beijing Key Laboratory of Digital Media, Beihang University, Beijing 100191, PR China (e-mail: shizhenwei@buaa.edu.cn).

Ying Wu is with the Department of Electrical Engineering and Computer Science, Northwestern University, Evanston, IL 60208 USA (e-mail: yingwu@eecs.northwestern.edu).

Changshui Zhang is with the State Key Laboratory on Intelligent Technology and Systems, Tsinghua National Laboratory for Information Science and Technology (TNList), Department of Automation, Tsinghua University, Beijing 100084, China (e-mail: zcs@mail.tsinghua.edu.cn).

The work was supported by the National Natural Science Foundation of China under the Grants 61273245 and 91120301, the 973 Program under the Grant 2010CB327904, the Program for New Century Excellent Talents in University of Ministry of Education of China under the Grant NCET-11-0775, the funding project of State Key Laboratory of Virtual Reality Technology and Systems, Beihang University under the Grant VR-2014-ZZ-02, and the Fundamental Research Funds for the Central Universities under the Grant YWF-14-YHXY-028.

I. Introduction

With the development of imaging spectroscopy, hyperspectral remote sensors have been evolved to collect spectra with hundreds of narrow and contiguous wavelength bands from the visible region through the infrared region. However, because of the limited spatial resolution of the sensors and the combination of distinct materials into a homogeneous mixture, mixed pixels widely exist in the obtained hyperspectral data [1]. Hyperspectral unmixing, a great challenging task underlying many hyperspectral imagery applications, then aims at identifying the constituent spectra (endmembers) in the mixed pixel and estimating their corresponding fractions (abundances) [2, 3].

For the past few years, the linear unmixing model, which assumes that distinct endmembers are not interfered with each other, has been widely used for spectral unmixing [4, 5]. It has advantages such as ease of implementations and flexibility in various applications. Under this model, several unmixing approaches have been developed based on geometry [6–8], statistics [9] and nonnegative matrix factorization (NMF) [10, 11]. The advantage of these methods is that they are all unsupervised and need little additional a priori knowledge. Nevertheless, as these methods extract endmembers merely from the hyperspectral data, they either could obtain virtual endmembers with no physical meaning [12] or assume the presence in the data of at least one pure pixel per endmember, which is usually difficult to guarantee.

To improve the unmixing accuracy, the characteristics or a priori knowledge concerned with the hyperspectral data are often incorporated into the unmixing framework. For example, the NMF has been widely applied in hyperspectral unmixing in the last few years because it provides a part-based representation of the data, which makes the decomposition matrices intuitive and interpretable [13]. However, NMF could lead to non-unique solutions for the existence of local minima caused by the non-convexity of the objective function [14]. To alleviate this difficulty, the spatial smoothness of the abundances is considered in [15], the sparsity of the abundances is considered in [16], and the manifold structure of the hyperspectral data is considered in [17], all of which exploit the characteristics of the hyperspectral data to improve the unmixing results.

Recently, as more spectral libraries become open available, a semi-supervised approach, which takes the spectral library as a priori knowledge, has been proposed to tackle the problem concerned with the generation of virtual endmembers and the unavailability of pure pixels [18–20]. It aims at using only a few spectra in a given spectral library to model each mixed pixel in the hyperspectral imagery. As the number of actual endmembers present in the hyperspectral scene is usually much smaller than the number of spectra in the library, this approach (called sparse unmixing) often leads to a sparse solution. Several sparse unmixing methods based on sparse regression techniques [21, 22] have been proposed [18, 23–26] and show their advantages over the unsupervised approaches as they need not extract endmembers from the hyperspectral data and can estimate the abundances more accurately. Most recently, efforts have been paid to incorporate the spatial-contextual information [27–29] and subspace nature[1] [30–33] of the hyperspectral images into sparse unmixing algorithms to obtain much better results. The effectiveness of sparse unmixing indicates that it is critical and necessary to make full use of a priori knowledge of hyperspectral

[1]Subspace nature means all the pixels in the hyperspectral data share the same active set of endmembers included in the spectral library.

imagery to improve the unmixing accuracy.

Besides the spectral library, we draw attention to another kind of a priori knowledge concerned with the hyperspectral imagery, which we call spectral a priori information. With the research of hyperspectral data advancing, a priori knowledge of some endmember signatures could be available. Besides, in some scenes like mining areas, substances such as copper and iron could be identified by field investigation [34, 35]. Thus, in many situations, we can know that some materials exist in the hyperspectral scene. We can expect to obtain better unmixing results by incorporating this a priori information into hyperspectral unmixing. Indeed, in our previous work [36], with the assumption that some spectral signatures in the hyperspectral imagery are known in advance, we incorporate the spectral a priori information into the unsupervised NMF framework. Experimental results on both synthetic and real data demonstrate that this kind of a priori knowledge is beneficial to the extraction of other unknown endmembers and the estimation of abundances corresponding to all the endmembers.

In practical applications, the high mutual coherence of the spectral library, which is defined as the largest cosine between any two spectral signatures in the library [22], limits the performance of the sparse regression techniques. If we already know some actual substances present in the scene, which is indeed difficult for the sparse unmixing algorithms to identify in the large spectral library, it should be easier for us to estimate their abundances. In this way, the high mutual coherence of the spectral library should have a weaker impact on the unmixing. Thus, in this paper, we consider incorporating the spectral a priori information into the sparse unmixing model. Specifically, we assume that some materials in the spectral library are known to exist in the hyperspectral scene. Then, we propose a new sparse unmixing model to make full use of the spectral a priori information. Based on the alternating direction method of multipliers (ADMM) [37, 38], we develop a new sparse unmixing algorithm, called sparse unmixing using spectral a priori information (SUnSPI), to solve the newly formulated problem. It is worth mentioning that our work is an extension of [23] and [31]. The SUnSPI model is a combination of the model of sparse unmixing by variable splitting and augmented Lagrangian (SUnSAL) [23] and the newly proposed spectral a priori information (SPI) regularizer, which is based on the mixed $l_{2,1}$ norm [31, 39–41]. All the algorithms of SUnSAL, Collaborative SUnSAL (CLSUnSAL) [31] and the proposed SUnSPI are instances of the methodology introduced in [38]. Experimental results demonstrate that SUnSPI can exploit the spectral a priori information effectively.

The rest of the paper is structured as follows. Section II introduces the sparse unmixing model briefly. In Section III we present the model of sparse unmixing using the spectral a priori information. Experimental results are shown in Section IV. Finally, we conclude in Section V.

II. Sparse Unmixing Model

Let $\mathbf{A} \in R^{L\times m}$ denote the spectral library, where L is the number of bands and m is the number of spectral signatures in the library. The linear sparse unmixing model assumes that the observed spectrum vector of a mixed pixel $\mathbf{y} \in R^L$ is a linear combination of only a few spectral signatures in the spectral library:

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n} \tag{1}$$

where $\mathbf{x} \in R^m$ is the abundance vector with regard to the library $\mathbf{A}$ and $\mathbf{n} \in R^L$ is the vector of error term. Because the number of endmembers present in the hyperspectral scene is much smaller than the number of spectral signatures in the spectral library, the vector $\mathbf{x}$ should be sparse. Besides, subject to physical constraints, the model has the following two constraints (x_i is the i-th element of $\mathbf{x}$):

$$x_i \geq 0, \forall i \tag{2}$$

$$\sum_{i=1}^{m} x_i = 1 \tag{3}$$

which are called abundance nonnegativity constraint and sum-to-one constraint, respectively [42]. But, because the signature variability, which widely exists in the real hyperspectral images, will introduce nonnegative scaling coefficients from pixel to pixel in the signatures present in the hyperspectral scene, the sum-to-one constraint should be replaced by a so-called generalized sum-to-one constraint [18, 43]:

$$\sum_{i=1}^{m} \omega_i x_i = 1 \tag{4}$$

where ω_i ($i = 1 \dots m$) are the pixel-dependent scale coefficients. Besides, the nonnegativity of the sources automatically imposes a generalized sum-to-one constraint [18]. Thus, we do not explicitly consider the sum-to-one constraint in the sparse unmixing model in this paper.

Sparse unmixing aims at using only a few spectral signatures in the library to express each mixed pixel in the hyperspectral data. Thus, the optimization problem of sparse unmixing is

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{subject to} \quad \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \leq \delta, \mathbf{x} \geq 0 \tag{5}$$

where $\|\mathbf{x}\|_0$ (called l_0 norm) denotes the number of nonzero components in $\mathbf{x}$, $\delta \geq 0$ is the error tolerance due to the noise and modeling errors. Due to the discrete and nonconvex nature of l_0 norm, the problem in Eq. (5) is NP-hard [22], which means it can not be solved in polynomial time. A popular strategy is to replace l_0 norm with the l_1 norm which is naturally its best convex approximate. Then the new objective is

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{subject to} \quad \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \leq \delta, \mathbf{x} \geq 0. \tag{6}$$

Suppose that the hyperspectral data set contains K pixels organized in the matrix $\mathbf{Y} \in R^{L \times K}$ each column of which represents an observed mixed pixel. Then, Eq. (1) can be written in a compact matrix form:

$$\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{N} \tag{7}$$

where $\mathbf{X} \in R^{m \times K}$ is the abundance matrix each column of which corresponds with the abundance fractions of a mixed pixel and $\mathbf{N} \in R^{L \times K}$ is the matrix of error term. With an appropriate Lagrange multiplier, the optimization problem of sparse unmixing for the whole hyperspectral image can be written as follows

$$\min_{\mathbf{X}} \frac{1}{2}\|\mathbf{A}\mathbf{X} - \mathbf{Y}\|_F^2 + \lambda \sum_{i=1}^{K} \|\mathbf{x}_i\|_1$$

$$\text{subject to: } \mathbf{X} \geq 0 \tag{8}$$

where $\|\cdot\|_F$ is the Frobenius norm of matrix, $\mathbf{x}_i$ is the i-th column of $\mathbf{X}$, $\lambda \geq 0$ is the regularization parameter that controls the sparsity of the solution. The sparse unmixing by variable splitting and augmented Lagrangian (SUnSAL) [23], which is based on the alternating direction method of multipliers (ADMM) [37, 38], has been proposed to solve the problem in Eq. (8).

Recently, a collaborative sparse unmixing model [31] has been proposed to exploit the subspace nature of the hyperspectral data. Specifically, it assumes that all the pixels in the hyperspectral image share the same active set of endmembers, which means the abundance matrix $\mathbf{X}$ should have only a few nonzero rows. Then the optimization problem of the collaborative sparse unmixing model can be written as

$$\min_{\mathbf{X}} \|\mathbf{X}\|_{\text{row-0}} \quad \text{subject to} \quad \|\mathbf{Y} - \mathbf{A}\mathbf{X}\|_F \leq \delta, \mathbf{X} \geq 0 \tag{9}$$

where $\|\mathbf{X}\|_{\text{row-0}}$ (called row-l_0 quasi-norm [44]) is the number of nonzero rows in matrix $\mathbf{X}$. The problem in Eq. (9) is also NP-hard. In the field of compressive sensing, row-l_0 quasi-norm is usually relaxed to the $l_{2,1}$ mixed norm [45], i.e. $\sum_{i=1}^{m} \|\mathbf{x}^i\|_2$. Thus, with an appropriate Lagrange multiplier, the collaborative sparse unmixing model becomes

$$\begin{aligned} &\min_{\mathbf{X}} \frac{1}{2}\|\mathbf{A}\mathbf{X} - \mathbf{Y}\|_F^2 + \lambda \sum_{i=1}^{m} \|\mathbf{x}^i\|_2 \\ &\text{subject to: } \mathbf{X} \geq 0 \end{aligned} \tag{10}$$

where $\mathbf{x}^i$ is the i-th row of the abundance matrix $\mathbf{X}$, $\lambda \geq 0$ is the regularization parameter that controls the row-sparse level of $\mathbf{X}$. In [31], the collaborative SUnSAL (CLSUnSAL), which is also based on the ADMM, has been proposed to solve the problem in Eq. (10).

The collaborative sparse unmixing model in Eq. (10) has several advantages over the model in Eq. (8): 1) as researched in the literature of multiple-measurement vectors (MMV) [46, 47], the collaborative sparse regression model is more likely to have a unique solution and succeed to get the solution; 2) the constraint that all the pixels in the hyperspectral data share the same set of endmembers can alleviate the sparse regression limitations caused by the usually high mutual coherence of the spectral library [31].

III. Sparse Unmixing Using Spectral a Priori Information

In this section, we first introduce the newly formulated sparse unmixing model that can exploit the spectral a priori information. Then, the SUnSPI algorithm will be presented based on the ADMM.

A. Sparse Unmixing Model Using Spectral a Priori Information

In the SUnSPI model, we assume that some materials in the spectral library are known to exist in the hyperspectral scene. Suppose $\mathbf{S} = \{1, \ldots, m\}$ is the set of the indices of all the spectral signatures in the spectral library; $\mathbf{P} \subset \mathbf{S}$ is the set of the indices corresponding to the materials that are known to exist in the scene; $\mathbf{S}/\mathbf{P} = \{x \in \mathbf{S} | x \notin \mathbf{P}\}$ denotes the set difference of $\mathbf{S}$ and $\mathbf{P}$. Let $\mathbf{X}^{\mathbf{P}}$ denote the rows of $\mathbf{X}$ that have the indices from the set $\mathbf{P}$. To incorporate the spectral a priori information into the sparse unmixing model, we should guarantee that the spectral

signatures indexed by $\mathbf{P}$ are active in representing the hyperspectral data; the other spectral signatures may or may not be active, but only a few of them should be active because of the subspace nature of the hyperspectral data. This line of reasoning leads us to consider enforcing the row-sparsity of $\mathbf{X}^{\mathbf{S/P}}$ while leaving $\mathbf{X}^{\mathbf{P}}$ alone. Then, we propose to solve the following optimization problem:

$$\min_{\mathbf{X}} \frac{1}{2}\|\mathbf{AX}-\mathbf{Y}\|_F^2 + \lambda_S \sum_{i=1}^{K} \|\mathbf{x}_i\|_1 + \lambda_P \sum_{i\in\mathbf{S/P}} \|\mathbf{x}^i\|_2$$
$$\text{subject to: } \mathbf{X} \geq 0 \tag{11}$$

where $\lambda_S \geq 0$ and $\lambda_P \geq 0$ are regularization parameters.

The first regularizer (called l_1 norm regularizer) in the SUnSPI model in Eq. (11) is the same with that in the SUnSAL model in Eq. (8), which enforces each column of $\mathbf{X}$ to be sparse. The second regularizer (called SPI regularizer) in the SUnSPI model is proposed to enforce the row-sparsity of $\mathbf{X}^{\mathbf{S/P}}$ which can incorporate the spectral a priori information indirectly. Fig. 1 illustrates the effect of the SPI regularizer in the SUnSPI model.

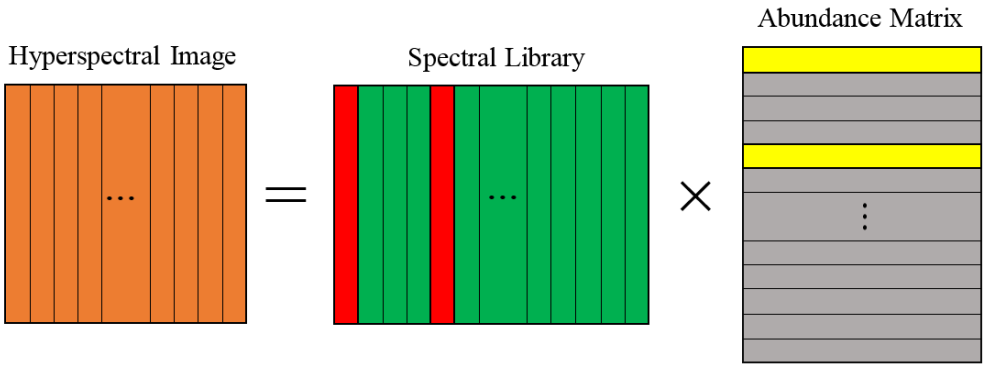

Fig. 1. Illustration of the effect of the SPI regularizer. The red columns in the spectral library matrix correspond with the materials that are known to exist in the scene. Due to the constraint of the SPI regularizer, the gray lines in the abundance matrix should be row-sparse, while the yellow lines which correspond with the known materials will not be row-sparse.

The SUnSPI model is much related to the SUnSAL model. Indeed, the SUnSPI model can be considered as the combination of the SUnSAL model with the SPI regularizer. When $\lambda_P = 0$, SUnSPI becomes the SUnSAL model. Another interpretation is that we first get the following objective function

$$\min_{\mathbf{X}} \frac{1}{2}\|\mathbf{AX}-\mathbf{Y}\|_F^2 + \lambda_P \sum_{i\in\mathbf{S/P}} \|\mathbf{x}^i\|_2$$
$$\text{subject to: } \mathbf{X} \geq 0 \tag{12}$$

which considers the spectral a priori information. By solving the optimization problem in Eq. (12), we can get the active endmembers, which correspond with the non-zero rows of $\mathbf{X}$. However, we do not get the optimal sparse solution without considering that each column of $\mathbf{X}$ is also sparse. For this purpose, we need to add the l_1 norm

regularizer into the objective function in Eq. (12), which leads to the SUnSPI model. In this way SUnSPI model can incorporate the spectral a priori information, while making sure that each column of $\mathbf{X}$ is also sparse.

The SUnSPI model is also related to the CLSUnSAL model in Eq. (10). When $\lambda_S = 0$ and no spectral a priori information is available (i.e. $\mathbf{P} = \emptyset$), the SUnSPI model becomes the CLSUnSAL model. It is worth noting that the CLSUnSAL model does not add the l_1 norm regularizer to enforce the sparsity of each column of the abundance matrix. This is because the $l_{2,1}$ norm regularizer already imposes such sparsity of each column of $\mathbf{X}$ and the l_1 norm regularizer somehow becomes redundant [31]. By contrast, the $l_{2,1}$ norm regularizer only acts on the $\mathbf{X}^{\mathbf{S/P}}$ in the SUnSPI model. Thus, the l_1 norm regularizer becomes necessary to impact the abundances with respect to all the actual endmembers in the SUnSPI model to improve the unmixing results further. From our experience, the SPI regularizer is much more dominant than the l_1 norm regularizer to improve the unmixing results. But the l_1 norm regularizer still plays a role in making the abundances be estimated even more accurately. Also, a disadvantage brought about by this extra regularizer is that SUnSPI uses two parameters while both SUnSAL and CLSUnSAL use only one.

Another close form of prior, also known as the subspace nature of hyperspectral data, is the knowledge that, given a spectral library and a hyperspectral dataset, the active endmembers in the dataset are only a few spectral signatures present in the library. And these signatures can be determined by array processing methodologies such as MUSIC-CSR [33] or simultaneous greedy algorithms [29, 32]. However, this kind of prior knowledge is different with the spectral a priori information. Spectral a priori information is a subset of all the actual endmembers while the spectral signatures obtained by MUSIC-CSR and simultaneous greedy algorithms are supposed to include all the actual endmembers. In other words, spectral a priori information tells that part of the spectral library are surely present in the hyperspectral data; spectral signatures determined by the subspace nature of hyperspectral data give us a smaller library to replace the original large library for sparse unmixing. Thus, in practical applications, both of these two kinds of prior knowledge can be incorporated into the sparse unmixing framework at the same time, i.e. spectral a priori information is used for sparse unmixing with a smaller spectral library.

B. The SUnSPI Algorithm

Now we present the SUnSPI algorithm to solve the problem in Eq. (11). The algorithm is based on the ADMM which has been proved to be very effective to solve the sparse unmixing problem [23, 27, 31]. For convenience, we use $\|\mathbf{X}\|_1$ and $\|\mathbf{X}\|_{2,1}$ to denote $\sum_{i=1}^{K} \|\mathbf{x}_i\|_1$ and $\sum_{i=1}^{m} \|\mathbf{x}^i\|_2$, respectively. Suppose $\mathbf{H} \in R^{m\times m}$ is a diagonal matrix related to the set $\mathbf{P}$:

$$h_{ii} = \begin{cases} 0, \text{ if } i \in \mathbf{P} \\ 1, \text{ otherwise} \end{cases} \tag{13}$$

where h_{ii} is the i-th diagonal element of $\mathbf{H}$. Thus, we have $\sum_{i\in\mathbf{S/P}} \|\mathbf{x}^i\|_2 = \|\mathbf{HX}\|_{2,1}$. Then the SUnSPI model in Eq. (11) can be written in the following equivalent form:

$$\min_{\mathbf{X}} \frac{1}{2}\|\mathbf{AX} - \mathbf{Y}\|_F^2 + \lambda_S\|\mathbf{X}\|_1 + \lambda_P\|\mathbf{HX}\|_{2,1} + l_{R+}(\mathbf{X}) \tag{14}$$

where $l_{R+}(\mathbf{X})$ is the indicator function: $l_{R+}(\mathbf{X})$ is zero if $\mathbf{X} \geq 0$ is satisfied and $+\infty$ otherwise.

The optimization problem in Eq. (14) has the following equivalent formulation:

$$\begin{aligned} \min_{\mathbf{U},\mathbf{V}} \frac{1}{2}\|\mathbf{V}_1 - \mathbf{Y}\|_F^2 + \lambda_S\|\mathbf{V}_2\|_1 + \lambda_P\|\mathbf{V}_3\|_{2,1} + l_{R+}(\mathbf{V}_4) \\ \text{subject to: } \mathbf{V}_1 = \mathbf{A}\mathbf{U} \\ \mathbf{V}_2 = \mathbf{U} \\ \mathbf{V}_3 = \mathbf{H}\mathbf{U} \\ \mathbf{V}_4 = \mathbf{U} \end{aligned} \tag{15}$$

where

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}_1 \\ \mathbf{V}_2 \\ \mathbf{V}_3 \\ \mathbf{V}_4 \end{bmatrix}. \tag{16}$$

Suppose $\mathbf{I}$ is identity matrix with proper size. We can write Eq. (15) in a more compact form:

$$\begin{aligned} &\min_{\mathbf{U},\mathbf{V}} g(\mathbf{V}) \\ &\text{subject to:} \quad \mathbf{G}\mathbf{U} + \mathbf{B}\mathbf{V} = 0 \end{aligned} \tag{17}$$

where

$$\begin{aligned} g(\mathbf{V}) \equiv &\frac{1}{2}\|\mathbf{V}_1 - \mathbf{Y}\|_F^2 + \lambda_S\|\mathbf{V}_2\|_1 + \lambda_P\|\mathbf{V}_3\|_{2,1} + l_{R+}(\mathbf{V}_4) \\ \mathbf{G} = \begin{bmatrix} \mathbf{A} \\ \mathbf{I} \\ \mathbf{H} \\ \mathbf{I} \end{bmatrix}, &\mathbf{B} = \begin{bmatrix} -\mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & -\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & -\mathbf{I} \end{bmatrix}. \end{aligned} \tag{18}$$

The augmented Lagrangian for the problem in Eq. (17) is

$$\ell(\mathbf{U}, \mathbf{V}, \mathbf{D}) \equiv g(\mathbf{V}) + \frac{\mu}{2}\|\mathbf{G}\mathbf{U} + \mathbf{B}\mathbf{V} - \mathbf{D}\|_F^2 \tag{19}$$

where $\mu > 0$ is the augmented Lagrangian penalty parameter [48], and $\mu\mathbf{D}$ denotes the Lagrange multipliers related to the constraint $\mathbf{G}\mathbf{U} + \mathbf{B}\mathbf{V} = 0$. In each iteration, the ADMM algorithm sequentially minimizes ℓ with respect to $\mathbf{U}$ and $\mathbf{V}$, and then updates the Lagrange multipliers. Following similar deduction steps in many related works [23, 31, 38], we can get the detailed process for the SUnSPI easily which is shown in Algorithm 1. The detailed deductions of the SUnSPI algorithm can be found in the supplementary material at http://levir.buaa.edu.cn/publications.html.

The convergence of the SUnSPI algorithm can be analyzed in a similar way with [27] and [31]. Since matrix $\mathbf{G}$ is of full column rank and function g in Eq. (18) is closed, proper, and convex, the convergence conditions for the ADMM algorithm in [49] are satisfied. According to the convergence theorem [49], under these conditions, for any

Algorithm 1 Pseudocode of the SUnSPI algorithm

Initialization:

set $k=0$, choose $\mu>0$, $\mathbf{U}^0, \mathbf{V}_1^0, \mathbf{V}_2^0, \mathbf{V}_3^0, \mathbf{V}_4^0, \mathbf{D}_1^0, \mathbf{D}_2^0, \mathbf{D}_3^0, \mathbf{D}_4^0$

repeat:

$\mathbf{U}^{(k+1)} \leftarrow (\mathbf{A}^T\mathbf{A}+2\mathbf{I}+\mathbf{H})^{-1}[\mathbf{A}^T(\mathbf{V}_1^{(k)}+\mathbf{D}_1^{(k)}) + \mathbf{V}_2^{(k)}+\mathbf{D}_2^{(k)}+\mathbf{H}(\mathbf{V}_3^{(k)}+\mathbf{D}_3^{(k)})+\mathbf{V}_4^{(k)}+\mathbf{D}_4^{(k)}]$

$\mathbf{V}_1^{(k+1)} \leftarrow \frac{1}{1+\mu}[\mathbf{Y}+\mu(\mathbf{A}\mathbf{U}^{(k)}-\mathbf{D}_1^{(k)})]$

$\mathbf{V}_2^{(k+1)} \leftarrow \text{soft}(\mathbf{U}^{(k)}-\mathbf{D}_2^{(k)}, \frac{\lambda_S}{\mu})$;
{soft$(\cdot,\tau)$ denotes the component-wise application of the soft-threshold function $y \mapsto \text{sign}(y)\max\{|y|-\tau,0\}$.}

$(\mathbf{V}_3^{(k+1)})^r \leftarrow \text{vect-soft}(\mathbf{H}\mathbf{U}^{(k)}-\mathbf{D}_3^{(k)}, \frac{\lambda_P}{\mu})$;
{vect-soft$(\cdot,\tau)$ denotes the row-wise application of the vect-soft-threshold function $\mathbf{y} \mapsto \mathbf{y}\max\{\|\mathbf{y}\|_2-\tau,0\}/(\max\{\|\mathbf{y}\|_2-\tau,0\}+\tau)$.}

$\mathbf{V}_4^{(k+1)} \leftarrow \max\{\mathbf{U}^{(k)}-\mathbf{D}_4^{(k)},0\}$

$\mathbf{D}_1^{(k+1)} \leftarrow \mathbf{D}_1^{(k)}-\mathbf{A}\mathbf{U}^{(k+1)}+\mathbf{V}_1^{(k+1)}$

$\mathbf{D}_2^{(k+1)} \leftarrow \mathbf{D}_2^{(k)}-\mathbf{U}^{(k+1)}+\mathbf{V}_2^{(k+1)}$

$\mathbf{D}_3^{(k+1)} \leftarrow \mathbf{D}_3^{(k)}-\mathbf{H}\mathbf{U}^{(k+1)}+\mathbf{V}_3^{(k+1)}$

$\mathbf{D}_4^{(k+1)} \leftarrow \mathbf{D}_4^{(k)}-\mathbf{U}^{(k+1)}+\mathbf{V}_4^{(k+1)}$

Update iteration: $k \leftarrow k+1$

until some stopping criterion is satisfied.

$\mu>0$, if the problem in Eq. (18) has a solution $\mathbf{U}^*$, then the sequence $\{\mathbf{U}^{(k)}\}$ will converge to $\mathbf{U}^*$; otherwise, at least one of the sequences $\{\mathbf{U}^{(k)}\}$ and $\{\mathbf{D}^{(k)}\}$ will diverge. In practical implementations, the ADMM algorithm stops when the maximum iteration number is reached or

$$\|\mathbf{G}\mathbf{U}^{(k)}+\mathbf{B}\mathbf{V}^{(k)}\|_F \le \epsilon \tag{20}$$

where $\epsilon=\sqrt{(3m+L)K}\epsilon'$ is the predefined error tolerance and ϵ' is the scaled error tolerance unrelated to the image size and library size [23, 27].

The setting of the augmented Lagrangian penalty parameter μ influences the convergence speed of the ADMM algorithm much. Here, we adopt the same update strategy with [31] (see also [50–53]). Specifically, μ is updated by keeping the ratio between the ADMM primal residual norm and dual residual norm within a given positive interval, as they both converge to zero. The primal residual ($r^{(k)}$) and dual residual ($d^{(k)}$) which can measure how well the iterates of Algorithm 1 satisfy the KKT conditions can be defined as [53]:

$$r^{(k)}=\mathbf{G}\mathbf{U}^{(k)}+\mathbf{B}\mathbf{V}^{(k)} \tag{21}$$

$$d^{(k)}=\mu\mathbf{G}^T\mathbf{B}(\mathbf{V}^{(k)}-\mathbf{V}^{(k-1)}). \tag{22}$$

The detailed deductions of the primal and dual residuals can be found in the supplementary material at http://levir.buaa.edu.cn/publicatio

IV. EXPERIMENTS

In this section, we conduct three synthetic data experiments and one real data experiment to test the validity of the proposed model. In the first synthetic data experiment, the SUnSPI algorithm[2] is compared with the SUnSAL [23] and CLSUnSAL [31] to test whether the spectral a priori information is beneficial to the sparse unmixing and whether our proposed SUnSPI algorithm can exploit this kind of information effectively. In the second synthetic data experiment, the SUnSPI is tested using different spectral a priori information when the correlations between the prior endmembers and the spectral library vary. In the third synthetic data experiment, the SUnSPI is tested when the total fractions corresponding to all the prior endmembers in the hyperspectral image are set differently. The supplementary material, which can be found at http://levir.buaa.edu.cn/publications.html, contains another synthetic data experiment for visual comparison of the performances of the different algorithms. All the considered algorithms have been tuned to their best performances by using different values of regularization parameters: 0, 0.001, 0.005, 0.01, 0.05, 0.1, 0.5, 1, 3, 5. The stopping criterions of SUnSAL and CLSUnSAL are set to be the same as that of SUnSPI: the scaled error tolerance ϵ' and maximum iteration number are respectively set to 10^{-4} and 300 which are enough for the algorithms to converge.

The root mean square error (RMSE) is used to evaluate the abundance estimations. For the i-th endmember, RMSE is defined as

$$\text{RMSE}_i = \sqrt{\frac{1}{K}\sum_{j=1}^{K}(\mathbf{X}_{ij} - \hat{\mathbf{X}}_{ij})^2}. \tag{23}$$

Here, $\mathbf{X}$ denotes the true abundances and $\hat{\mathbf{X}}$ represents the estimated ones. The mean value of all the endmembers' RMSEs will be computed. Generally speaking, the smaller the RMSE is, the more the estimation approximates the truth.

The remainder of this section is organized as follows. The first three subsections will respectively present the three synthetic data experiments. Real data experiment is conducted in Section IV-D. In Section IV-E we discuss the parameters setting of the SUnSPI algorithm. Finally, we analyse the computational complexities of the considered algorithms and compare their running time in Section IV-F.

A. Synthetic Data Experiment 1

For our first synthetic data experiment, we evaluate the performances of the sparse unmixing algorithms in situations of different noise types, different signal-to-noise ratios (SNR $\equiv 10\log_{10}\frac{\|\mathbf{Ax}\|_2^2}{\|\mathbf{n}\|_2^2}$) of noise, different spectral libraries and different endmember numbers. Specifically, the synthetic data sets are contaminated by Gaussian white noise and correlated noise[3] with different levels of SNR: 20, 30 and 40 dB. Two spectral libraries are considered in our experiments:

[2]The MATLAB code of SUnSPI is available at http://levir.buaa.edu.cn/publications.html.

[3]The Gaussian white noise is generated using the *awgn* function in MATLAB. The correlated noise is generated using the *correlatedGaussianNoise* function that is available online: http://www.mathworks.com/matlabcentral/fileexchange/21156-correlated-gaussian-noise/content/correlatedGaussianNoise.m. The correlation matrix is set as default.

- $\mathbf{A}_1 \in R^{224\times 498}$: The Chapter 1 of United States Geological Survey (USGS) [54] digital spectral library (splib06a). The reflectance values of 498 materials are measured for 224 spectral bands distributed uniformly in the interval $0.4 - 2.5$ μm.
- $\mathbf{A}_2 \in R^{224\times 240}$: Subset of $\mathbf{A}_1$, where 240 spectral signatures are selected randomly from $\mathbf{A}_1$.

Besides, we also test the performance of SUnSPI when different amounts of endmembers are known to exist in the hyperspectral scene: SUnSPI(0) means no spectral a priori information is available; SUnSPI(1) means $\frac{1}{3}$ of the actual endmembers are known to exist in the imagery; SUnSPI(2) means $\frac{2}{3}$ of the actual endmembers are known to exist. For example, suppose there are totally 6 actual endmembers in the hyperspectral scene, SUnSPI(1) knows 2 of them are active in the spectral library and SUnSPI(2) knows 4 of them are active in the spectral library. The 4 endmembers SUnSPI(2) knows will include the 2 endmembers SUnSPI(1) knows and are selected randomly. In this way, we can see whether more spectral a priori information will lead to better unmixing results. Note though SUnSPI(0) gets no spectral a priori information, it is different with the SUnSAL and CLSUnSAL. Indeed, SUnSPI(0) can be considered as a direct combination of these two algorithms. Moreover, to show how the l_1 norm regularizer influences the results obtained by SUnSPI, we also consider solving the problem in Eq. (12) which we call nonnegative constrained least squares with spectral a priori information (NCLS-SPI). Note when the $\lambda_S = 0$, SUnSPI becomes NCLS-SPI. When no priori information is available, NCLS-SPI is essentially CLSUnSAL. NCLS-SPI(1) and NCLS-SPI(2) have the same spectral a priori information with SUnSPI(1) and SUnSPI(2), respectively.

Nine spectral signatures are randomly chosen from $\mathbf{A}_2$ to generate the synthetic hyperspectral images: Rhodochrosite HS67, Axinite HS342.3B, Chrysocolla HS297.3B, Niter GDS43 (K-Saltpeter), Anthophyllite HS286.3B, Neodymium Oxide GDS34, Monazite HS255.3B, Samarium Oxide GDS36, Pigeonite HS199.3B. Finally, the synthetic hyperspectral data sets are generated using different endmember numbers: $k_1 = 3$, $k_2 = 6$, $k_3 = 9$.

The first synthetic hyperspectral data set (SD1) contains 900 pixels randomly generated following a Dirichlet distribution [8, 18]. To make sure that no pure pixel exists in the hyperspectral data, we force all the abundances to be no larger than 0.7.

Tab. I shows the RMSEs obtained by different algorithms using $\mathbf{A}_1$ on SD1 corrupted by white noise. The performances of all the algorithms degrade as the noise gets stronger. We can see that in most cases, CLSUnSAL behaves better than SUnSAL, which indicates that it is important to incorporate the subspace nature of the hyperspectral data into sparse unmixing. Besides, as expected, SUnSPI(0) outperforms SUnSAL and CLSUnSAL in all the cases because SUnSPI(0) not only considers the subspace nature of hyperspectral data but also gets the improved sparse solution for each pixel. However, the advantage of SUnSPI(0) over CLSUnSAL is not large, which is in accordance with the results in [31]. This is because the $l_{2,1}$ norm already imposes sparsity on the solution. Now, we consider the spectral a priori information in sparse unmixing. From Tab. I, we can find that in all the cases SUnSPI(1) and SUnSPI(2) behave better than SUnSPI(0), which suggests that the spectral a priori information is beneficial to sparse unmixing and the proposed model can exploit this information to improve the unmixing results. We can also observe that SUnSPI(2) always has a better performance than SUnSPI(1), which indicates that more spectral a priori information will lead to better results. Comparing SUnSPI and NCLS-SPI, we find that when the

TABLE I

RMSEs OBTAINED BY DIFFERENT ALGORITHMS USING $\mathbf{A}_1$ ON SD1 CORRUPTED BY WHITE NOISE

	SNR (dB)	SUnSAL	CLSUnSAL	NCLS-SPI(1)	NCLS-SPI(2)	SUnSPI(0)	SUnSPI(1)	SUnSPI(2)
SD1	20	0.0682	0.0395	0.0232	0.0216	0.0369	0.0231	0.0216
($k_1 = 3$)	30	0.0223	0.0169	0.0092	0.0088	0.0156	0.0092	0.0086
	40	0.0071	0.0061	0.0036	0.0036	0.0059	0.0036	0.0035
SD1	20	0.0647	0.0602	0.0595	0.0447	0.0589	0.0581	0.0403
($k_2 = 6$)	30	0.0227	0.0223	0.0218	0.0155	0.0214	0.0209	0.0143
	40	0.0078	0.0092	0.0081	0.0055	0.0078	0.0076	0.0044
SD1	20	0.0674	0.0663	0.0636	0.0619	0.0639	0.0595	0.0573
($k_3 = 9$)	30	0.0399	0.0401	0.0386	0.0370	0.0395	0.0363	0.0358
	40	0.0304	0.0300	0.0297	0.0249	0.0297	0.0286	0.0249

endmember number is 3, the RMSEs obtained by them are nearly the same. But, when the endmember number and noise increase, the improvement caused by the l_1 regularizer becomes more obvious. This phenomenon suggests that when there are only a few endmembers in the scene or the noise of the hyperspectral data is very weak, λ_S in the SUnSPI model can be simply set to 0 to achieve satisfactory results and avoid tuning an extra parameter. But when the situation is more complex (i.e. high endmember number and noise), the l_1 norm regularizer can make the estimation more accurate.

Tab. II shows the RMSEs obtained by different algorithms using $\mathbf{A}_1$ on SD1 corrupted by correlated noise. Since the noise in the real hyperspectral images is usually correlated, this case is closer to the practical ones. From Tab. II we can see that all the algorithms tend to get worse performances in the correlated noise case compared with in the white noise case. In nearly half of the cases, SUnSAL behaves better than CLSUnSAL. This phenomenon suggests that the correlated noise could have a more negative effect on the CLSUnSAL than the SUnSAL. Thus, the combination of these two algorithms is necessary to get the improved sparse solution. We can observe that SUnSPI(2) and SUnSPI(1) outperform the other algorithms, which indicates that the spectral a priori information can help the sparse unmixing algorithms mitigate the difficulty brought by the correlated noise.

Tab. III and Tab. IV show the RMSEs obtained by different algorithms using $\mathbf{A}_2$ on SD1 corrupted by white noise and correlated noise, respectively. All the algorithms behave better when using $\mathbf{A}_2$ than using $\mathbf{A}_1$. This is because the mutual coherence of $\mathbf{A}_2$ is much weaker than that of $\mathbf{A}_1$ as $\mathbf{A}_2$ is a subset of $\mathbf{A}_1$. Similarly, we can find that the spectral a priori information benefits the sparse unmixing algorithms. We can also observe that when the spectral library becomes smaller, the improvement brought by the l_1 regularizer also declines.

B. Synthetic Data Experiment 2

From Tabs. I - IV, we can observe that the improvement of abundance estimation brought by the spectral a priori information is not always large. This is because the prior endmembers in the first synthetic data experiments are selected randomly and some prior endmembers are not so related to the other spectra in the library, which

TABLE II

RMSEs OBTAINED BY DIFFERENT ALGORITHMS USING $\mathbf{A}_1$ ON SD1 CORRUPTED BY CORRELATED NOISE

	SNR (dB)	SUnSAL	CLSUnSAL	NCLS-SPI(1)	NCLS-SPI(2)	SUnSPI(0)	SUnSPI(1)	SUnSPI(2)
SD1	20	0.0811	0.0623	0.0351	0.0211	0.0598	0.0351	0.0211
($k_1 = 3$)	30	0.0268	0.0268	0.0239	0.0212	0.0268	0.0165	0.0138
	40	0.0084	0.0096	0.0086	0.0075	0.0083	0.0047	0.0040
SD1	20	0.0694	0.0734	0.0706	0.0568	0.0672	0.0663	0.0469
($k_2 = 6$)	30	0.0242	0.0271	0.0255	0.0183	0.0232	0.0227	0.0160
	40	0.0088	0.0090	0.0088	0.0075	0.0088	0.0087	0.0045
SD1	20	0.0680	0.0720	0.0668	0.0657	0.0653	0.0609	0.0585
($k_3 = 9$)	30	0.0408	0.0434	0.0409	0.0380	0.0407	0.0374	0.0369
	40	0.0304	0.0305	0.0298	0.0250	0.0297	0.0286	0.0250

TABLE III

RMSEs OBTAINED BY DIFFERENT ALGORITHMS USING $\mathbf{A}_2$ ON SD1 CORRUPTED BY WHITE NOISE

	SNR (dB)	SUnSAL	CLSUnSAL	NCLS-SPI(1)	NCLS-SPI(2)	SUnSPI(0)	SUnSPI(1)	SUnSPI(2)
SD1	20	0.0460	0.0279	0.0203	0.0192	0.0279	0.0203	0.0186
($k_1 = 3$)	30	0.0157	0.0094	0.0063	0.0057	0.0092	0.0063	0.0057
	40	0.0058	0.0035	0.0023	0.0022	0.0031	0.0023	0.0021
SD1	20	0.0526	0.0510	0.0469	0.0337	0.0447	0.0417	0.0323
($k_2 = 6$)	30	0.0183	0.0183	0.0176	0.0110	0.0158	0.0151	0.0110
	40	0.0066	0.0059	0.0050	0.0037	0.0049	0.0046	0.0036
SD1	20	0.0611	0.0620	0.0553	0.0466	0.0585	0.0528	0.0459
($k_3 = 9$)	30	0.0378	0.0380	0.0355	0.0178	0.0360	0.0344	0.0178
	40	0.0291	0.0292	0.0280	0.0055	0.0288	0.0280	0.0055

TABLE IV

RMSEs OBTAINED BY DIFFERENT ALGORITHMS USING $\mathbf{A}_2$ ON SD1 CORRUPTED BY CORRELATED NOISE

	SNR (dB)	SUnSAL	CLSUnSAL	NCLS-SPI(1)	NCLS-SPI(2)	SUnSPI(0)	SUnSPI(1)	SUnSPI(2)
SD1	20	0.0475	0.0344	0.0256	0.0172	0.0337	0.0256	0.0172
($k_1 = 3$)	30	0.0154	0.0115	0.0075	0.0057	0.0108	0.0072	0.0057
	40	0.0060	0.0047	0.0030	0.0025	0.0037	0.0027	0.0023
SD1	20	0.0535	0.0644	0.0554	0.0363	0.0474	0.0440	0.0323
($k_2 = 6$)	30	0.0174	0.0223	0.0202	0.0102	0.0156	0.0148	0.0101
	40	0.0071	0.0073	0.0060	0.0039	0.0053	0.0048	0.0035
SD1	20	0.0582	0.0697	0.0563	0.0476	0.0569	0.0512	0.0462
($k_3 = 9$)	30	0.0369	0.0402	0.0367	0.0178	0.0362	0.0349	0.0171
	40	0.0291	0.0297	0.0282	0.0054	0.0289	0.0277	0.0054

means they are relatively easy to identify for the sparse unmixing algorithms such as CLSUnSAL and SUnSAL. By contrast, if some prior endmembers are very related to some other library members, they will be difficult to distinguish and the improvement brought by the spectral a priori information will be considerable. Thus, in this experiment, to explore this phenomenon further, we test SUnSPI using different spectral a priori information when the correlations between the prior endmembers and the spectral library vary. Specifically, the correlation between the i-th library member (denoted by $\mathbf{a}_i$) and the spectral library can be defined as the maximum cosine between $\mathbf{a}_i$ and any other member in the library:

$$cor_i = \max_{j \in \mathbf{S}, j \neq i} \frac{\mathbf{a}_i^T \mathbf{a}_j}{\|\mathbf{a}_i\|_2 \|\mathbf{a}_j\|_2}. \tag{24}$$

The spectral library used in this experiment is $\mathbf{A}_1$. Using Eq. (24), the correlation between each spectral signature and the spectral library can be obtained easily. We manually choose 6 spectral signatures from $\mathbf{A}_1$ to generate the second synthetic hyperspectral data set (SD2). The 6 endmembers can be divided into 3 groups:

- Actinolite HS315.4B and Monazite HS255.3B (group 1): among the top 5% least correlated spectral signatures with the spectral library. The cors of these two signatures are 0.9812 and 0.9809, respectively.
- Nacrite GDS88 and Pyrope WS474 (group 2): among the top 5% most correlated spectral signatures with the spectral library. The cors of these two signatures are both 0.9990.
- Glauconite HS313.3B and Olivine HS285.4B (group 3): moderate correlated spectral signatures with the spectral library. The cors of these two signatures are 0.9894 and 0.9905, respectively.

The abundances are generated in the same way as the first synthetic data. The endmember number is 6. SUnSPI(0) means no spectral a priori information is available; SUnSPI(1S), SUnSPI(1L) and SUnSPI(1M) take group 1, group 2 and group 3 as spectral a priori information, respectively. As SUnSPI(0) is a generalization of SUnSAL and CLSUnSAL and behaves better than them (as can be seen in the first synthetic experiments), we do not consider SUnSAL and CLSUnSAL in this experiment. Tabs. V and VI display the results obtained by SUnSPI using different spectral a priori information on SD2 corrupted by white noise and correlated noise, respectively. We can observe that when other conditions are equal, the more correlated the prior endmembers are with the spectral library, the better the abundances will be estimated. Even when the prior endmembers are among the library members that are least correlated with the spectral library, the improvements of abundance estimation brought by them are obvious in most cases. When the prior endmembers are more correlated with the spectral library, the improvements become more considerable. Thus, as the high correlation between the endmembers and the spectral library is the greatest challenge faced by the sparse unmixing algorithms, it is necessary to incorporate the available spectral a priori information into them.

C. Synthetic Data Experiment 3

Another factor that could affect the abundance estimation is the total abundance fractions corresponding to the prior endmembers in the hyperspectral image. Intuitively, if these fractions are small, the improvement of abundance estimation brought by the spectral a priori information could also be small. To see how this factor can affect the

TABLE V

RMSEs OBTAINED BY SUnSPI USING $\mathbf{A}_1$ ON SD2 CORRUPTED BY WHITE NOISE WHEN THE CORRELATIONS BETWEEN THE PRIOR ENDMEMBERS AND THE SPECTRAL LIBRARY VARY

SNR (dB)	SUnSPI(0)	SUnSPI(1S)	SUnSPI(1M)	SUnSPI(1L)
20	0.1124	0.1063	0.0871	0.0779
30	0.0502	0.0460	0.0351	0.0340
40	0.0142	0.0115	0.0103	0.0090

TABLE VI

RMSEs OBTAINED BY SUnSPI USING $\mathbf{A}_1$ ON SD2 CORRUPTED BY CORRELATED NOISE WHEN THE CORRELATIONS BETWEEN THE PRIOR ENDMEMBERS AND THE SPECTRAL LIBRARY VARY

SNR (dB)	SUnSPI(0)	SUnSPI(1S)	SUnSPI(1M)	SUnSPI(1L)
20	0.1287	0.1263	0.1178	0.0801
30	0.0581	0.0549	0.0443	0.0318
40	0.0181	0.0163	0.0136	0.0085

abundance estimation, we test the SUnSPI algorithm on synthetic data sets with different total abundance fractions corresponding to the prior endmembers.

The spectral library used in this experiment is still $\mathbf{A}_1$ and 6 spectral signatures are selected randomly to construct the synthetic hyperspectral data (SD3). SUnSPI(0) means no spectral a priori information is available; SUnSPI(1) means 3 random endmembers are known to exist. The abundances of 900 pixels are randomly generated following a Dirichlet distribution. Then, the abundances are scaled to meet the following two conditions: 1) the percentage of the total abundance fractions corresponding to the prior endmembers in the hyperspectral image is a given constant, i.e. 5%, 10%, 20%, 50% or 80%; 2) the abundance fractions corresponding to all the endmembers satisfy the nonnegativity constraint and the sum-to-one constraint.

Tab. VII and Tab. VIII display the results obtained by SUnSPI using $\mathbf{A}_1$ on SD3 corrupted by white noise and correlated noise, respectively. Note that the hyperspectral data change as the percentage of the total fractions corresponding to the prior endmembers is set differently. Thus, to make the experimental results more comparable, we also report the improvement of abundance estimation brought by the spectral a priori information which is calculated as

$$\text{improvement} = \frac{\text{RMSE(0) - RMSE(1)}}{\text{RMSE(0)}} \tag{25}$$

where RMSE(0) and RMSE(1) denote the RMSEs obtained by SUnSPI(0) and SUnSPI(1), respectively. We can observe that as the percentage of the abundance fractions corresponding to the prior endmembers increases, the improvement of abundance estimation brought by the spectral a priori information also increases, but not always significantly. It is worth noting that even when this percentage is only 10%, the spectral a priori information can lead to considerable improvement to the abundance estimation, which demonstrates that it is necessary and critical

TABLE VII

RESULTS OBTAINED BY SUNSPI USING $\mathbf{A}_1$ ON SD3 CORRUPTED BY WHITE NOISE WHEN THE TOTAL FRACTIONS OF THE PRIOR ENDMEMBERS ARE SET DIFFERENTLY

SNR (dB)	percentage (prior)	SUnSPI(0)	SUnSPI(1)	improvement
20	80%	0.0522	0.0372	28.70%
	50%	0.0508	0.0365	28.20%
	20%	0.0446	0.0359	19.40%
	10%	0.0390	0.0337	13.70%
	5%	0.0310	0.0312	−0.57%
30	80%	0.0209	0.0138	34.00%
	50%	0.0207	0.0138	33.11%
	20%	0.0193	0.0133	30.89%
	10%	0.0175	0.0128	26.51%
	5%	0.0157	0.0125	20.39%
40	80%	0.0078	0.0055	30.11%
	50%	0.0073	0.0052	28.67%
	20%	0.0071	0.0051	28.57%
	10%	0.0068	0.0050	25.60%
	5%	0.0066	0.0050	24.40%

to exploit this a priori knowledge for sparse unmixing. But when the proportion of the prior endmembers is too small (i.e. lower than 5%), their abundances could be overwhelmed by the strong noise and thus the improvement cannot be guaranteed.

D. Real Data Experiment

In this subsection, we use the real data to test the validity of the proposed algorithm. Because the true abundance maps of the real hyperspectral data are unavailable, we use the classification maps of the hyperspectral scene to make a qualitative analysis. We display the abundances estimated by SUnSAL, CLSUnSAL and SUnSPI for visual comparison.

The real data set we used is the well-known AVIRIS Cuprite data set[4]. In our experiment, we use a 204×151 pixels subset with 188 spectral bands (low-SNR bands are removed). The spectral library used here is $\mathbf{A}_1$ which has been used in the synthetic data experiments but with the corresponding bands removed. Fig. 2 shows the minerals map[5] which was produced by a Tricorder 3.3 software product. Note that the publicly available AVIRIS Cuprite data were collected in 1997 but the Tricorder map was produced in 1995. Thus, we can only make a qualitative analysis of the performances of different sparse unmixing algorithms by comparing their estimated abundances with the minerals map.

[4]http://aviris.jpl.nasa.gov/html/aviris.freedata.html

[5]http://speclab.cr.usgs.gov/PAPERS/tetracorder/

TABLE VIII

RESULTS OBTAINED BY SUNSPI USING $\mathbf{A}_1$ ON SD3 CORRUPTED BY CORRELATED NOISE WHEN THE TOTAL FRACTIONS OF THE PRIOR ENDMEMBERS ARE SET DIFFERENTLY

SNR (dB)	percentage (prior)	SUnSPI(0)	SUnSPI(1)	improvement
20	80%	0.0653	0.0425	34.80%
	50%	0.0601	0.0413	31.36%
	20%	0.0554	0.0402	27.36%
	10%	0.0459	0.0386	15.76%
	5%	0.0381	0.0357	6.28%
30	80%	0.0249	0.0148	40.35%
	50%	0.0233	0.0147	36.95%
	20%	0.0220	0.0146	33.73%
	10%	0.0206	0.0143	30.25%
	5%	0.0183	0.0140	23.74%
40	80%	0.0086	0.0059	31.38%
	50%	0.0082	0.0056	31.10%
	20%	0.0081	0.0056	30.29%
	10%	0.0079	0.0056	29.24%
	5%	0.0077	0.0055	28.17%

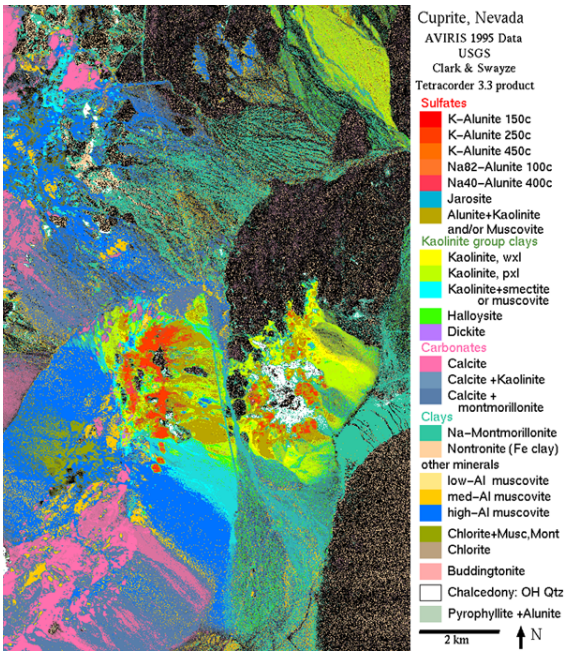

Fig. 2. USGS map showing the distribution of different minerals in the Cuprite mining district in Nevada.

There are about 14 kinds of mineral in the considered hyperspectral scene [8]. In order to test whether the proposed algorithm can exploit the spectral a priori information to improve the unmixing results in the real data, we assume that SUnSPI knows four minerals exist in the scene: Alunite, Buddingtonite, Chalcedony and Montmorillonite. The λ_P and λ_S in the SUnSPI algorithm are set empirically to 0.1 and 0.001, respectively. The λ's for SUnSAL and CLSUnSAL are set to 0.001 and 0.01, respectively (after empirical optimization). Fig. 3 shows a qualitative comparison among the fractional abundance maps of the four considered minerals in the hyperspectral scene estimated by the SUnSAL, CLSUnSAL and SUnSPI. Note that the abundance maps estimated by the sparse unmixing algorithms and the Tricorder maps are not exactly the same. This is because the Tricorder maps are classification maps which consider each pixel in the hyperspectral data pure and classify it as member of a specific class related to the representative mineral in the pixel; but unmixing is a subpixel level classification and the abundances for a mixed pixel represent the degree of presence of the mineral in the pixel. But, it can be observed in Fig. 3 that the highest abundances estimated by the sparse unmixing algorithms generally correspond with those pixels belong to the respective class of minerals. Thus, we can qualitatively conclude that SUnSPI is a valid tool for sparse unmixing of real hyperspectral data. It is also worth noting that the abundances estimated by SUnSPI are generally comparable or higher in the regions classified as respective minerals in comparison to SUnSAL and CLSUnSAL. This observation is more obvious for the abundances corresponding to Chalcedony and Montmorillonite than the abundances of Alunite and Buddingtonite. Generally, we can conclude that SUnSPI can exploit the spectral a priori information to improve the unmixing results in the real hyperspectral image.

E. Discussion of Parameters Setting

There are two regularization parameters to be set for the SUnSPI algorithm, i.e. the λ_P and λ_S. It should be clear that it is impossible to choose the optimal parameters for the convex relaxation methods directly. In all the synthetic experiments, the SUnSPI is tested using different values of λ_S and λ_P: 0, 0.001, 0.005, 0.01, 0.05, 0.1, 0.5, 1, 3, 5. All possible combinations of these parameters are considered. The optimal parameters are defined as parameter pairs that lead to the best performances of SUnSPI. Here, based on the experimental results and intuitive analysis, we summarize several rules concerned with how to choose relatively good parameters. The following rules are satisfied in most cases of our experiments. First, the SPI regularizer ($\sum_{i\in\mathbf{S}/\mathbf{P}} \|\mathbf{x}^i\|_2$) in Eq. (11) is more dominant than the l_1 norm regularizer ($\sum_{i=1}^{K} \|\mathbf{x}_i\|_1$). Indeed, for the optimal parameter pair (λ_P, λ_S), λ_P is usually much larger than λ_S. There are some cases in which the optimal λ_S is zero. This phenomenon indicates that the SPI regularizer plays a more important role in sparse unmixing. Second, as the noise gets weaker or SNR gets higher, both the optimal λ_P and optimal λ_S will decrease. This is natural because weaker noise means more fitness between $\mathbf{Y}$ and $\mathbf{AX}$. Thus, the term $\frac{1}{2}\|\mathbf{AX}-\mathbf{Y}\|_F^2$ in Eq. (11) will become more critical to the SUnSPI model. By our experience, when the noise is very weak or absent, even the nonnegative constrained least squares model ($\lambda_P = \lambda_S = 0$) can get very good estimation of the abundances. In these cases, larger λ_P or λ_S could have a negative effect on the results. Third, when the amount of spectral a priori information increases, the optimal λ_P stays unchanged or increases in our experiments. As the number of active endmembers in the spectral library $\mathbf{A}$ is fixed, when the set

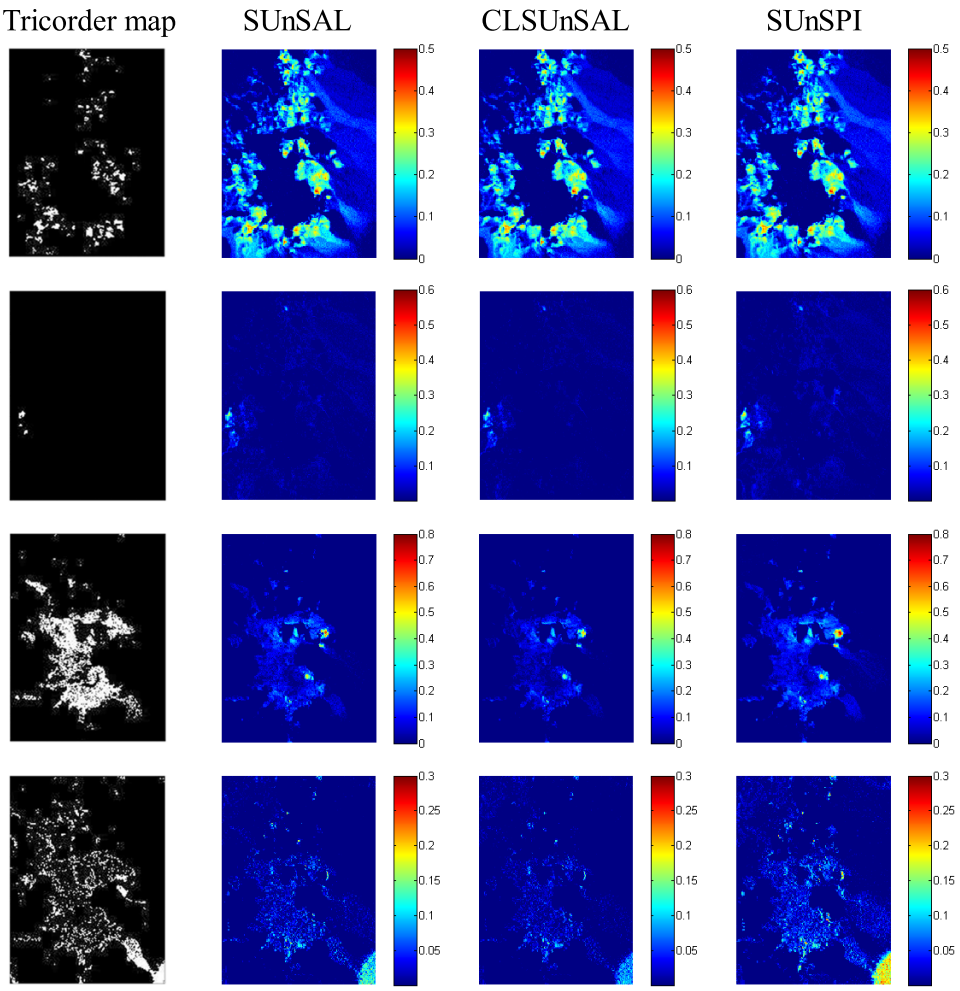

Fig. 3. The fractional abundance maps estimated by SUnSAL, CLSUnSAL, SUnSPI, and the distribution maps produced by Tricorder software for the 204×151 pixels subset of AVIRIS Cuprite scene. From left column to right column are the maps produced or estimated by Tricorder software, SUnSAL, CLSUnSAL and SUnSPI, respectively. From top row to bottom row are the maps corresponding to Alunite, Buddingtonite, Chalcedony and Montmorillonite, respectively.

$\mathbf{P}$ which corresponds with the spectral a priori information gets larger, the number of active endmembers in the sub-library $\mathbf{A}_{\mathbf{S}/\mathbf{P}}$ will decrease. Thus, $\mathbf{A}_{\mathbf{S}/\mathbf{P}}$ will be row-sparser, which will lead to larger optimal λ_P. But, as the parameter candidates in our experiments (i.e. 0, 0.001, 0.005, 0.01, 0.05, 0.1, 0.5, 1, 3, 5) are discontinuous samples with relatively large sampling intervals, the optimal λ_P in our experiments stays unchanged occasionally rather than increases monotonously when the amount of spectral a priori information increases. Fourth, when other conditions are equal, the optimal parameter pair (λ_P, λ_S) varies slightly when the hyperspectral data are corrupted by white noise or correlated noise. For example, Tab. IX and Tab. X display the optimal parameter pair (λ_P, λ_S) obtained by the SUnSPI algorithm using spectral library $\mathbf{A}_1$ on the SD1 with white noise and correlated noise, respectively. For illustrative purpose, Fig. 4 shows the RMSE obtained by SUnSPI as a function of parameters λ_P

TABLE IX

OPTIMAL PARAMETER PAIR (λ_P, λ_S) OBTAINED BY THE SUNSPI ALGORITHM USING SPECTRAL LIBRARY $\mathbf{A}_1$ ON THE SD1 WITH WHITE NOISE

	SNR (dB)	SUnSPI(0)	SUnSPI(1)	SUnSPI(2)
SD1	20	(1, 0.01)	(1, 0.001)	(3, 0)
($k_1 = 3$)	30	(0.5, 0.01)	(0.5, 0)	(0.5, 0.01)
	40	(0.1, 0.001)	(0.1, 0)	(0.1, 0.001)
SD1	20	(1, 0.05)	(1, 0.1)	(1, 0.05)
($k_2 = 6$)	30	(0.5, 0.01)	(0.5, 0.01)	(0.5, 0.01)
	40	(0.05, 0.005)	(0.1, 0.001)	(0.1, 0.001)
SD1	20	(0.1, 0.1)	(0.5, 0.05)	(0.5, 0.05)
($k_3 = 9$)	30	(0.01, 0.01)	(0.1, 0.01)	(0.1, 0.01)
	40	(0.01, 0.005)	(0.05, 0.005)	(0.05, 0)

TABLE X

OPTIMAL PARAMETER PAIR (λ_P, λ_S) OBTAINED BY THE SUNSPI ALGORITHM USING SPECTRAL LIBRARY $\mathbf{A}_1$ ON THE SD1 WITH CORRELATED NOISE

	SNR	SUnSPI(0)	SUnSPI(1)	SUnSPI(2)
SD1	20	(0.1, 0.1)	(3, 0)	(5, 0)
($k_1 = 3$)	30	(0.01, 0.01)	(0.5, 0.05)	(1, 0.05)
	40	(0.01, 0.005)	(0.1, 0.005)	(0.1, 0.005)
SD1	20	(1, 0.1)	(1, 0.1)	(1, 0.1)
($k_2 = 6$)	30	(0.5, 0.05)	(0.5, 0.05)	(0.5, 0.01)
	40	(0.05, 0.005)	(0.05, 0.005)	(0.1, 0)
SD1	20	(0.1, 0.1)	(0.1, 0.1)	(1, 0.1)
($k_3 = 9$)	30	(0.05, 0.05)	(0.1, 0.1)	(0.5, 0.05)
	40	(0.01, 0.005)	(0.01, 0.005)	(0.1, 0.01)

and λ_S on SD1 using $\mathbf{A}_1$ when the endmember number is 6 and 2 of the endmembers are known to exist. We can observe that as the noise becomes stronger, the regularization parameters increase and the improvement brought by the spectral a priori information is more obvious. These observations are in accordance with our analyses above.

F. Complexity Analysis

In the following analysis, suppose L and K are the band number and pixel number in the hyperspectral image, respectively; m is the number of spectral signatures in the spectral library. The most expensive step in the SUnSPI algorithm is the calculation of $\mathbf{U}$ in Step 4 of Algorithm 1. To update $\mathbf{U}$, we need to solve a linear system of equations of size $m \times m$. However, the matrix involved in this system of equations (i.e. $\mathbf{A}^T\mathbf{A} + 2\mathbf{I} + \mathbf{H}$) is fixed and then can be precomputed with low complexity. Then the complexity of updating $\mathbf{U}$ is dominated by the matrix multiplication in this step whose complexity is $O(mLK)$. Thus, the overall complexity of SUnSPI is $O(mLK)$ which is the same with those of SUnSAL and CLSUnSAL.

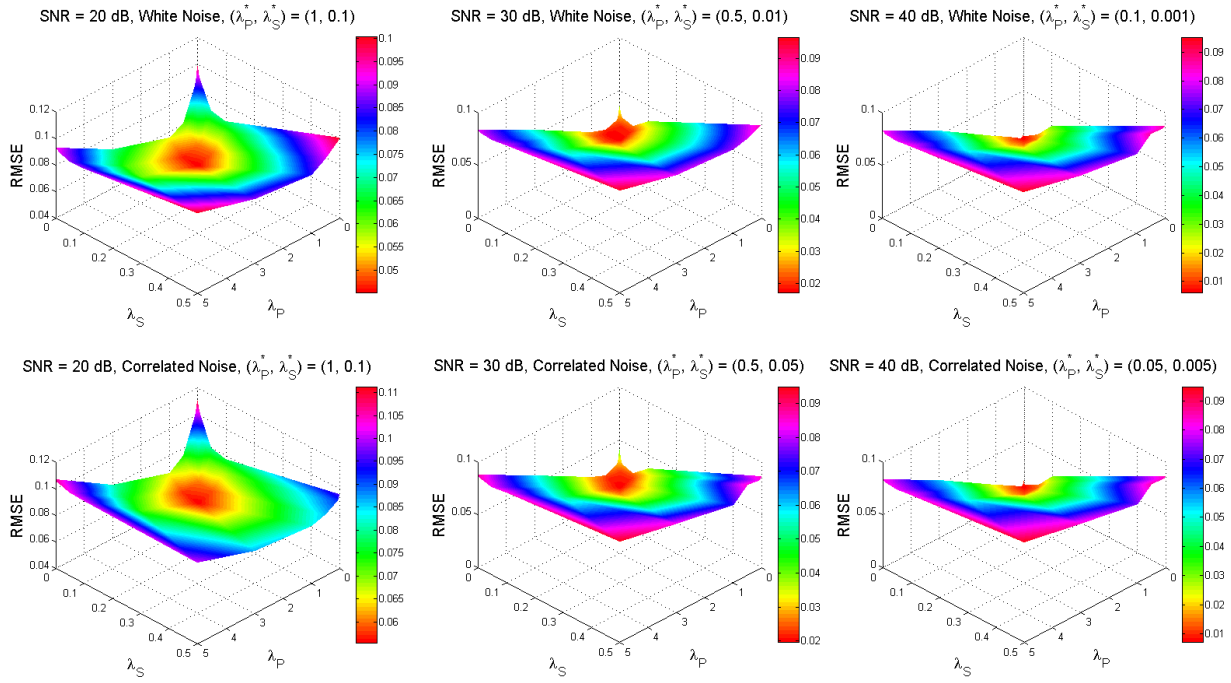

Fig. 4. RMSE obtained by SUnSPI as a function of parameters λ_P and λ_S on SD1 using $\mathbf{A}_1$ when the endmember number is 6 and 2 of the endmembers are known to exist. $(\lambda_P^*, \lambda_S^*)$ means the optimal parameter pair.

TABLE XI

PROCESSING TIMES AND NUMBERS OF ITERATIONS MEASURED AFTER APPLYING THE ALGORITHMS TO SD1 WITH 30 DB WHITE NOISE CONSTRUCTED BY 6 ENDMEMBERS USING $\mathbf{A}_1$

Algorithms	SUnSAL	CLSUnSAL	SUnSPI(0)	SUnSPI(1)	SUnSPI(2)
Time (s)	10.04	16.92	37.90	23.26	28.19
Iterations	160	140	150	90	110

Tab. XI reports the processing time and iteration numbers (the average values after running each algorithm 10 times) measured after applying the algorithms to SD1 with 30 dB white noise constructed by 6 endmembers using $\mathbf{A}_1$. A desktop PC equipped with an Intel Core 2 Duo CPU (2.66 GHz) and 4 GB of RAM memory was used to implement all the algorithms on MATLAB R2012b. It can be observed that SUnSPI could be slower than SUnSAL and CLSUnSAL due to its more complex model.

V. CONCLUSION

In this paper, we consider the spectral a priori information in sparse unmixing, which assumes that some materials are known to exist in the hyperspectral scene. Then we propose a novel sparse unmixing model to exploit the spectral a priori information. Then, based on the ADMM, we present a new algorithm termed sparse unmixing using spectral a priori information (SUnSPI) for sparse unmixing. Experimental results obtained on both synthetic data and real

data demonstrate that the spectral a priori information is beneficial to sparse unmixing and our proposed algorithm can effectively exploit this information to improve the unmixing results.

ACKNOWLEDGMENT

The authors would like to thank M.-D. Iordache, J. Bioucas-Dias and A. Plaza for sharing their codes for the algorithm of SUnSAL and CLSUnSAL. Besides, the authors would like to gratefully thank the Associate Editor and the four Anonymous Reviewers for their very helpful comments and suggestions, which greatly helped us improve the technical quality and presentation of our manuscript.

REFERENCES

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

[2] P. Shippert, "Why use hyperspectral imagery?" *Photogrammetric engineering and remote sensing*, vol. 70, no. 4, pp. 377–396, 2004.

[3] D. Landgrebe, "Hyperspectral image data analysis," *IEEE Signal Processing Magazine*, vol. 19, no. 1, pp. 17–28, 2002.

[4] M. Petrou and P. G. Foschi, "Confidence in linear spectral unmixing of single pixels," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 1, pp. 624–626, 1999.

[5] Y. Hu, H. Lee, and F. Scarpace, "Optimal linear spectral unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 1, pp. 639–644, 1999.

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

[7] M. Winter, "Fast autonomous spectral end-member determination in hyperspectral data," in *Proc.* 13^{th} *Int. Conf. Appl. Geologic Remote Sens.*, vol. 2, Vancouver, BC, Canada, Apr. 1999, pp. 337–344.

[8] J. M. Nascimento and J. Bioucas-Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 4, pp. 898–910, Apr. 2005.

[9] M. Berman, H. Kiiveri, R. Lagerstrom, A. Ernst, R. Dunne, and J. F. Huntington, "Ice: A statistical approach to identifying endmembers in hyperspectral images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 10, pp. 2085–2095, 2004.

[10] D. D. Lee and H. S. Seung, "Learning the parts of objects by nonnegative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, Oct. 1999.

[11] V. P. Pauca, J. Piper, and R. J. Plemmons, "Nonnegative matrix factorization for spectral data analysis," *Linear algebra and its applications*, vol. 416, no. 1, pp. 29–47, 2006.

[12] X. Chen, J. Chen, X. Jia, B. Somers, J. Wu, and P. Coppin, "A quantitative analysis of virtual endmembers' increased impact on the collinearity effect in spectral unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 8, pp. 2945–2956, 2011.

[13] A. Cichocki, R. Zdunek, A. H. Phan, and S.-i. Amari, *Nonnegative matrix and tensor factorizations: applications to exploratory multi-way data analysis and blind source separation*. Wiley. com, 2009.

[14] D. Seung and L. Lee, "Algorithms for non-negative matrix factorization," *Advances in neural information processing systems*, vol. 13, pp. 556–562, 2001.

[15] S. Jia and Y. Qian, "Constrained nonnegative matrix factorization for hyperspectral unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 1, pp. 161–173, 2009.

[16] Y. Qian, S. Jia, J. Zhou, and A. Robles-Kelly, "Hyperspectral unmixing via $l_{1/2}$ sparsity-constrained nonnegative matrix factorization," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, pp. 4282–4297, 2011.

[17] X. Lu, H. Wu, Y. Yuan, P. Yan, and X. Li, "Manifold regularized sparse nmf for hyperspectral unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, pp. 1–12, Oct. 2012.

[18] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Sparse unmixing of hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 6, pp. 2014–2039, 2011.

[19] M. Iordache, "A sparse regression approach to hyperspectral unmixing," *Universidade Técnica de Lisboa, PhD thesis*, 2011.

[20] J. B. Greer, "Sparse demixing of hyperspectral images," *IEEE Transactions on Image Processing*, vol. 21, no. 1, pp. 219–228, 2012.

[21] J. A. Tropp and S. J. Wright, "Computational methods for sparse solution of linear inverse problems," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 948–958, 2010.

[22] M. Elad, *Sparse and redundant representations*. Springer, 2010.

[23] J. M. Bioucas-Dias and M. A. Figueiredo, "Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing," in *Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 2010 2nd Workshop on*. IEEE, 2010, pp. 1–4.

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

[25] F. Chen and Y. Zhang, "Sparse hyperspectral unmixing based on constrained $l_p - l_2$ optimization," *IEEE Geoscience and Remote Sensing Letters*, vol. 10, pp. 1142 – 1146, Sept. 2013.

[26] W. Tang, Z. Shi, and Z. Duren, "Sparse hyperspectral unmixing using an approximate l0 norm," *Optik-International Journal for Light and Electron Optics*, vol. 125, no. 1, pp. 31–38, Jan. 2014.

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

[28] X.-L. Zhao, F. Wang, T.-Z. Huang, M. K. Ng, and R. J. Plemmons, "Deblurring and sparse unmixing for hyperspectral images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 7, pp. 4045–4058, 2013.

[29] W. Tang, Z. Shi, and Y. Wu, "Regularized simultaneous forward-backward greedy algorithm for sparse unmixing of hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 9, pp. 5271–5288, Sep. 2014.

[30] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Dictionary pruning in sparse unmixing of hyperspectral data," in *2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, 2010, pp. 1–4.

[31] M.-D. Iordache, J. M. B. Dias, and A. Plaza, "Collaborative sparse regression for hyperspectral unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 1, pp. 341–354, Feb. 2014.

[32] Z. Shi, W. Tang, Z. Duren, and Z. Jiang, "Subspace matching pursuit for sparse unmixing of hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 6, pp. 3256–3274, Jun. 2013.

[33] M.-D. Iordache, J. M. Bioucas-Dias, A. Plaza, and B. Somers, "Music-csr: Hyperspectral unmixing via multiple signal classification and collaborative sparse regression," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 7, pp. 4364 – 4382, Jul. 2014.

[34] R. N. Clark, G. A. Swayze, K. E. Livo, R. F. Kokaly, S. J. Sutley, J. B. Dalton, R. R. McDougal, and C. A. Gent, "Imaging spectroscopy: Earth and planetary remote sensing with the usgs tetracorder and expert systems," *Journal of Geophysical Research: Planets (1991–2012)*, vol. 108, no. E12, 2003.

[35] G. A. Swayze, "The hydrothermal and structural history of the cuprite mining district, southwester nevada: An integrated geological and geophysical approach," Ph.D. dissertation, Univ. Of Colorado, 1997.

[36] W. Tang, Z. Shi, and Z. An, "Nonnegative matrix factorization for hyperspectral unmixing using prior knowledge of spectral signatures," *Optical Engineering*, vol. 51, no. 8, pp. 087 001–1–087 001–10, 2012. [Online]. Available: http://dx.doi.org/10.1117/1.OE.51.8.087001

[37] E. Esser, "Applications of lagrangian-based alternating direction methods and connections to split bregman," *CAM report*, vol. 9, p. 31, 2009.

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

[39] G. Obozinski, B. Taskar, and M. Jordan, "Multi-task feature selection," *Statistics Department, UC Berkeley, Tech. Rep*, 2006.

[40] G. Obozinski, M. J. Wainwright, and M. I. Jordan, "High-dimensional support union recovery in multivariate regression." in *NIPS*, 2008, pp. 1217–1224.

[41] J. Liu, S. Ji, and J. Ye, "Multi-task feature learning via efficient l 2, 1-norm minimization," in *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence*. AUAI Press, 2009, pp. 339–348.

[42] D. C. Heinz and C. Chang, "Fully constrained least-squares linear spectral mixture analysis method for material quantification in hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 3, pp. 529–545, Mar. 2001.

[43] C. A. Bateson, G. P. Asner, and C. A. Wessman, "Endmember bundles: a new approach to incorporating

endmember variability into spectral mixture analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no. 2, pp. 1083–1094, 2000.

[44] J. A. Tropp, A. C. Gilbert, and M. J. Strauss, "Algorithms for simultaneous sparse approximation. part i: Greedy pursuit," *Signal Processing*, vol. 86, no. 3, pp. 572–588, 2006.

[45] M. Kowalski and B. Torrésani, "Sparsity and persistence: mixed norms provide simple signal models with dependent coefficients," *Signal, image and video processing*, vol. 3, no. 3, pp. 251–264, 2009.

[46] J. Chen and X. Huo, "Theoretical results on sparse representations of multiple-measurement vectors," *IEEE Transactions on Signal Processing*, vol. 54, no. 12, pp. 4634–4643, Dec. 2006.

[47] Y. C. Eldar and H. Rauhut, "Average case analysis of multichannel sparse recovery using convex relaxation," *IEEE Transactions on Information Theory*, vol. 56, no. 1, pp. 505–519, 2010.

[48] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed. New York: Springer, 2006.

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

[50] B. He, H. Yang, and S. Wang, "Alternating direction method with self-adaptive penalty parameters for monotone variational inequalities," *Journal of Optimization Theory and applications*, vol. 106, no. 2, pp. 337–356, 2000.

[51] S. Wang and L. Liao, "Decomposition method with a variable parameter for a class of monotone variational inequality problems," *Journal of optimization theory and applications*, vol. 109, no. 2, pp. 415–429, 2001.

[52] 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.

[53] T. Goldstein, B. ODonoghue, and S. Setzer, "Fast alternating direction optimization methods," *CAM report*, pp. 12–35, 2012.

[54] R. N. Clark, G. A. Swayze, R. Wise, E. Livo, T. Hoefen, R. Kokaly, and S. J. Sutley, *USGS digital spectral library splib06a*. US Geological Survey Denver, CO, 2007.