MUSIC-CSR: Hyperspectral Unmixing via Multiple Signal Classification and Collaborative Sparse Regression

Marian-Daniel Iordache, José M. Bioucas-Dias, *Member, IEEE*, Antonio Plaza, *Senior Member, IEEE*, and Ben Somers

***Abstract*—Spectral unmixing aims at finding the spectrally pure constituent materials (also called *endmembers*) and their respective fractional abundances in each pixel of a hyperspectral image scene. In recent years, sparse unmixing has been widely used as a reliable spectral unmixing methodology. In this approach, the observed spectral vectors are expressed as linear combinations of spectral signatures assumed to be known *a priori* and presented in a large collection, termed *spectral library* or *dictionary*, usually acquired in laboratory. Sparse unmixing has attracted much attention as it sidesteps two common limitations of classic spectral unmixing approaches, namely, the lack of pure pixels in hyperspectral scenes and the need to estimate the number of endmembers in a given scene, which are very difficult tasks. However, the high mutual coherence of spectral libraries, jointly with their ever-growing dimensionality, strongly limits the operational applicability of sparse unmixing. In this paper, we introduce a two-step algorithm aimed at mitigating the aforementioned limitations. The algorithm exploits the usual low dimensionality of the hyperspectral data sets. The first step, which is similar to the *multiple signal classification* array signal processing algorithm, identifies a subset of the library elements, which contains the endmember signatures. Because this subset has cardinality much smaller than the initial number of library elements, the sparse regression we are led to is much more well conditioned than the initial one using the complete library. The second step applies *collaborative sparse regression*, which is a form of structured sparse regression, exploiting the fact that only a few spectral signatures in the library are active. The effectiveness of the proposed approach, termed MUSIC-CSR, is extensively validated using both simulated and real hyperspectral data sets.**

***Index Terms*—Array signal processing, collaborative sparse regression (CSR), dictionary pruning, hyperspectral imaging, hyperspectral unmixing, multiple signal classification (MUSIC), sparse regression, sparse unmixing, spectral libraries.**

I. Introduction

SPECTRAL unmixing is an important technique for hyperspectral data exploitation [1]. It decomposes the (possibly mixed) pixel spectra measured by an imaging spectrometer into a collection of pure constituent spectra (called *endmembers*) and their corresponding fractional abundances, which quantify the proportion of each pure material in the pixel [2]. Mixed pixels appear due to the relatively low spatial resolution of the sensor flying at high altitudes, or because the materials form intimate mixtures [3]. In a linear spectral unmixing scenario, the mixed pixels can be expressed as a linear combination of the endmember signatures present in the scene weighted by their respective fractional abundances. The exploitation of this model, in spite of its simplicity, has fostered a large amount of research leading to a plethora of endmember extraction and abundance estimation algorithms developed with (and without) the assumption that pure pixels[1] can be found in the original hyperspectral image. A detailed review of techniques developed for spectral unmixing in recent years is available at [4].

Linear spectral unmixing has been recently addressed under a sparse regression framework [5], [6]. The core assumption in this framework is that the observed (generally mixed) spectra measured by a hyperspectral imaging instrument is well approximated by a linear combination of a small (i.e., sparse) subset of spectral signatures selected from a large (usually overcomplete) library. As shown in [4], this sparse unmixing formulation has attracted much attention, as it sidesteps two well known obstacles in classic spectral unmixing methods. First, sparse unmixing does not require the presence of pure signatures in the data, as the endmembers used for spectral mixture modeling are collected from a library of pure signatures. Second, sparse unmixing does not require the estimation of the

Manuscript received March 6, 2013; revised July 22, 2013; accepted August 19, 2013 Date of publication October 9, 2013; date of current version March 3, 2014. This work was supported in part by the European Community's Marie Curie Research Training Networks Programme under Contract MRTN-CT-2006-035927, Hyperspectral Imaging Network (HYPER-I-NET), by the Spanish Ministry of Science and Innovation (CEOS-SPAIN project under Reference AYA2011-29334-C02-02), by the Portuguese Science and Technology Foundation under Project PEst-OE/EEI/LA0008/2013, and by the French National Research Agency, project Hypanema under Grant ANR-12-BS03-0003. The contribution of Ben Somers was supported in part by the Belgian Science Policy Office in the framework of the STEREO II Programme—Projects REMEDY (SR/67/164) and ESSENSE (SR/00/151).

M.-D. Iordache is with the Centre for Remote Sensing and Earth Observation Processes (TAP), Flemish Institute for Technological Research (VITO), 2400 Mol, Belgium.

J. M. Bioucas-Dias is with the Instituto de Telecomunicações and Instituto Superior Técnico, TU Lisbon, 1049-001 Lisboa, Portugal.

A. Plaza is with the Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, Escuela Politécnica, University of Extremadura, 10071 Cáceres, Spain.

B. Somers is with the Centre for Remote Sensing and Earth Observation Processes (TAP), Flemish Institute for Technological Research (VITO), 2400 Mol, Belgium, and also with the Department of Earth and Environmental Sciences, Division Forest, Nature and Landscape Research, Katholieke Universiteit Leuven, 3000 Leuven, Belgium.

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/TGRS.2013.2281589

[1] A pure pixel contains just one endmember.

number of endmembers in a given scene, which has been shown to be a challenging process in the literature.

The ability to obtain meaningful unmixing results by seeking sparse solutions of underdetermined linear systems of equations depends on the degree of sparseness of the mixtures[2] and on the coherence of the library signatures, measured, e.g., in terms of the so-called *mutual coherence* [7] or of the *restricted isometric constants* (RICs) [8]. Qualitatively, the higher the *mutual coherence*, the lower the degree of sparseness ensuring perfect unmixing. Unfortunately, it happens that, in hyperspectral applications, the mutual coherence is often very high (i.e., close to 1) [9], [10], thus, limiting the success of the unmixing via sparse regression. This drawback is somehow mitigated by the very high degree of sparseness (or low number of endmembers) that most hyperspectral applications exhibit. For an extensive study on these issues, see [5].

The aforementioned limitation has been partially mitigated by promoting some type of structured sparsity in the unmixed solutions via suitable regularization terms. Three relevant examples are the total variation spatial regularizer [11], which promotes piecewise-smooth fractional abundances, the group-based regularizer in [12] and [13], which promotes predefined groups of active endmembers in a single fractional abundance vector, and the mixed norm regularizers, which promotes the same set of active endmembers across all fractional abundance vectors [14] [thus, the name collaborative sparse regression (CSR)]. The method we propose in this paper goes beyond this goal, being intended not only to mitigate the coherence drawback, but also to obtain a significant decrease in the running time of the sparse regression algorithms.

A. Related Work

Let us assume that the hyperspectral data set to be unmixed is well approximated by the linear mixing model. In this case, the objective of sparse regression is the determination of the unknown fractional abundance vectors that share a common sparse support. This is the so-called multiple measurement vector (MMV) problem with applications, for example, in distributed compressive sensing, direction-of-arrival (DOA) estimation in radar, magnetic resonance imaging with multiple coils, diffuse optical tomography using multiple illumination patterns (see [15]–[17] and references therein).

The determination of sparse solutions for the MMV problem has been actively pursued in recent years. Relevant examples are greedy methods based on the simultaneous orthogonal matching pursuit [18]–[20], convex relaxation methods using mixed norms [14], [20]–[22], Bayesian methods enforcing a common sparsity profile via suitable prior [23], randomized methods [24], and model-based compressive sensing using block sparsity [25], [26].

The designation "MMV problem" was coined with the advent of sparse regression in the 1990s. However, the same problem, termed as DOA or the *bearing estimation problem* has been addressed since the 1970s by the array signal processing community. The *multiple signal classification* (MUSIC) algorithm independently introduced in [27] and [28] is the most successful method to solve DOA problems. MUSIC first estimates an orthonormal basis for the signal subspace based on the empirical covariance matrix and then identifies the DOAs exploiting the fact that the vectors impinged on the array by the sources are orthogonal to the noise subspace. If the sources are uncorrelated (fractional abundances in our case) and the noise is Gaussian, independent (among sources and snapshots) and identically distributed (i.i.d.), the MUSIC estimator is a large sample realization (large number of spectral vectors in our case) of the maximum likelihood estimator [29]. In addition, in the absence of noise and, if the rank of the observations is equal to the number of active elements in the support of vectors, the MUSIC yields perfect reconstruction [16].

When the rank of the observations is less than the number of active elements in the support of vectors, which is termed the "coherent source" problem within the sensor array signal processing context [30], the MUSIC algorithm fails. However, in this case, the sparse regression algorithms are able to provide useful results even with just one observation. Several sufficient conditions for perfect reconstruction have been derived based on the RICs of the dictionary (see [16] and references therein).

In conclusion, we have two families of methods to solve the MMV problems. On one hand, we have the sparse regression family, which is able to recover, at least in part, the support of the observed vectors, provided that the dictionary has suitable RICs and regardless of the number of observations. On the other hand, we have the MUSIC-based approaches, which yield very good results, provided that the rank of the observations is no less than the number of elements of the support of the vectors. Very recently, the complementarity between MUSIC and sparse regression approaches to solve MMV problems has been exploited [15], [16] by first applying sparse regression methods that identify a subset of the support and then apply MUSIC-based methods. In this paper, we also exploit MUSIC and sparse regression approaches to unmix hyperspectral data, but in a reversed order. Specifically, given a hyperspectral data set and a spectral library with hundreds or thousands of signatures, we aim at pruning the library to only a few tens of spectral signatures, then apply sparse regression methods in much better recoverability conditions and in a much faster way than without applying the pruning. Conceptually, the pruning scheme we propose here is similar to the MUSIC algorithms, but not equal. The pruning step, which was not introduced before in hyperspectral unmixing, represents the principal component of our proposed algorithm. Here, we detail our contribution.

B. Paper Contribution

In this paper, we push forward the research boundary on hyperspectral sparse unmixing. We exploit a simple, yet fundamental, characteristic of hyperspectral data sets: the number of endmembers present in a given scene is often much less than the number of library signatures. Based on this characteristic, we introduce a two-step algorithm aimed at mitigating the aforementioned limitations of hyperspectral sparse unmixing. The first step starts by identifying the signal subspace and then runs

[2]i.e., the number of library signatures with nonzero weights participating in the mixture.

a binary test over the library signatures to identify the endmembers. The signal subspace is estimated with the *hyperspectral subspace identification by minimum error* (HySime) algorithm [31], and the binary test is similar to the MUSIC [27], [28] array signal processing algorithm. If the linear mixing model was an exact fit for observed spectral vectors and there was no noise, then the MUSIC step would, under mild assumptions, exactly identify the endmembers. However, we do have modeling errors and noise. Because of these degrading factors, the binary test in the MUSIC step is designed with a relatively high false alarm probability yielding a set of signatures that, with high probability, contains the endmembers but is, nevertheless, much smaller than the complete library.

The second step is the CSR algorithm introduced in [14] which, in addition to promoting joint sparsity over the abundance fraction vectors (i.e., collaborative sparseness), operates now on a pruned library. Because the dimensionality of the pruned library is, usually, much smaller than the dimensionality of the original library available, the conditioning of resulting sparse regression is naturally improved, and this has a strong positive impact on the quality of the obtained unmixing results as will be shown in this paper.

The remainder of the paper is structured as follows. Section II introduces our MUSIC-based approach to select a subset of the library signatures. Section III emphasizes recent developments in sparse unmixing. Section IV describes the proposed methodology. Section V analyzes the performance of the proposed approach with simulated data. Section VI discusses the performance with real hyperspectral data. Section VII concludes the paper with some remarks and hints at plausible future research lines.

II. Multiple Measurement Problem

Let $\mathbf{y} \in \mathbb{R}^L$ denote an L-dimensional observed spectral vector from a given pixel of a hyperspectral image with L spectral bands and $\mathbf{A} := [\mathbf{a}_1, \ldots, \mathbf{a}_m] \in \mathbb{R}^{L \times m}$ denoting a spectral library with m spectral signatures available *a priori*. Under the linear mixing model, the observed vector $\mathbf{y}$ can be expressed as a linear combination of spectral signatures taken from the library $\mathbf{A}$ as (see [5] for more details)

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

where the vector $\mathbf{x} \in \mathbb{R}^m$ holds the fractional abundances and the vector $\mathbf{n} \in \mathbb{R}^L$ holds the errors affecting the measurements at each spectral band. Because the abundance fractions are nonnegative and sum to 1, the constraints $\mathbf{x} \geq \mathbf{0}$, to be understood in the component-wise sense, and $\mathbf{1}_m^T \mathbf{x} = 1$ ($\mathbf{1}_m$ stands for a column vector with m ones) called abundance nonnegativity constraint (ANC) and abundance sum-to-one constraint (ASC), respectively, are often imposed into model (1).

Assuming that the data set contains $n \gg L$ pixels organized in the matrix $\mathbf{Y} := [\mathbf{y}_1, \ldots, \mathbf{y}_n]$, we can write

$$\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{N} \tag{2}$$

where $\mathbf{X} := [\mathbf{x}_1, \ldots, \mathbf{x}_n]$ is the abundance fraction matrix, and $\mathbf{N} := [\mathbf{n}_1, \ldots, \mathbf{n}_n]$ is the noise matrix. Let us momentarily assume that $\mathbf{N} = 0$ and disregard the ANC and the ASC. In these conditions, finding the sparsest common support for $\mathbf{x}_i$, with $i = 1, \ldots, n$, is precisely an MMV problem, which, formally, corresponds to finding the solution of the optimization, i.e.,

$$\begin{aligned} &\min_{\mathbf{X}} \|\mathbf{X}\|_0 \\ &\text{subject to :} \quad \mathbf{Y} = \mathbf{A}\mathbf{X} \end{aligned} \tag{3}$$

where $\|\mathbf{X}\|_0 := |S|$, $S := \text{supp}(\mathbf{X}) := \{1 \leq i \leq m : \mathbf{x}^i \neq \mathbf{0}\}$, and $\mathbf{x}^i$ is the ith row of $\mathbf{X}$.

The MMV problem (3) has a unique solution if and only if (iff) [32], i.e.,

$$\|\mathbf{X}\|_0 < \frac{\text{spark}(\mathbf{A}) + \text{rank}(\mathbf{Y}) - 1}{2} \tag{4}$$

where $\text{spark}(\mathbf{A})$ denotes the smallest number of linearly dependent columns of $\mathbf{A}$. The term $(\text{rank}(\mathbf{Y}) - 1)/2$ represents the MMV gain; that is, by increasing $\text{rank}(\mathbf{Y})$, which, under suitable conditions, is achieved by increasing the number of measurements, we are able to uniquely solve less sparse problems. Because $\text{rank}(\mathbf{Y}) \leq \|\mathbf{X}\|_0$, we have $\|\mathbf{X}\|_0 < \text{spark}(\mathbf{A}) - 1$.

When $\text{rank}(\mathbf{Y}) < \|\mathbf{X}\|_0$, the MMV problem is combinatorial and therefore very hard to solve exactly. As already referred to in Section I-A, the determination of sparse solutions for the MMV problem has been actively investigated in recent years (e.g., greedy, convex relaxation, Bayesian, randomized, and block-sparsity methods). A much simpler scenario happens, however, when $\text{rank}(\mathbf{Y}) = \|\mathbf{X}\|_0 := k < L$. In this case, we can uniquely recover the support set $\text{supp}(\mathbf{X})$ as follows. Let us now write $\mathbf{Y} = \mathbf{A}_S \mathbf{X}^S$, where $\mathbf{A}_S \in \mathbb{R}^{L \times k}$ and $\mathbf{X}^S$ are the matrices holding, respectively, the columns of $\mathbf{A}$ and the rows of $\mathbf{X}$, whose indixes are in S. Assuming that $\text{spark}(\mathbf{A}) > k + 1$ and that the rows of $\mathbf{X}^S$ are in general position, then $\mathbf{A}_S$ has full column rank and $\mathbf{X}^S$ has full row rank implying that $\text{range}(\mathbf{Y}) = \text{range}(\mathbf{A}_S)$. Therefore, an orthogonal basis for $\text{range}(\mathbf{A}_S)$ can be obtained from the singular value decomposition of $\mathbf{Y}$ or of $\mathbf{Y}\mathbf{Y}^T$. In the latter, we have

$$\mathbf{Y}\mathbf{Y}^T = \mathbf{U}\,\text{diag}(\lambda_1, \ldots, \lambda_k)\,\mathbf{U}^T \tag{5}$$

where $\mathbf{U} \in \mathbb{R}^{L \times k}$, $\mathbf{U}^T\mathbf{U} = \mathbf{I}_k$, $\lambda_1 \geq \lambda_2, \ldots, \lambda_k > 0$, and $\text{range}(\mathbf{U}) = \text{range}(\mathbf{A}_S)$. Now noting that $\mathbf{a}_j \in \text{range}(\mathbf{A}_S)$ iff $j \in S$ (otherwise, we would have $\text{spark}(\mathbf{A}) \leq k + 1$), we conclude that

$$j \in S \quad \text{iff} \quad \mathbf{P}^{\perp}_{\mathbf{A}_S} \mathbf{a}_j = \mathbf{0} \tag{6}$$

where $\mathbf{P}^{\perp}_{\mathbf{A}_S} := \mathbf{I} - \mathbf{U}\mathbf{U}^T$ is the projector on $\text{range}(\mathbf{A}_S)^{\perp}$. Equivalence (6) is the core of the MUSIC algorithm, which in its original version considers vectors $\mathbf{a}$ parameterized by continuous parameters representing DOAs; whereas, in our hyperspectral application, $\mathbf{a}$ is indexed by a finite index set.

In conclusion, we have proved the following result, which is a minor modification of Theorem 3.4 of [16]:

Theorem 1: Given the hyperspectral data set $\mathbf{Y} = \mathbf{A}\mathbf{X} \in \mathbb{R}^{L \times n}$, with $\mathbf{X} \in \mathbb{R}^{m \times n}$, assume that $\text{rank}(\mathbf{Y}) = \|\mathbf{X}\|_0 := k < L$, $\text{spark}(\mathbf{A}) > k + 1$, and the rows of $\mathbf{X}^S$ are in general position. Then, for $j \in \{1, \ldots, n\}$, we have that $j \in \text{supp}(\mathbf{X})$ iff $\mathbf{P}^{\perp}_{\mathbf{A}_S} \mathbf{a}_j = \mathbf{0}$.

A. Support Identification in the Presence of Noise

In real applications, we do have noise, i.e., $\mathbf{N} \neq \mathbf{0}$ in (2), which hinders the estimation of range$(\mathbf{A}_S)$. To shed light on the effect of noise, let us assume that the spectral vectors $\mathbf{y}_i$, for $i = 1, \ldots, n$, are independent samples of a random $\mathbf{y}$ vector with correlation matrix $\mathbf{R}_y := \mathbb{E}[\mathbf{y}\mathbf{y}^T]$ and sample correlation matrix $\widehat{\mathbf{R}}_y := \mathbf{Y}\mathbf{Y}^T/n$. In these conditions, it can be easily concluded that $\mathbb{E}[\widehat{\mathbf{R}}_y - \mathbf{R}_y] = \mathbf{0}$ and that $\mathbb{E}[\|\widehat{\mathbf{R}}_y - \mathbf{R}_y\|^2] = \alpha/n$ for some $\alpha > 0$. Therefore, for n large enough, the sample correlation matrix $\widehat{\mathbf{R}}_y$ can be taken as a good approximation for the correlation matrix $\mathbf{R}_y$. Assuming that the noise and the fractional abundances are independent with correlation matrices $\mathbf{R}_n$ and $\mathbf{R}_x$, respectively, then it follows that

$$\mathbf{R}_y = \boldsymbol{A}_S \mathbf{R}_x(S) \boldsymbol{A}_S^T + \mathbf{R}_n \tag{7}$$

where $\mathbf{R}_x(S)$ is the submatrix of $\mathbf{R}_x$ containing the rows and columns with indexes in S.

White Noise: If $\mathbf{R}_n = \sigma_n^2 \mathbf{I}$, i.e., the noise is i.i.d., then the eigendecomposition of $\mathbf{R}_y$ is

$$\mathbf{R}_y = [\mathbf{U}, \mathbf{V}] \operatorname{diag}\left(\lambda_1 + \sigma_n^2, \ldots, \lambda_k + \sigma_n^2, \sigma_n^2, \ldots, \sigma_n^2\right) [\mathbf{U}, \mathbf{V}]^T \tag{8}$$

where λ_i, for $i = 1, \ldots, k$ is the ith eigenvalue of $\boldsymbol{A}_S \mathbf{R}_x(S) \boldsymbol{A}_S^T$, $\mathbf{U} := [\mathbf{u}_1, \ldots, \mathbf{u}_k]$ holds the first k eigenvectors of $\mathbf{R}_y$ and $\mathbf{V}$ holds the remaining eigenvalues. If $\mathbf{A}_S$ and $\mathbf{R}_x(S)$ are full rank, then $\lambda_i > 0$ for $i = 1, \ldots, k$ and then range$(\mathbf{U})$ = range$(\mathbf{A}_S)$. In conclusion, if the noise is i.i.d., then the signal subspace can be easily estimated from the sample correlation matrix $\widehat{\mathbf{R}}_y$, provided that n is large enough.

Colored or Correlated Noise: If $\mathbf{R}_n \neq \sigma_n^2 \mathbf{I}$, the estimation of range$(\mathbf{A}_S)$ is hard because there is no more a simple relation between the eigenvectors of $\mathbf{R}_y$ and those of $\boldsymbol{A}_S \mathbf{R}_x(S) \boldsymbol{A}_S^T$ [33], [34]. Recently, much attention has been devoted to the estimation of the noise covariance or correlation matrices in the context of DOA problems using, for example, autoregressive or autoregressive moving average models and parametric models [34]. A complementary approach focuses on properties of the signal assuming, for example, that it is non-Gaussian, or that its temporal correlation interval is significantly larger than that of noise, or assuming that the signals are linear combinations of a certain set of known basis functions [34].

In this paper, we adopt the HySime method [31] to estimate the signal subspace range$(\mathbf{A}_S)$. HySime was conceived to exploit the characteristics of hyperspectral data sets. The method starts by estimating the noise correlation and the signal correlation matrices. The noise correlation matrix is estimated via multiple spectral band regression [35], which exploits the high correlation between neighboring spectral bands. HySime then selects the subset of eigenvectors that best represents the signal subspace in the minimum mean square error sense. The application of this criterion leads to the minimization of a two-term objective function. One term corresponds to the power of the signal projection error and is a decreasing function of the subspace dimension; the other term corresponds to the power of the noise projection and is an increasing function of subspace dimension.

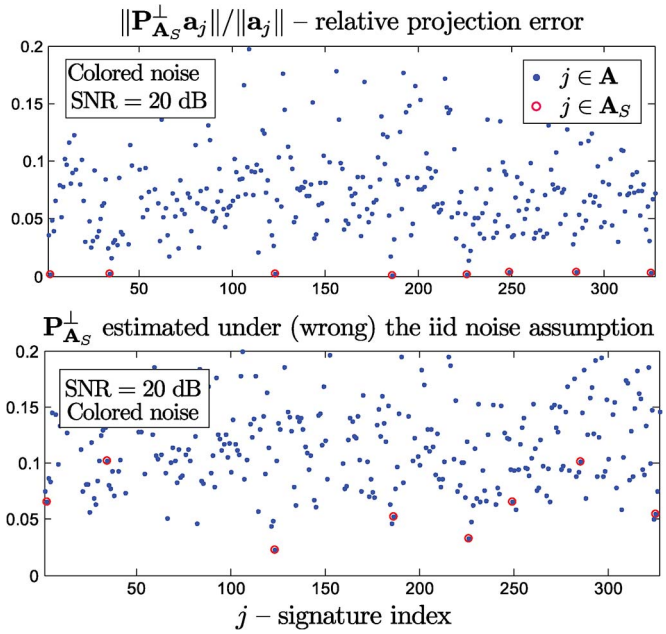

Fig. 1. Norm of projection of the spactral signatures onto the range$(\mathbf{A}_S)^{\perp}$ normalized by the norm of $\mathbf{a}_j$. The values corresponding to $j \in S$ (i.e., the active spactral signatures) are circled. (Top) Range$(\mathbf{A}_S)^{\perp}$ based on the results yielded by the HySime algorithm, thus having into account that the noise is colored. (Bottom) Range$(\mathbf{A}_S)^{\perp}$ computed from the sample correlation matrix under the assumption that the noise is i.i.d.

Fig. 1 provides a numerical illustration of the proposed MUSIC-based method and of the impact of (wrongly) assuming that the noise is i.i.d., when, in fact, it is colored. We have generated a hyperspectral data set according to model (2) with $n = 100\,000$ hyperspectral vectors and library $\mathbf{A}$ formed by a subset of $m = 327$ spectral signatures of size $L = 224$ taken from the U.S. Geological Survey (USGS) splib06[3] library. The subset was chosen such that the angles between any two signatures are no smaller than 3°. The fractional abundances are uniformly distributed over the simplex defined by the active spectral signatures in numbers of $k = 8$ randomly taken from $\mathbf{A}$. The noise $\mathbf{n}$ is Gaussian zero-mean independent from band to band and from vector to vector with a Gaussian-shaped variance along the bands; the Gaussian shape is centered at the middle band and has a spread[4] of 20 bands. The signal-to-noise ratio (SNR $:= \mathbb{E}\|\mathbf{A}\mathbf{x}\|_2^2 / \mathbb{E}\|\mathbf{n}\|_2^2$) is set to 20 dB.

For each spectral signature $\mathbf{a}_j$, with $i = 1, \ldots, m$, we compute the norm of projection onto the range$(\mathbf{A}_S)^{\perp}$ normalized by the norm of $\mathbf{a}_j$, i.e.,

$$\varepsilon_j := \frac{\left\|\mathbf{P}^{\perp}_{\mathbf{A}_S}\mathbf{a}_j\right\|_2}{\|\mathbf{a}_j\|_2}. \tag{9}$$

On the top of Fig. 1, the projection matrix $\mathbf{P}^{\perp}_{\mathbf{A}_S}$ was computed from the range$(\mathbf{A}_S)$ provided by the HySime algorithm. The eight smaller values of ε_i occur, as desirable, for $j \in S$. In contrast, with this very good performance, some values of ε_i for $j \in S$ on the bottom of Fig. 1 are larger than many other errors for $j \notin S$. The reason for this poorer performance is

[3]Available online: http://speclab.cr.usgs.gov/spectral.lib06.

[4]Length of the interval between half variance bands.

the assumption that the noise is i.i.d. and, thus, computing the range$(\mathbf{A}_S)$ based on the sample correlation matrix as described in Section II-A1.

III. Sparse Unmixing of Nonnegative Fractional Abundances Sharing the Same Support

In the previous section, we showed that the active signatures $\mathbf{A}_S$ are effectively detected by the MUSIC algorithm adapted to the hyperspectral scenario. From $\mathbf{A}_S$, we can estimate the corresponding fractional abundances by solving, e.g., the nonnegative constrained least squares (NCLS) optimization problem

$$\begin{aligned} \min_{\mathbf{Z}} \quad & \|\mathbf{Y} - \mathbf{A}_S\mathbf{Z}\|_F \\ \text{subject to :} \quad & \mathbf{Z} \geq \mathbf{0} \end{aligned} \tag{10}$$

where $\|\mathbf{B}\|_F := \sqrt{\text{trace}\{\mathbf{B}\mathbf{B}^T\}}$. We have used the $\mathbf{Z} \in \mathbb{R}^{|S|\times n}$ instead of $\mathbf{X}^S$ to keep the notation light. Notice that we have not included the ASC $\mathbf{1}^T\mathbf{Z} = \mathbf{1}_n$ in (10) because it is hardly satisfied in real applications, namely, owing to spectral variability. See [4] for more details.

In real hyperspectral imaging applications, the detection of support S is hindered by a few degradation mechanisms, which include nonlinearities not captured by the linear model (1), calibration errors between the signatures available in the spectral library, and spectral variability within a given scene. These degradation mechanisms imply errors between the estimated range$(\mathbf{A}_S)$ inferred from data $\mathbf{Y}$ and the subspace spanned by the corresponding elements in the library. These errors may lead to incorrect detection of support S; that is, some of the k smaller indixes ε_i given by (9) may correspond to inactive spectral signatures.

To avoid missing active signatures, we take a conservative approach by setting the probability of missing a signature to a low value. In doing so, we end up with a set R of cardinality $r := |R|$ larger than k but nevertheless much smaller than m, the number of signatures in library $\mathbf{A}$. In addition, we have with high probability $S \subset R$, i.e., we do not miss any active signature. We implement a twofolded strategy, which assists in achieving this goal: first, in order to define the data subspace, we use a number of eigenvectors, which are returned by HySime, larger than the subspace dimensionality inferred by the same algorithm; then, the number of spectral signatures composing the pruned library is set to higher values than the number of endmembers resulting from the HySime estimations. We will denote the reduced library with respect to R as $\mathbf{A}_R$.

The hyperspectral unmixing with respect to $\mathbf{A}_R$, although a much simpler problem than the original one, is still not trivial because matrix $\mathbf{A}_R$ tends to be bad conditioned. We attack this drawback by adding a regularization term to (10), which promotes sparsity among the rows of $\mathbf{Z}$. More specifically, we solve the optimization problem, i.e.,

$$\begin{aligned} \min_{\mathbf{Z}} \quad & \|\mathbf{Y} - \mathbf{A}_R\mathbf{Z}\|_F^2 + \lambda_C\|\mathbf{Z}\|_{2,1} \\ \text{subject to :} \quad & \mathbf{Z} \geq 0 \end{aligned} \tag{11}$$

where $\|\mathbf{Z}\|_{2,1} := \sum_{i=1}^{r} \|\mathbf{z}^i\|_2$ is the mixed $\ell_{2,1}$ norm, which promotes sparseness among the rows of $\mathbf{Z}$. Problem (11) is similar to the collaborative sparse coding problem described in [16], [17], [22], and [36]. The main difference is the introduction of the constraint $\mathbf{Z} \geq 0$. Notice that the NCLS optimization (10) corresponds to (11) with $\lambda_C = 0$.

Let us assume that $r > L$, the noise is zero and that we want to find the sparsest solution of $\mathbf{Y} = \mathbf{A}_R\mathbf{Z}$. In this case, it is shown in [17] that collaborative or multichannel, sparse recovery yields a probability of recovery failure that exponentially decays in the number of channels. In other words, sparse methods have more chances to succeed when the number of acquisition channels increases. Herein, we summarize the results of Theorem [4.4] in [17], which assumes that dictionary $\mathbf{A}_R$ is normalized and composed by i.i.d. Gaussian entries, the observations are generated by a set of atoms whose support is $S \subset \{1, 2, \ldots, m\}$ of cardinality k (i.e., there are at most k rows in the solution matrix, which are not identically zero) and $\|\mathbf{A}_S^+\mathbf{a}_l\|_2 \leq \alpha < 1$ holds for all $l \notin S$, where $\mathbf{A}_S^+$ is the pseudoinverse of the matrix $\mathbf{A}_S$ containing the atoms from $\mathbf{A}_R$ corresponding to the indixes in S. The same theorem states that, under these assumptions, solution $\mathbf{Z}$ of the linear system of equations $\mathbf{Y} = \mathbf{A}_R\mathbf{Z}$ is recovered by solving an $\ell_{2,1}$-norm optimization problem with a probability of at least $1 - m\exp(-(L/2)(\alpha^{-2} - \log(\alpha^{-2}) - 1))$. The exponential decay of the error is obvious as $\alpha < 1$. Although the conditions from the aforementioned theorem are not met in common hyperspectral data, in which the dictionary atoms (that is, the pure spectral signatures) are highly correlated, leading to high values of $\|\mathbf{A}_S^+\mathbf{a}_l\|_2$, we have systematically observed the same type of behavior in our applications. Even when $r \leq L$ and, thus, the system matrix $\mathbf{Y} = \mathbf{A}_R\mathbf{Z}$ is determined or overdetermined, the inclusion of the $\ell_{2,1}$ regularizer in (11) has shown to be beneficial as is tends to set to zero the components of the rows of $\mathbf{Z}$ activated by the noise. In the next section, we give experimental evidence of the advantages of including the $\ell_{2,1}$ regularizer in (11). We compute the solution of (11) with the *collaborative sparse unmixing algorithm via variable splitting* and *augmented Lagrangian* (CLSUnSAL) algorithm [14]. CLSUnSAL is an elaboration of the SUnSAL algorithm [37] designed to enforce the sparsity across pixel vectors. Both SUnSAL and CLSUnSAL exploit the alternating direction method of multipliers developed in [38] and originally introduced in [39].

IV. MUSIC-CSR Algorithm

Here, we describe the proposed hyperspectral unmixing methodology via MUSIC and CSR, which we call MUSIC-CSR. Algorithm 1 shows MUSIC-CSR pseudocode. The inputs to the algorithm are a hyperspectral data set, i.e., $\mathbf{Y}$; a library holding the spectral signatures, i.e., $\mathbf{A}$; the number of signatures to be retained, i.e., r; and the regularization parameter, i.e., λ_C, for the CLSUnSAL algorithm. The output is a pruned library $\mathbf{A}_R$ with r signatures and the fractional signatures $\mathbf{X}^R$. Here, we summarize the main steps of the MUSIC-CSR algorithm.

- *Signal Subspace (line 2)*: Infers the subspace in which the hyperspectral data $\mathbf{Y}$ lives. This is done using the HySime algorithm [31].
- *Projection errors (line 5)*: Computes the distance from each member of the library to the estimated subspace. The

error indicator we use is the normalized Euclidean distance between one member of the library and the estimated subspace in which the data lives.

- *Active set detection (lines 6 and 7)*: Sorts the normalized projection errors by increasing order and retain the indexes of first r in the set R.
- *CSR optimization (line 8)*: Solve a CSR using the identified pruned library $\mathbf{A}_R$ using the CLSUnSAL algorithm [14].

Algorithm 1: MUSIC-CSR

Input: $\mathbf{A} \in \mathbb{C}^{L\times m}$ (library), $\mathbf{Y} \in \mathbb{R}^{L\times N}$ (hyperspectral image), r (number of signatures to be retained), λ_C (regularization parameter)

Output: $\mathbf{A}_R$ (detected signatures), $\mathbf{X}^R$ (fractional abundances with respect to $\mathbf{A}_R$)

1 **begin**

2 $\mathbf{E} :=$ HySime($\mathbf{Y}$) (estimate an orthonormal basis for range($\mathbf{A}_S$) using the HySime algorithm [31])

3 $\mathbf{P}^{\perp}_{\mathbf{A}_S} := \mathbf{I} - \mathbf{E}\mathbf{E}^T$ (projector on range($\mathbf{A}_S^{\perp}$))

4 **for** $j := 1$ **to** m **do**

5 $\varepsilon_j := \dfrac{\|\mathbf{P}^{\perp}_{\mathbf{A}_S}\mathbf{a}_j\|_2}{\|\mathbf{a}_j\|_2}$

6 $\pi :=$ permutation$\{1,\ldots,n : \varepsilon_{\pi(i)} \le \varepsilon_{\pi(j)}, i \le j\}$

7 $R := \{\pi(i), i = 1,\ldots,r\}$

8 Solve the collaborative sparse regression optimization

$$\mathbf{X}^R := \arg\min_{\mathbf{Z}} \|\mathbf{Y} - \mathbf{A}_R\mathbf{Z}\|_F^2 + \lambda_C\|\mathbf{Z}\|_{2,1} \quad \text{subject to:} \quad \mathbf{Z} \ge 0,$$

using the CLSUnSAL algorithm [14].

In the following, we illustrate the proposed MUSIC-CSR algorithm with two toy examples.

First Toy Example: The data set that we use in this example was generated using five randomly selected endmembers from a library containing 302 mineral spectra from the USGS library, which is denoted by splib06,[5] released in September 2007. The number of spectral bands is $L = 224$. We generated 5000 spectra by assuming that the fractional abundances follow a Dirichlet distribution. The datacube was contaminated with i.i.d. Gaussian noise. The variance was set to a value yielding SNR $= 20$ dB. This very high level of noise makes the problem very difficult.

In order to estimate the data subspace in this problem, we used the HySime algorithm, which provided an estimated number of endmembers equal to 5. The first five eigenvectors returned by HySime were used to define the subspace, and the library members were projected to the estimated subspace. For each member, the Euclidean distance to the subspace was then computed. Fig. 2 shows the obtained projection errors for all members. The errors corresponding to true endmembers are highlighted with red circles. Note that these have the lowest projection errors among all the library members. The projection errors were then ordered in ascendent order, so that the first five correspond to the true endmembers, as expected from Fig. 2. The reduced library $\mathbf{A}_R$ was then built by retaining the members corresponding to the lowest $r = 13$ projection errors, which correspond to a maximum allowed error of 0.045. The members are displayed in Fig. 3 (in this figure, the true endmembers are represented in black color). The mutual coherence of the obtained library (0.9983) was slightly lower than that of the original library (0.9986).

[5]Available online: http://speclab.cr.usgs.gov/spectral.lib06.

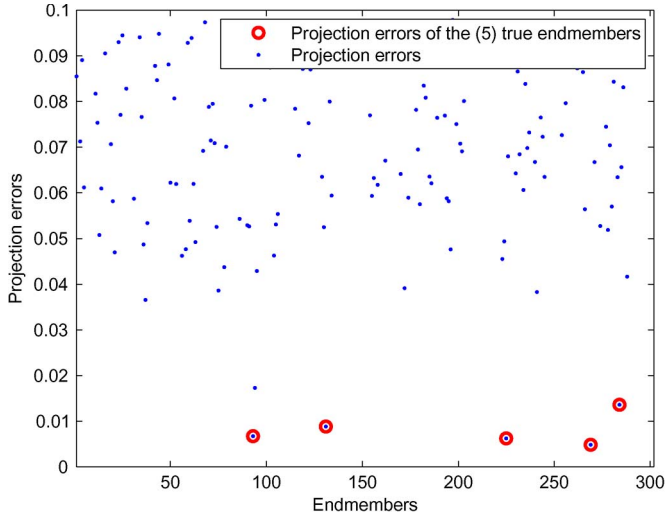

Fig. 2. Projection errors of the library members for the considered simulated toy example. (Red circles) Projection errors of the true endmembers.

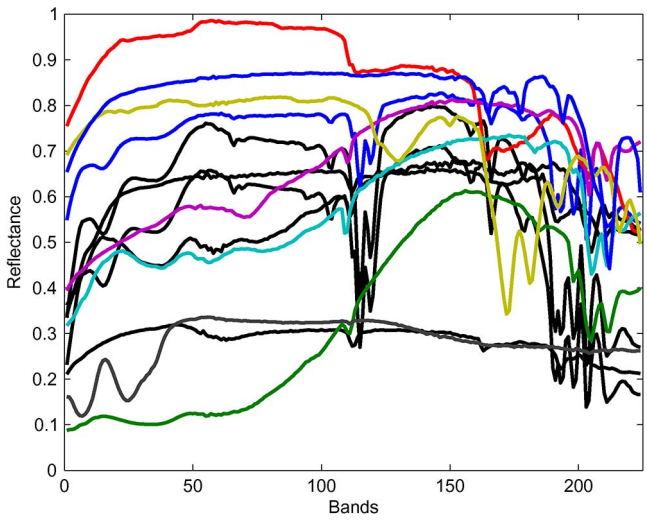

Fig. 3. Spectral signatures from the USGS spectral library, which were retained after the pruning process.

Finally, the unmixing was performed using CLSUnSAL in two different situations, namely, using $\mathbf{A}$, i.e., without pruning (CSR) and using $\mathbf{A}_R$, i.e., after dictionary pruning (MUSIC-CSR). Fig. 4(a) shows the true fractional abundances in the considered simulated toy example. Fig. 4(b) shows the cumulative abundance values (sum of the abundances for all pixels) corresponding to each library member. Fig. 4(c) shows the fractional abundances estimated by CSR with the cumulative abundance values displayed in Fig. 4(d). Finally, Fig. 4(e) shows the fractional abundances estimated by MUSIC-CSR with the cumulative abundance values displayed in Fig. 4(f). In all cases, the algorithm was tuned for optimal performance in the considered scenarios, namely, the parameter λ_C was set to 0.1 for the unmixing cases using $\mathbf{A}$ and to 10^{-2} for the unmixing cases using $\mathbf{A}_R$. However, after testing the MUSIC-CSR for a relatively large range of values of the regularization parameter λ_C (between 10^{-5} and 0.12), we concluded that the obtained performance does not have significant variations, which is a proof of the robustness of the algorithm to this parameter.

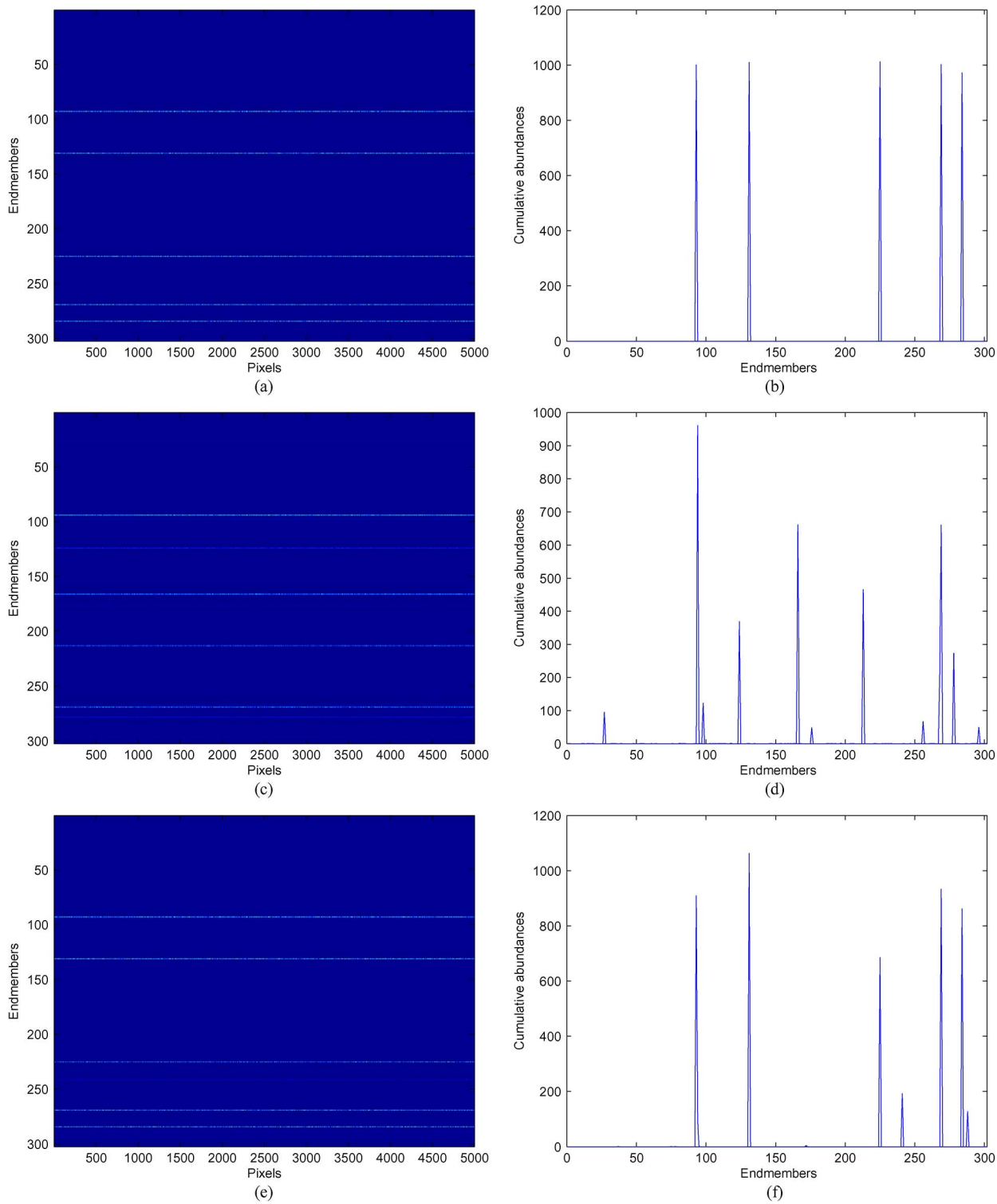

Fig. 4. (a) True fractional abundances in the considered simulated toy example. (b) Cumulative abundance values (sum of the abundances for all pixels) corresponding to each material. (c) Fractional abundances estimated by CSR. (d) Cumulative abundance values estimated using CSR. (e) Fractional abundances estimated by MUSIC-CSR. (f) Cumulative abundance values estimated using MUSIC-CSR.

From Fig. 4, we can conclude that the fractional abundances inferred by MUSIC-CSR are indeed closer to the true ones than those inferred by the CSR. The qualitative differences between the two strategies can be appreciated both in the abundance plots and in the amplitude of the cumulative abundances. Most importantly, the MUSIC-CSR could properly identify the correct set of endmembers in order to explain the observed data. The positions of the true endmembers in the original library were 93, 131, 225, 269, and 284. If we take into account only the highest five values of the cumulative abundances, Fig. 4 indicates that these correspond to the members, i.e., 94, 124, 166, 213, and 269 for the CSR and to members, i.e., 93, 131, 225,

269, and 284 for the MUSIC-CSR. As a result, the MUSIC-CSR is able to identify all the true endmembers that were used to generate the data, whereas the CSR could only correctly identify one of them. This is due to the fact that the original library contained groups of spectral signatures representing the same mineral. The existence of such groups of materials is an already well known feature of the spectral libraries, and it opens many possibilities in unmixing, as this structure can be efficiently exploited by incorporating appropriate terms in the SR objective function. An example is the *group sparse unmixing via variable splitting* and *augmented Lagrangian* [12] algorithm, which is designed to refine the solution at group level (in each pixel, instead of enforcing sparsity on individual materials, it drops entire groups when necessary). As our paper is focused on the pruning strategy and not on the use of any structural characteristic of the spectral libraries, we do not exploit here this property, but we consider it is an interesting direction for future work. However, we mention that the pruning part of MUSIC-CSR can be applied prior to any sparse unmixing algorithm.

Second Toy Example: This experiment is related to an interesting observation on the mutual coherence, which can be drawn from the first toy example. After pruning, the library coherence did not dramatically decrease, as one would expect. A relevant question can be raised: if after the pruning, the mutual coherence is still very high, what is then the advantage of applying the proposed MUSIC-CSR method?

First of all, there clearly is a gain in sparse regression (SR) computation time of orders of magnitude. This is *per se* a huge advantage. However, there is more. The recovery guarantees based on mutual coherence and also on other matrix properties, such as the RICs [40], are sufficient but not necessary conditions. In practice, and in SR unmixing applications, we have systematically observed that the quality of the unmixing improves as the size of the library decreases, although the mutual coherence remains close to 1 [41]. In this example, we investigate in more detail this finding, applied to our specific problem.

In this toy example, the data set consists of a sequence of simulated data sets using an increasing (1–10) number of endmembers from the USGS library. The endmembers were selected such that the angle between them is no larger 3.4°, and the SNR was set to 30 dB, whereas the number of spectral bands is 200. For each number of endmembers, we generated 100 spectral vectors with abundances uniformly distributed in the simplex. In the pruning step, the number of retained spectra was set to 20. The regularization parameter was set to $\lambda_C = 10^{-2}$ in all experiments. Any value of λ_C in the interval $[10^{-4}, 10^{-2}]$ yields comparable SRE results.

Fig. 5 shows two SRE(dB) (i.e., the signal to reconstruction error [5]: $\text{SRE} \equiv \mathbb{E}[\|\mathbf{x}\|_2^2]/\mathbb{E}[\|\mathbf{x} - \widehat{\mathbf{x}}\|_2^2]$, measured in decibels: $\text{SRE(dB)} \equiv 10\log_{10}(\text{SRE})$) plots obtained with the CLSUnSAL algorithm—without and with pruning (continuous and dashed lines, respectively), and the total running time of the algorithms (for all data sets). It should be noted that the mutual coherence of the original library is 0.9999, whereas the one of the pruned library is 0.9989.

From the plots in Fig. 5, we may take two clear conclusions: 1) the results obtained with the pruned dictionary are, approximately, 5 dB above the ones obtained without pruning; and 2) the computation time is approximately 20 times lower with pruning. These findings clearly illustrate the advantage of the proposed approach despite the negligible decrease in the mutual coherence obtained after pruning.

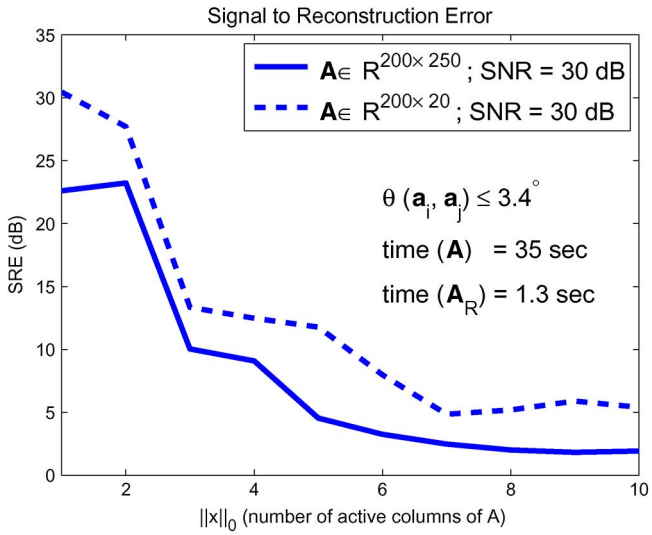

Fig. 5. SRE(dB) and running time obtained by CSR (continuous line) and MUSIC-CSR (dashed line) in the second toy data set.

V. Experimental Results Using Simulated Data

Here, we test the effectiveness of the proposed dictionary pruning method in various simulated scenarios. The section is organized as follows. First, we describe the spectral libraries used in our simulated data experiments and the generated data sets. Then, we describe the considered performance discriminators. We use only data sets affected by noise, as the case in which the observations are not affected by noise is trivial. Next, the sparse unmixing methods presented in Section III are applied to the simulated data sets using both the full library and pruned versions with different numbers of signatures (i.e., in the latter case, step (8) of the MUSIC-CSR algorithm performs unmixing using also NCLS and SUnSAL algorithms, in order to exemplify the applicability of the proposed dictionary pruning methodology shown in steps (2)–(7) of the same algorithm in generic scenarios where per-pixel computation might be necessary). The obtained results are discussed from the viewpoint of estimation accuracy and computational performance. The section concludes with a discussion on the most important observations resulting from the experiments conducted with simulated data.

A. Spectral Libraries

In our experiments with simulated data, we used two spectral libraries with very different characteristics. This allowed us to evaluate the proposed pruning methodology in different scenarios.

For clarity, we establish the terminology used in our experiments, related to the spectral libraries: a) one *endmember class* or *class of materials* or simply *class* represents a specific

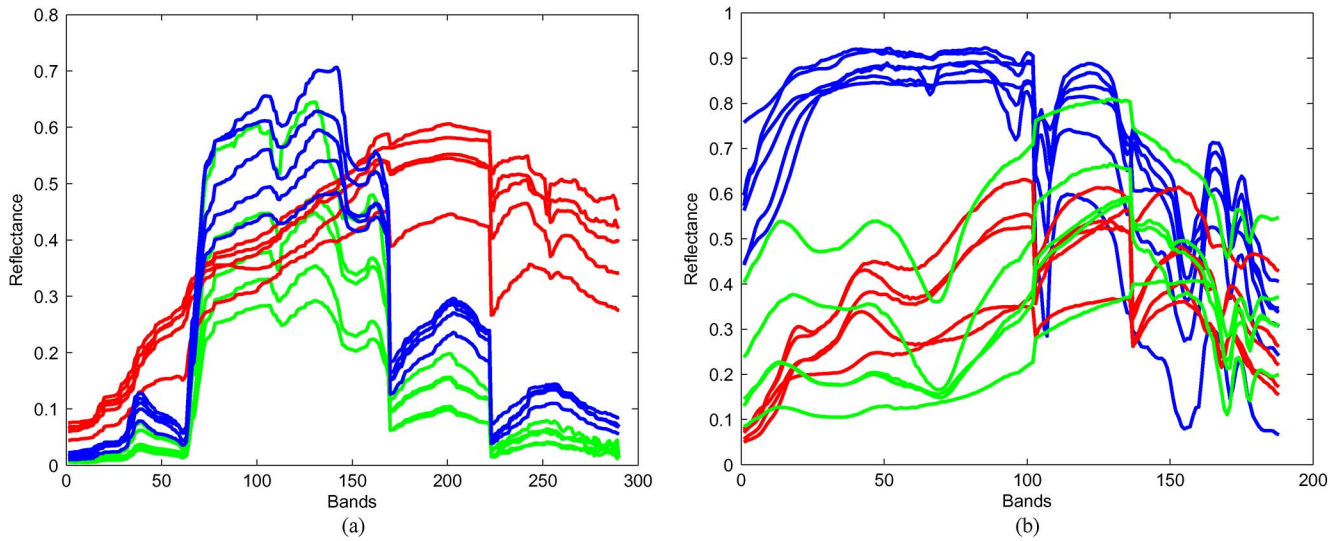

Fig. 6. (a) Five randomly selected spectral signatures from each class in the spectral library $\mathbf{A}_1$. (b) Five randomly selected spectral signatures from each class in the spectral library $\mathbf{A}_2$. (a) $\mathbf{A}_1$ library with (green) tree, (blue) weed, and (red) soil spectra. (b) $\mathbf{A}_2$ library with (green) actinolite, (blue) alunite, and (red) nontronite spectra.

material, structurally distinct from the others (e.g., *tree, soil*); b) one *member of the class* or *library member* or simply *member* is one signature collected in the spectral library; and c) one *true endmember* or *endmember* represents one member contributing to the observed pixel (and implicitly present in the observed scene). Only in the figures, we use the denomination *endmembers* for the library members, to suggest that they represent *a priori* hypothetical endmembers in the image, from which only a few should have nonzero abundances. In the following, we provide a description of the two considered libraries in our experiments with simulated data.

- The first library, which is denoted hereinafter as $\mathbf{A}_1$, contains 50-soil, 200-citrus canopy, and 50-weed canopy reflectance spectra measured in a commercial citrus orchard near Wellington, South Africa, using an analytical spectral device (ASD) field spectroradiometer with a 25° foreoptic, covering the 350- to 2500-nm spectral range. The library was compiled from field campaigns conducted at different periods during the growing season [42]. The major water absorption regions, which were sensitive to changing atmospheric water vapor content, were excluded from the analysis. Then, 290 spectral bands were randomly retained (out of 1798). This was done in order to deal with a very difficult problem, as in this case the mutual coherence of the library approaches one. The library is structured in three groups corresponding to the material classes, namely, soil, citrus canopy, and weed. Fig. 6(a) plots five randomly selected spectral signatures contained in $\mathbf{A}_1$ for each class. Note that, while the soil signatures can be easily distinguished, confusions between the other two (vegetation) classes can easily occur.
- The second library, which is denoted hereinafter as $\mathbf{A}_2$, contains 240 randomly selected signatures from the USGS library, comprising different mineral types. The spectral signatures in this library are made up of reflectance values given in 224 spectral bands and uniformly distributed in the interval 400–2500 nm. The mutual coherence of $\mathbf{A}_2$ also approaches one. The library contains 55 groups of materials, each containing a variable number of signatures (between 1 and 17) describing different alterations of the same mineral. Fig. 6(b) plots five randomly selected spectral signatures contained in $\mathbf{A}_2$ for three different minerals, namely, actinolite, alunite, and notronite.

B. Simulated Data Cubes

The spectral libraries $\mathbf{A}_1$ and $\mathbf{A}_2$ were used to generate various simulated hyperspectral data sets. Specifically, three data cubes were generated using each library corresponding to different numbers of endmembers, namely, $k = \{3, 6, 9\}$. The first three data cubes, which are denoted by DC1, DC2, and DC3, were generated using signatures from $\mathbf{A}_1$; whereas the other three data cubes, which are denoted by DC3, DC4, and DC6, were generated using signatures from $\mathbf{A}_2$. The endmembers were randomly selected as follows. In order to construct DC1, DC2, and DC3, we randomly selected one, two, and three endmembers from each available class, respectively. The main reason for this selection was to avoid the case, in which all the endmembers belong to the same class. On the other hand, at most, one endmember was selected from each class in the construction of DC4, DC5, and DC6. This was done in order to avoid using more than one spectral signature for the same class. All the simulated data cubes comprised 5000 simulated spectra, where the fractional abundances of endmembers follow a Dirichlet distribution uniformly over the probability simplex [43]. The data cubes were then contaminated with white noise, using three levels of SNR, namely, 30, 40, and 50 dB. For clarity, Table I shows the principal characteristics of the simulated data cubes.

C. Performance Discriminators

In our experiments, we use several performance discriminators in order to substantiate the accuracy of the sparse

TABLE I
MAIN CHARACTERISTICS OF THE SIX SIMULATED DATA CUBES CONSIDERED IN OUR STUDY

Simulated Data cube	**DC1**	**DC2**	**DC3**	**DC4**	**DC5**	**DC6**
Spectral library	$\mathbf{A}_1$	$\mathbf{A}_1$	$\mathbf{A}_1$	$\mathbf{A}_2$	$\mathbf{A}_2$	$\mathbf{A}_2$
Number of endmembers (k)	3	6	9	3	6	9
Number of pixels	5000	5000	5000	5000	5000	5000
SNR (dB)	30/40/50	30/40/50	30/40/50	30/40/50	30/40/50	30/40/50

unmixing process, with and without dictionary pruning. An important parameter for the pruned libraries is the number of the correctly retained endmembers and their projection errors, which will be given for all our tests. Regarding the quality of the reconstruction of spectral mixtures, the performance discriminator adopted in this paper is the signal to reconstruction error [5], i.e., $\text{SRE} \equiv \mathbb{E}[\|\mathbf{x}\|_2^2]/\mathbb{E}[\|\mathbf{x}-\widehat{\mathbf{x}}\|_2^2]$, measured in decibels: $\text{SRE(dB)} \equiv 10\log_{10}(\text{SRE})$. We use this measure instead of the root-mean-square error [44] as it gives more information regarding the power of the error in relation with the power of the signal. The higher the SRE(dB), the better the unmixing performance. We also computed the "probability of success," i.e., p_s, which we define as the probability that the relative error power be smaller than a certain threshold and is formally expressed as follows: $p_s \equiv P(\|\widehat{\mathbf{x}}-\mathbf{x}\|^2/\|\mathbf{x}\|^2 \leq threshold)$. This performance measure gives an indication about the stability of the estimation that is not directly inferable from the SRE (which is an average).

In this paper, as we are more interested in finding the correct fractional abundances of the endmember classes (not necessarily of a specific member of the library), we will consider also the case, in which the abundance of one endmember is represented by the sum of abundances of all the members of the class (group of materials) associated to the endmember. In other words, we highlight the performance indicators not only for the individual members of the library (i.e., *per member* assessment), but also for each group considered as a unique endmember (i.e., *per group* assessment). In the former case, the *threshold* value used to compute the p_s is set to 5 dB. This is because we have shown in previous work that solutions attaining this value can be considered of high quality [5]. On the other hand, we also emphasize that, when computing the p_s metric *per group*, most of the unmixing results lead to a higher SRE(dB). This means that the p_s is equal or very close to 1. In this situation, we set a higher quality threshold of 15 dB. By doing this, we can better discriminate between the accuracy achieved by different sparse unmixing methods. The computation times will be also reported for all our experiments.

D. Performance Evaluation

The proposed dictionary pruning methodology was applied to the simulated data cubes by retaining different numbers of signatures $r = \{20, 40, 60\}$. The algorithms discussed in Section III were used to solve the unmixing problem with the full and with the pruned libraries. All algorithms (except NCLS, which is parameter-free) were empirically tuned for optimal performance by carefully adjusting the parameters in each test and reporting only the best obtained results. Four different experiments have been conducted, intended to analyze the quality of the selected endmembers after pruning, the quality of the estimated fractional abundances per member and per group, and the computational performance of the different techniques tested.

1) Experiment 1—Quality of Selected Endmembers After Pruning: Table II shows the number of correctly selected members from the original spectral libraries after the pruning process. From Table II, it can be seen that the proposed pruning method is able to identify the correct endmembers in most cases. Only some difficulties were encountered in situations, in which the number of endmembers is high ($k = 9$), and the SNR is very low (30 dB). On the other hand, the method is able to correctly retain all the endmembers when the SNR is higher than 40 dB, which is a reasonable SNR level in most hyperspectral applications. On the other hand, Fig. 7 illustrates the projection errors measured for the members of $\mathbf{A}_1$ and $\mathbf{A}_2$ after analyzing the data cubes DC1 and DC4, respectively. In the figure, the projection errors measured for the retained members are highlighted with red circles, whereas the green circles indicate the position of the actual endmembers. From Fig. 7, it can be concluded that the true endmembers always exhibit very low projection errors. Although $\mathbf{A}_1$ is characterized by a high variability of the three classes of materials, only a few members are really close to the inferred subspace. Despite $\mathbf{A}_2$ is more heterogeneous (i.e., it contains a larger number of classes, as compared with $\mathbf{A}_1$), a clear gap can be observed in the projection errors measured for the true endmembers and those measured for the other members in the library. This confirms the results reported on Table II.

2) Experiment 2—Accuracy of Estimated Abundances Per Member: Table III shows the SRE(dB) and the p_s achieved by the considered unmixing techniques *per member* for the simulated data cubes generated using the library $\mathbf{A}_1$. Similarly, Table IV shows the same performance indicators for the simulated data cubes generated using $\mathbf{A}_2$. From Tables III and IV, it can be seen that the SRE(dB) and p_s values computed *per member* are generally higher for $\mathbf{A}_2$. This is because $\mathbf{A}_2$ is more heterogeneous in nature than $\mathbf{A}_1$. As expected, the methods that explicitly enforce sparsity (SUnSAL and CLSUnSAL) perform better than NCLS. In general, the performances of all the algorithms systematically improve after dictionary pruning. Not surprisingly, the performances of the algorithms improve when the retained number of signatures approaches the true number of endmembers for both $\mathbf{A}_1$ and $\mathbf{A}_2$. The only exception corresponds to the simulated data cube DC3 (using $\mathbf{A}_1$), in which only 20 members from the original library are retained. However, this behavior is not related to the pruning methodology itself but to the ability of the chosen subspace estimator

TABLE II
NUMBER OF CORRECTLY EXTRACTED MEMBERS FROM THE ORIGINAL LIBRARIES AFTER THE PRUNING PROCESS USING SIMULATED DATA CUBES CONSTRUCTED USING DIFFERENT SNR VALUES

Data Cube (true number of endmembers)	Size of the library after pruning	SNR=30dB	SNR=40dB	SNR=50dB
DC1 (3)	20	3	3	3
	40	3	3	3
	60	3	3	3
DC2 (6)	20	6	6	6
	40	6	6	6
	60	6	6	6
DC3 (9)	20	7	9	9
	40	9	9	9
	60	9	9	9
DC4 (3)	20	3	3	3
	40	3	3	3
	60	3	3	3
DC5 (6)	20	6	6	6
	40	6	6	6
	60	6	6	6
DC6 (9)	20	8	9	9
	40	9	9	9
	60	9	9	9

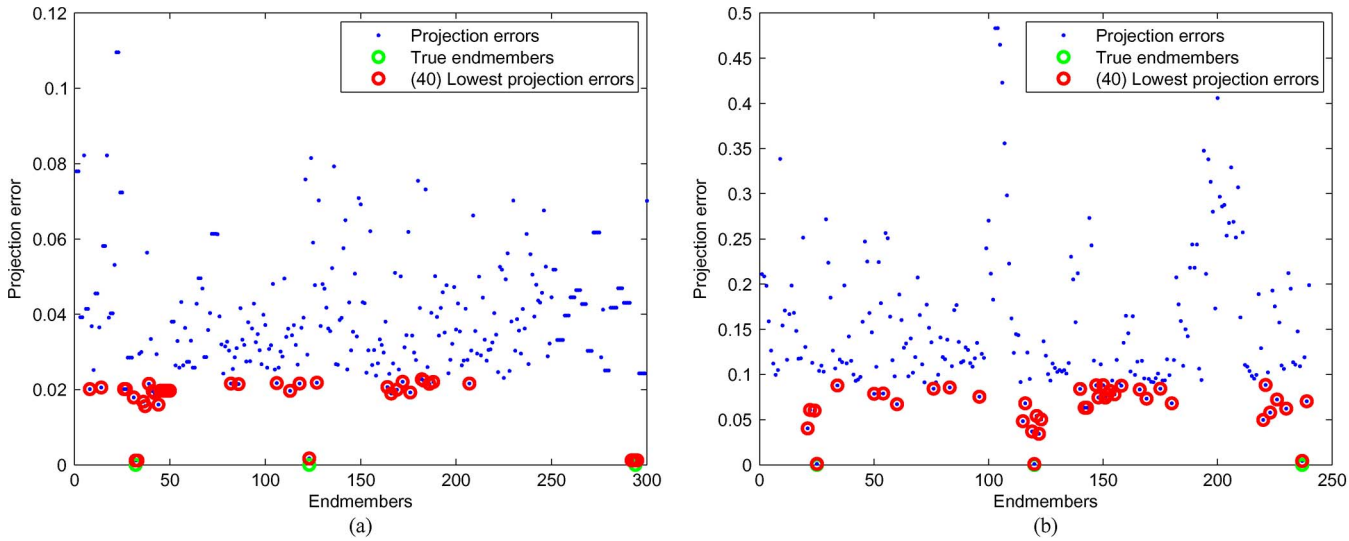

Fig. 7. Projection errors for the members of $\mathbf{A}_1$ and $\mathbf{A}_2$, when the simulated datacubes DC1 and DC4 are contaminated with noise having SNR $=$ 30 dB and 40 members from each library are retained in the pruning process. (a) DC1. (b) DC4.

(in our case, the HySime method) to infer the correct subspace in such environment affected by strong signature variability. Moreover, the spectral confusion between the endmembers is high in this particular case since DC3 was generated using three signatures from each available class of materials (soil, tree, and weed). The same decrease in terms of accuracy is observed for DC5 and DC6 when the observations are affected by high noise (SNR $=$ 30 dB), and only 20 members of the original library were retained. Even so, it can be observed that the unmixing results for any pruning level improve in terms of accuracy when compared with the results obtained with the original library. It should be noted that SNR values of 30 dB are not typically encountered in practice since modern hyperspectral instruments provide high SNR values [45]. Another important observation from Tables III and IV is that the considered unmixing algorithms perform better when the number of endmembers in the image is low (see results for DC1 and DC4). This was already observed in the original sparse unmixing formulation [5]. After analyzing the accuracy of the unmixing results computed *per member*, we conclude that dictionary pruning can significantly improve the obtained unmixing accuracies in most cases.

TABLE III
SRE(dB) AND p_s PER MEMBER BY THE CONSIDERED UNMIXING TECHNIQUES FOR SIMULATED DATA CUBES, NAMELY, DC1, DC2, AND DC3

SRE (dB)										
Method	Library size	DC1 ($k = 3$)			DC2 ($k = 6$)			DC3 ($k = 9$)		
		SNR=30dB	SNR=40dB	SNR=50dB	SNR=30dB	SNR=40dB	SNR=50dB	SNR=30dB	SNR=40dB	SNR=50dB
NCLS	300 (full)	1.25	3.27	3.77	0.10	2.75	4.07	-0.78	0.93	2.58
	60	2.23	3.57	3.82	1.25	3.47	4.26	0.39	2.17	3.45
	40	2.53	3.63	3.82	1.40	3.52	4.27	1.22	2.74	3.56
	20	2.82	3.69	3.83	2.20	3.91	4.34	1.35	3.51	3.75
SUnSAL	300 (Full)	1.44	3.28	3.78	1.19	3.22	4.10	-0.25	0.97	2.76
	60	2.26	3.57	3.82	2.21	3.91	4.26	0.76	2.17	3.45
	40	2.56	3.63	3.82	2.71	4.06	4.34	1.36	2.74	3.56
	20	2.94	3.69	3.83	3.21	4.21	4.36	1.36	3.51	3.75
CLSUnSAL	300 (Full)	2.78	3.68	3.83	2.57	4.05	4.36	0.61	1.86	3.35
	60	3.08	3.73	3.84	3.46	4.21	4.38	2.14	3.20	3.67
	40	3.19	3.75	3.84	3.80	4.26	4.39	2.59	3.36	3.69
	20	3.43	3.78	3.84	3.92	4.32	4.40	1.49	3.54	3.75
p_s										
Method	Library size	DC1 ($k = 3$)			DC2 ($k = 6$)			DC3 ($k = 9$)		
		SNR=30dB	SNR=40dB	SNR=50dB	SNR=30dB	SNR=40dB	SNR=50dB	SNR=30dB	SNR=40dB	SNR=50dB
NCLS	300 (Full)	0.09	0.25	0.31	0.04	0.19	0.39	0	0.01	0.08
	60	0.15	0.29	0.32	0.10	0.30	0.41	0.02	0.08	0.22
	40	0.18	0.29	0.32	0.12	0.30	0.41	0.05	0.12	0.24
	20	0.22	0.30	0.32	0.22	0.36	0.42	0.02	0.22	0.27
SUnSAL	300 (Full)	0.10	0.25	0.31	0.08	0.25	0.39	0	0.02	0.12
	60	0.15	0.29	0.32	0.14	0.36	0.41	0.03	0.08	0.22
	40	0.19	0.29	0.32	0.19	0.38	0.42	0.05	0.12	0.24
	20	0.23	0.30	0.32	0.29	0.40	0.42	0.05	0.22	0.27
CLSUnSAL	300 (Full)	0.21	0.30	0.32	0.13	0.38	0.42	0	0.02	0.20
	60	0.23	0.31	0.32	0.27	0.41	0.43	0.04	0.17	0.25
	40	0.25	0.31	0.32	0.35	0.41	0.43	0.10	0.19	0.26
	20	0.28	0.31	0.32	0.37	0.42	0.43	0.03	0.23	0.27

3) Experiment 3—Accuracy of Estimated Abundances Per Group: Table V shows the SRE(dB) and the p_s achieved by the considered unmixing techniques *per group* for the simulated data cubes generated using the library $\mathbf{A}_1$. Similarly, Table VI shows the same performance indicators for the simulated data cubes generated using $\mathbf{A}_2$. We recall that, in this case, the threshold used to compute p_s was set to 15, which decreases the probability of this performance discriminator to approach one. Tables V and VI are consistent with our observations related to the performance discriminators computed *per member*. Specifically, we can observe improvements in unmixing performance in all cases, particularly, when the noise levels are not very high. For the cases with a low number of endmembers in the original data, the probability of success approaches optimal performance (see DC1 and DC4). The low p_s observed in some cases is due to the high threshold set in computing this performance measure. At the same time, the values of SRE(dB) are generally higher than 5 dB (this value corresponds to high quality in the estimation of the fractional abundances). On the other hand, the unmixing results obtained with pruned libraries are always more accurate than those obtained with the corresponding full library. This is particularly the case for the simulations using $\mathbf{A}_2$ as the baseline library.

4) Experiment 4—Computational Performance of the Considered Algorithms: Table VII reports the computation times measured for the considered algorithms. The times are expressed in milliseconds and correspond to the average running times, per pixel, for fixed values of the regularization parameters ($\lambda = \lambda_C = 10^{-4}$). The maximum number of iterations was set to 1000 in all cases. The NCLS solution was calculated using the SUnSAL algorithm, by setting the regularization parameter to $\lambda = 0$. The algorithms were executed on a desktop PC with an Intel Core Duo CPU @2.5 GHz and 4 GB of RAM memory. The times reported for the unmixing algorithms with dictionary pruning already include the computation time of the pruning process. From Table VI, we can conclude that a significant decrease in computation time can be observed for all algorithms when dictionary pruning is performed. In some cases, the decrease is weakly correlated with the number of members retained from the original library (see, for example, the computation times measured for the SUnSAL method applied to the data cubes generated using $\mathbf{A}_1$). This is because, when the library is pruned, not only the number of computations dramatically decreases but also the algorithms converge faster. As a result, in this case the algorithms do not reach the imposed limit of 1000 iterations. This is a very encouraging result, which further confirms the advantages that can be gained by applying the proposed methodology.

Here, summarizing, we have conducted extensive tests to evaluate the potential of the proposed dictionary pruning methodology using simulated data sets. The libraries used in our tests exhibit distinct characteristics with regard to the type of materials represented in those libraries. While $\mathbf{A}_1$ contains three groups of signatures, corresponding to soil, citrus canopy, and weed, $\mathbf{A}_2$ contains 55 groups of mineral signatures. Our results indicate that dictionary pruning significantly improves

TABLE IV
SRE(dB) AND p_s PER MEMBER BY THE CONSIDERED UNMIXING TECHNIQUES FOR SIMULATED DATA CUBES, NAMELY, DC4, DC5, AND DC6

SRE (dB)

Method	Library size	DC4 ($k = 3$)			DC5 ($k = 6$)			DC6 ($k = 9$)		
		SNR=30dB	SNR=40dB	SNR=50dB	SNR=30dB	SNR=40dB	SNR=50dB	SNR=30dB	SNR=40dB	SNR=50dB
NCLS	240 (Full)	7.86	15.60	21.62	0.77	4.53	9.24	0.25	3.91	8.80
	60	9.70	18.59	28.25	3.45	8.43	15.46	4.52	10.45	18.48
	40	12.63	21.66	31.18	3.53	8.77	16.08	5.46	11.62	19.90
	20	14.01	23.09	32.60	5.07	11.20	19.05	4.79	13.23	21.54
SUnSAL	240 (Full)	7.96	16.06	25.66	1.82	5.75	11.16	1.39	4.25	10.20
	60	10.30	18.59	28.25	3.45	8.43	15.46	4.52	10.45	18.48
	40	12.63	21.66	31.18	3.54	8.77	16.08	5.46	11.62	19.90
	20	14.01	23.09	32.60	5.07	11.20	19.05	4.87	13.23	21.54
CLSUnSAL	240 (Full)	9.62	17.77	27.69	3.51	8.03	15.77	2.75	6.39	12.55
	60	13.33	22.26	31.81	5.80	11.54	18.43	6.37	12.50	20.44
	40	13.41	22.37	31.86	6.02	12.06	19.15	6.71	13.15	21.49
	20	14.34	23.43	32.90	8.89	13.92	20.98	5.76	14.86	23.53

p_s

Method	Library size	DC4 ($k = 3$)			DC5 ($k = 6$)			DC6 ($k = 9$)		
		SNR=30dB	SNR=40dB	SNR=50dB	SNR=30dB	SNR=40dB	SNR=50dB	SNR=30dB	SNR=40dB	SNR=50dB
NCLS	240 (Full)	0.83	1.00	1.00	0.14	0.51	0.97	0.08	0.42	0.96
	60	0.92	1.00	1.00	0.38	0.88	1.00	0.51	0.97	1.00
	40	0.99	1.00	1.00	0.40	0.90	1.00	0.63	0.99	1.00
	20	1.00	1.00	1.00	0.59	0.96	1.00	0.58	0.99	1.00
SUnSAL	240 (Full)	0.84	1.00	1.00	0.20	0.64	0.99	0.11	0.45	0.99
	60	0.94	1.00	1.00	0.38	0.88	1.00	0.51	0.97	1.00
	40	0.99	1.00	1.00	0.40	0.90	1.00	0.63	0.99	1.00
	20	1.00	1.00	1.00	0.59	0.96	1.00	0.58	0.99	1.00
CLSUnSAL	240 (Full)	0.96	1.00	1.00	0.36	0.91	1.00	0.18	0.75	1.00
	60	1.00	1.00	1.00	0.65	1.00	1.00	0.75	1.00	1.00
	40	1.00	1.00	1.00	0.69	1.00	1.00	0.78	1.00	1.00
	20	1.00	1.00	1.00	0.91	1.00	1.00	0.68	1.00	1.00

the unmixing results in all cases, both in terms of the accuracy of the solutions and the time to obtain them. The accuracy of the solutions always improved after the pruning process, regardless of whether they were computed *per member* or *per group*. We have also shown that dictionary pruning allows unmixing algorithms to converge faster to optimal solutions. Although our results with computer simulations are encouraging, further experiments with real hyperspectral data should be conducted.

VI. EXPERIMENTAL RESULTS USING REAL DATA

This section exemplifies the applicability of the proposed dictionary pruning methodology in real environments. In our previous work [41], we conducted a qualitative evaluation of the proposed pruning methodology using the well known Cuprite data set collected by the Airborne Visible Infrared Imaging Spectrometer, which is available online in reflectance units.[6] The aforementioned real data experiment is related to the tests conducted in Section V with the $\mathbf{A}_2$ library, in which a dictionary entirely composed of mineral signatures was used. The results included in [41] demonstrated that the unmixing results (obtained with both NCLS and SUnSAL algorithms) were correlated with the ground-truth data available, and the minerals of interest exhibited good spatial distribution. In that scene, a quantitative evaluation was not possible, as the ground-truth information available is a classification map obtained by the USGS Tricorder algorithm, in which each pixel is assigned to a certain endmember class without providing information about the fractional abundance of each endmember.

In this paper, we use a different real hyperspectral data set for which the true fractional abundances of endmembers are available. The data set comprises *in situ* measurements of reflectance spectra collected over mixed ground plots (i.e., covered by more than one material class) in a commercial citrus orchard near Wellington, South Africa. Mixed pixel spectra were measured for different material class combinations: i) tree–soil; ii) tree–weed; and iii) tree–soil–weed. For each mixture combination, 25 mixed pixel spectra were measured. For each pixel, specific endmember spectra and ground-cover fraction distributions were determined. The reflectance spectra were measured from nadir at a height of 4 m using an ASD field spectroradiometer with a 25° foreoptic, covering the 350- to 2500-nm spectral range. This resulted in a circular area coverage of approximately 2 m. The ASD measurements cover the 350- to 2500-nm spectral range with a spectral resolution of 3 nm (full-width at half-maximum (FWHM)) and a 1.4-nm sampling interval across the 350- to 1050-nm spectral range. The FWHM and the sampling interval for the 1051- to 2500-nm spectral range were 20 and 2 nm, respectively. Resulting data were interpolated during collection by the ASD software to produce values at each nanometer interval. Reflectance was calibrated using a white spectralon panel (Labsphere, Inc., North Sutton, NH). The major water absorption regions, which were sensitive to changing atmospheric water vapor content, were excluded from

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

TABLE V
SRE(dB) AND p_s PER GROUP BY THE CONSIDERED UNMIXING TECHNIQUES FOR SIMULATED DATA CUBES, NAMELY, DC1, DC2, AND DC3

SRE (dB)										
Method	Library size	DC1 ($k = 3$)			DC2 ($k = 6$)			DC3 ($k = 9$)		
		SNR=30dB	SNR=40dB	SNR=50dB	SNR=30dB	SNR=40dB	SNR=50dB	SNR=30dB	SNR=40dB	SNR=50dB
NCLS	300 (Full)	9.06	16.66	24.77	7.93	14.57	22.31	9.22	14.12	19.66
	60	18.27	27.31	37.21	12.92	19.51	28.10	13.08	19.37	28.46
	40	18.58	27.70	37.67	12.88	19.71	28.29	14.04	21.43	30.00
	20	17.89	27.03	36.94	13.93	21.78	30.92	9.83	23.34	32.53
SUnSAL	300 (Full)	11.44	16.84	26.17	12.26	20.66	21.13	8.66	11.22	17.29
	60	19.28	27.31	37.21	12.71	23.61	26.34	11.34	16.34	25.43
	40	19.91	27.70	37.67	13.13	22.28	28.51	11.82	18.40	26.97
	20	19.97	27.03	36.94	17.12	25.49	28.82	8.80	20.31	29.50
CLSUnSAL	300 (Full)	13.80	22.57	32.20	14.67	22.63	31.52	9.51	13.34	21.58
	60	19.80	29.46	38.80	17.14	25.65	34.20	12.37	18.26	26.89
	40	20.60	30.00	39.59	17.85	25.98	35.22	12.67	19.29	27.12
	20	21.24	30.66	40.35	17.92	26.45	35.39	10.83	20.50	29.51
p_s										
Method	Library size	DC1 ($k = 3$)			DC2 ($k = 6$)			DC3 ($k = 9$)		
		SNR=30dB	SNR=40dB	SNR=50dB	SNR=30dB	SNR=40dB	SNR=50dB	SNR=30dB	SNR=40dB	SNR=50dB
NCLS	300 (Full)	0.33	0.75	1.00	0.27	0.61	0.99	0.31	0.59	0.94
	60	0.86	1.00	1.00	0.60	0.90	1.00	0.55	0.91	1.00
	40	0.87	1.00	1.00	0.60	0.91	1.00	0.63	0.97	1.00
	20	0.86	1.00	1.00	0.70	0.95	1.00	0.37	0.99	1.00
SUnSAL	300 (Full)	0.49	0.76	1.00	0.57	0.99	0.99	0.30	0.60	0.95
	60	0.90	1.00	1.00	0.64	1.00	1.00	0.59	0.91	1.00
	40	0.92	1.00	1.00	0.65	1.00	1.00	0.65	0.97	1.00
	20	0.92	1.00	1.00	0.89	1.00	1.00	0.31	0.99	1.00
CLSUnSAL	300 (Full)	0.60	0.99	1.00	0.77	1.00	1.00	0.42	0.78	1.00
	60	0.92	1.00	1.00	0.89	1.00	1.00	0.69	0.98	1.00
	40	0.95	1.00	1.00	0.93	1.00	1.00	0.72	0.99	1.00
	20	0.96	1.00	1.00	0.93	1.00	1.00	0.44	0.99	1.00

the analysis, thus obtaining a data set with 1798 spectral bands. The plot-specific endmembers were acquired by measuring the reflectance of a number of pure soil, sunlit crown, and weed spectra in each plot. Measurements were taken from nadir at 1 m above the object of interest. The ground-truth spectra were collected within the plots in such a way that most of the area in the plot, covered by the targeted endmember, was integrated in the measurements. Measurements were acquired from nadir at a height of 1 m above the object of interest. The average spectrum for each cover class was then used as a reference for the considered endmember in that plot. Fig. 8 shows the experimental setup. As it can be seen in this figure, the field of view of the spectral sensor is demarcated with caution tape on the two crossed wooden laths. S1, S2, and S3 represent subplots selected for the measurements of pure soil spectra. The average of the three reflectance spectra is the specific soil endmember for this plot. T1, T2, and T3 are the subplots selected for the measurements of pure tree spectra. The average of the three reflectance spectra is the specific tree endmember for this plot. The relative cover of each material was extracted from digital images of a Sony DCP-P8/3.2-megapixel camera. This data set has been previously used in [46] and has the advantage that true spectral endmembers and cover fraction distributions are available. For a more detailed description of the experimental setup conducted in order to obtain the reference information for this scene, we refer to [46].

The spectral library used in our experiments, which was denoted by $\mathbf{B}$, is a collection of 971 spectra acquired on the ground in different days, including the day in which the data were acquired. The library comprises 85 soil spectra, 839 tree spectra, and 57 weed spectra, with 1798 spectral bands in the spectral range from 350 to 2500 nm. From these set of spectra, 52 soil spectra, 76 tree spectra, and 49 weed spectra were acquired the same day as the data were collected. Thus, these endmembers can be considered as the ones generating the observations. The algorithms detailed in Section III were used to unmix the data using the full library and two pruned versions, namely, $\mathbf{B}_1$ and $\mathbf{B}_2$, respectively, containing subsets of 200 and 100 spectra from $\mathbf{B}$. In our experiments, we considered only the results computed *per group* as this strategy is the most appropriate, bearing in mind the structure of the hyperspectral data and the available library. Although the true abundance of one class is associated to only one ground-truth spectrum, we allow the solution to contain several active library members, as we are interested in an accurate estimation of the fractional abundances of each class, in all pixels. The threshold used in computing the probability of success is set to 10 dB, and the data subspace was estimated using the HySime algorithm. This approach estimated the subspace dimension to be 56. Consequently, the first 56 eigenvectors returned by HySime were retained to define the estimated data subspace.

Fig. 9 shows the projection errors measured for the members in $\mathbf{B}$, obtained after applying the proposed pruning methodology. The projection errors of the endmembers, acquired on the ground the same day as the hyperspectral data were collected, are highlighted with red circles. The green line represents the

TABLE VI
SRE(dB) AND p_s PER GROUP BY THE CONSIDERED UNMIXING TECHNIQUES FOR SIMULATED DATA CUBES, NAMELY, DC4, DC5, AND DC6

SRE (dB)										
Method	Library size	DC4 ($k = 3$)			DC5 ($k = 6$)			DC6 ($k = 9$)		
		SNR=30dB	SNR=40dB	SNR=50dB	SNR=30dB	SNR=40dB	SNR=50dB	SNR=30dB	SNR=40dB	SNR=50dB
NCLS	240 (Full)	8.40	16.14	22.07	2.48	6.28	10.97	1.20	4.90	9.94
	60	10.18	19.01	28.69	7.14	13.61	21.38	6.13	12.78	21.18
	40	13.63	22.59	32.10	7.34	13.83	21.73	7.39	14.19	22.62
	20	15.15	24.12	33.62	10.07	20.26	28.58	6.45	15.85	24.38
SUnSAL	240 (Full)	8,39	16,60	26,23	4,61	8,60	13,64	2,91	5,96	11,35
	60	10,88	19,01	28,69	7,87	14,28	22,05	6,13	12,78	21,18
	40	13,63	22,59	32,10	8,08	14,50	22,40	7,39	14,19	22,62
	20	15,15	24,12	33,62	10,85	20,93	29,25	6,52	15,85	24,38
CLSUnSAL	240 (Full)	9,84	18,20	27,94	5,94	10,22	18,82	4,01	7,76	13,91
	60	14,03	23,16	32,47	9,39	16,66	24,51	7,77	14,37	22,73
	40	14,24	23,35	32,60	11,25	17,99	25,15	8,62	15,38	23,80
	20	15,23	24,39	34,10	15,07	21,99	30,10	7,43	16,99	25,83
p_S										
Method	Library size	DC4 ($k = 3$)			DC5 ($k = 6$)			DC6 ($k = 9$)		
		SNR=30dB	SNR=40dB	SNR=50dB	SNR=30dB	SNR=40dB	SNR=50dB	SNR=30dB	SNR=40dB	SNR=50dB
NCLS	240 (Full)	0.15	0.71	0.99	0.01	0.13	0.35	0	0.01	0.10
	60	0.34	0.89	1,00	0.09	0.54	0.99	0.02	0.35	0.99
	40	0.54	0.99	1.00	0.12	0.58	0.99	0.06	0.52	1.00
	20	0.66	1.00	1.00	0.36	0.96	1.00	0.09	0.71	1.00
SUnSAL	240 (Full)	0.14	0.74	1.00	0.02	0.22	0.53	0	0.02	0.21
	60	0.35	0.89	1.00	0.09	0.54	0.99	0.02	0.35	0.99
	40	0.54	0.99	1.00	0.12	0.58	0.99	0.06	0.52	1.00
	20	0.66	1.00	1.00	0.36	0.96	1.00	0.08	0.71	1.00
CLSUnSAL	240 (Full)	0.19	0.88	1.00	0.05	0.40	0.93	0	0.03	0.41
	60	0.57	0.99	1.00	0.09	0.77	1.00	0.02	0.49	1.00
	40	0.58	0.99	1.00	0.22	0.89	1.00	0.05	0.64	1.00
	20	0.66	1.00	1.00	0.62	0.99	1.00	0.12	0.80	1.00

threshold used when 200 members were retained from the library $\mathbf{B}$ in the pruning process. Note that the members acquired at the same time as the image data always have projection errors close to zero. On the other hand, there is a clear gap between those members and all the other ones, which might help an end-user decide which signatures from the available library were the ones that really generate the considered hyperspectral data. Fig. 10 plots the best SRE(dB) scores obtained by the unmixing methods described in Section III in each pixel of the considered data set. The average SRE(dB) along with the probability of success p_s are also indicated in the plots. In all cases, the regularization parameters involved in each method were tuned for optimal performance and only the best results from each considered unmixing method are reported. From Fig. 10, we conclude that the obtained unmixing performances always improve after dictionary pruning, an observation that was already emphasized in our experiments with simulated data. Note that, despite the fact that there are 177 spectra contained in the library that generates the data, the obtained results indicated that the unmixing accuracy computed per group still improves when only 100 members are retained. This is due to the fact that the retained spectra can successfully explain the observed data without confusion resulting from the presence of supplementary members in the spectral library.

We remark that, owing to the presence of trees, the light undergoes multiple scattering, and thus, the measured spectrum at each pixel contains nonlinear terms (see [46] for more details). Nevertheless, we are still obtaining useful unmixing results with the proposed approach. The explanation for this unexpected, but desirable, behavior of MUSIC-CSR parallels that of underlying the success of the sparse representation classification (SRC) introduced in [47]. In short, the observed data lives in a low dimensional manifold; given an observed spectral vector in this manifold, the sparse regression identifies a small number of spectral vectors from the library that are close (in a distance sense) to the observed spectral vector and that provide a good linear approximation thereof. Under this perspective, the spectral signatures that are not in the span of the data do not contribute to the local linear approximations and are eliminated by the MUSIC step. This reinterpretation of the MUSIC-CSR method in the case of nonlinear data sets holds great potential as it opens the door to tackle nonlinear mixtures using linear tools as in the case of SRC [47]. A detailed study of this issue is, however, beyond the scope of this paper and will be the subject of future research.

To conclude this section, Table VIII reports the computation times (in milliseconds) measured after applying the proposed unmixing algorithms to the real hyperspectral data set (with and without pruning) on a desktop PC with an Intel Core Duo CPU @2.5 GHz and 4 GB of RAM memory. The times reported in Table VIII correspond to the average running times, per pixel, for fixed values of the regularization parameters, where the parameters involved were optimized and empirically set to $\lambda = \lambda_C = 10^{-3}$. In all cases, the algorithms were set to run at most 1000 iterations. From Table VIII, we can conclude that the computing time of the considered algorithms decreases

TABLE VII
COMPUTATION TIMES (IN MILLISECONDS) FOR THE UNMIXING ALGORITHMS, WITH AND WITHOUT PRUNING, IN SIMULATED EXPERIMENTS

Spectral library: $\mathbf{A}_1$										
Method	Library size	DC1 ($k = 3$)			DC2 ($k = 6$)			DC3 ($k = 9$)		
		SNR=30dB	SNR=40dB	SNR=50dB	SNR=30dB	SNR=40dB	SNR=50dB	SNR=30dB	SNR=40dB	SNR=50dB
NCLS	240 (Full)	12.95	13.86	12.86	13.04	8.52	8.25	13.61	4.24	4.21
	60	1.26	1.21	0.49	0.86	0.39	0.47	0.49	0.40	0.53
	40	0.78	0.89	0.27	0.58	0.26	0.27	0.25	0.26	0.24
	20	0.52	0.42	0.18	0.26	0.12	0.17	0.12	0.13	0.15
SUnSAL	240 (Full)	15.39	15.64	14.61	15.3	10.26	9.93	15.38	5.06	5.31
	60	1.89	1.76	1.18	0.63	0.59	0.51	0.62	0.62	0.56
	40	1.10	1.16	0.34	0.36	0.35	0.32	0.35	0.38	0.31
	20	0.58	0.41	0.21	0.21	0.20	0.21	0.23	0.19	0.17
CLSUnSAL	240 (Full)	47.32	46.72	47.12	46.96	46.91	47.68	47.08	46.77	47.46
	60	12.83	13.19	12.84	12.90	13.52	13.02	12.77	12.99	13.04
	40	10.85	10.60	10.73	10.68	10.67	10.75	10.97	10.81	10.77
	20	8.50	8.64	8.66	8.81	8.65	8.67	8.73	9.34	8.72
Spectral library: $\mathbf{A}_2$										
Method	Library size	DC4 ($k = 3$)			DC5 ($k = 6$)			DC6 ($k = 9$)		
		SNR=30dB	SNR=40dB	SNR=50dB	SNR=30dB	SNR=40dB	SNR=50dB	SNR=30dB	SNR=40dB	SNR=50dB
NCLS	240 (Full)	9.01	8.63	8.64	8.87	9.56	9.83	10.35	8.43	8.44
	60	1.4	1.15	1.23	1.12	1.39	1.01	1.11	0.88	0.44
	40	0.8	0.86	0.81	0.84	0.70	0.51	0.83	0.89	0.29
	20	0.36	0.37	0.35	0.39	0.39	0.21	0.30	0.29	0.21
SUnSAL	240 (Full)	10.39	10.53	11.09	10.31	10.66	11.01	10.38	10.66	11.04
	60	1.85	1.67	1.57	1.57	1.67	1.13	1.60	1.36	0.68
	40	1.16	1.05	0.99	1.07	1.06	0.69	1.05	1.11	0.39
	20	0.55	0.65	0.59	0.58	0.66	0.17	0.62	0.42	0.38
CLSUnSAL	240 (Full)	32.71	32.95	32.67	32.93	33.07	32.91	32.77	33.13	32.94
	60	13.42	12.70	12.93	13.04	13.04	12.61	13.15	13.69	12.74
	40	10.43	10.65	10.42	10.31	10.11	10.39	10.28	10.12	10.41
	20	7.96	8.16	8.03	8.10	8.01	7.93	8.02	8.09	7.93

Fig. 8. Experimental setup to determine a plot-specific soil and tree endmember in the real scene.

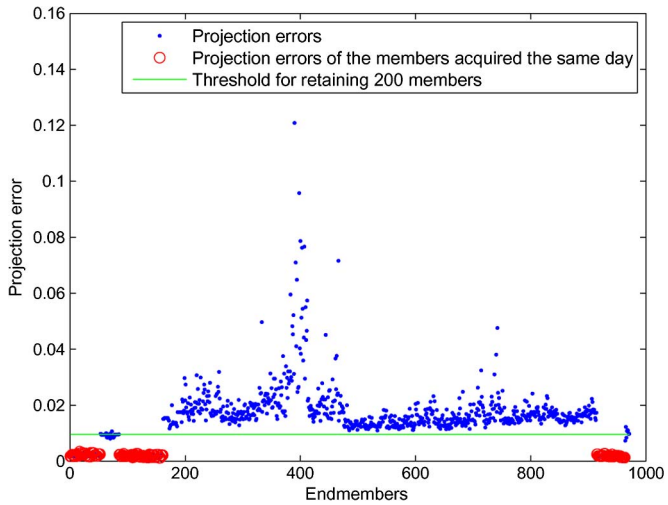

Fig. 9. Projection errors measured for the members in $\mathbf{B}$ in the experiment with real hyperspectral data, obtained after applying the proposed pruning methodology. The projection errors of the members acquired on the ground the same day as the hyperspectral data were collected are highlighted with red circles. The green line represents the threshold used when 200 members were retained from the library $\mathbf{B}$ in the pruning process.

dramatically after the pruning is performed (the pruning time is always included in the times reported in Table VIII). A significant decrease in computation time is observed for unmixing methods that act per pixel (NCLS and SUnSAL). Combined with the experimental results with simulated data, the experiments conducted with real data indicate that the proposed dictionary pruning methodology is able to improve the unmixing accuracy and computational performance of the considered spectral unmixing algorithms.

VII. CONCLUSION AND FUTURE WORK

In this paper, we have developed a new dictionary pruning methodology for spectral libraries intended to increase the accuracy of spectral unmixing algorithms while reducing their computation time. We exploit the fact that hyperspectral pixel

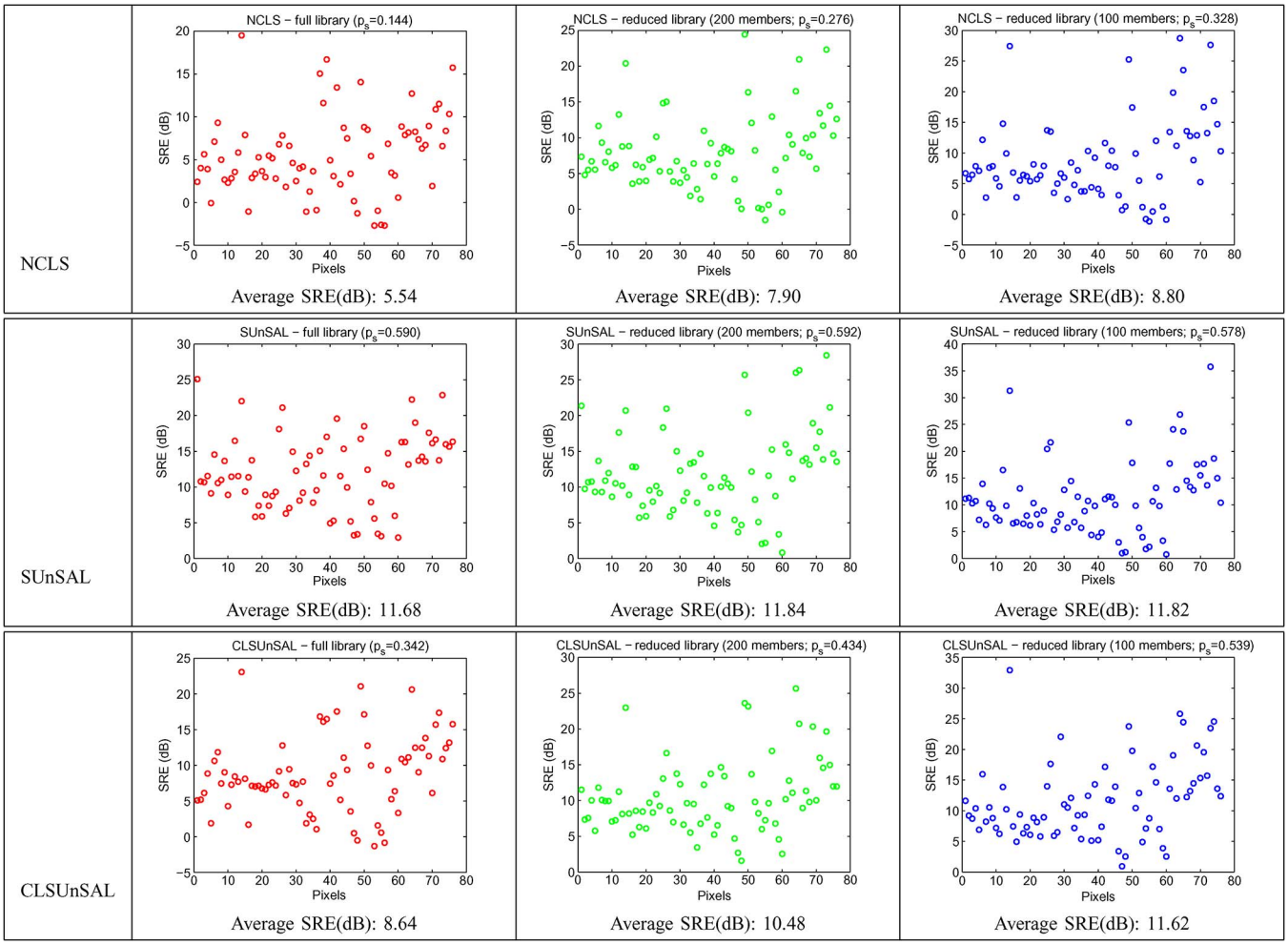

Fig. 10. Best SRE(dB) scores obtained by the considered unmixing methods in each pixel of the real hyperspectral data set. The average SRE(dB) along with the probability of success p_s are also indicated in the plots.

TABLE VIII
COMPUTATION TIMES (IN MILLISECONDS) FOR THE UNMIXING METHODS, WITH AND WITHOUT PRUNING, IN A REAL DATA EXPERIMENT

Method	$\mathbf{B}$ (971 members)	$\mathbf{B}_1$ (200 members)	$\mathbf{B}_2$ (100 members)
NCLS	24.20	1.47	0.80
SUnSAL	24.98	1.66	0.73
CLSUnSAL	125.65	32.08	20.32

vectors generally live in a lower dimensional subspace. The proposed approach has been extensively evaluated using both real and hyperspectral data sets. In all cases, dictionary pruning reveals as a relatively simple yet very powerful strategy to improve unmixing performance, particularly when combined with sparse unmixing algorithms that conduct the estimation of fractional abundances using large spectral libraries, thus circumventing important problems such as the estimation of the number of endmembers (that translates here to the identification of an optimal subset of members in the considered spectral library) and the identification of the spectral signatures of such endmembers when there are no pure observations in the hyperspectral data. With the proposed pruning strategy, we bring sparse unmixing to a new domain in terms of computational performance (now much more manageable than in previous efforts, such as [5], [11], and [14]) and also in terms of unmixing performance, thus enhancing the practical application of sparse unmixing techniques in real problems. A possible direction for future work is to adapt the proposed pruning method to identify spectral signatures based on the physical parameters of the endmembers on the ground. Another research line worth being explored in future developments is how to exploit possible alternative ways of discriminating between actual states of the same endmembers in different scenarios (e.g., seasonal variations of the same endmember in different temporal periods).

REFERENCES

[1] J. Adams, M. Smith, and P. Johnson, "Spectral mixture modeling: A new analysis of rock and soil types at the Viking Lander 1 site," *J. Geophys. Res.*, vol. 91, no. B8, pp. 8098–8112, Jul. 1986.
[2] D. Heinz and C.-I. Chang, "Fully constrained least squares linear mixture analysis for material quantification in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 3, pp. 529–545, Mar. 2001.
[3] N. Keshava and J. Mustard, "Spectral unmixing," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 44–57, Jan. 2002.
[4] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 354–379, Apr. 2011.

[5] 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.
[6] D. Iordache, "A sparse regression approach to hyperspectral unmixing," Ph.D. dissertation, Inst. Superior Tecnico, TU Lisbon, Lisbon, Portugal, 2011. [Online]. Available: http://www.lx.it.pt/~bioucas/files/PhD_Daniel_sparse_regression_2011.pdf
[7] D. Donoho and M. Elad, "Optimal sparse representation in general (nonorthogonal) dictionaries via l1 minimization," *Proc. Nat. Acad. Sci. USA*, vol. 100, no. 5, pp. 2197–2202, Mar. 2003.
[8] E. Candès, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Commun. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, Aug. 2006.
[9] F. D. van der Meer and X. Jia, "Collinearity and orthogonality of endmembers in linear spectral unmixing," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 18, pp. 491–503, Aug. 2012.
[10] 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 Trans. Geosci. Remote Sens.*, vol. 49, no. 8, pp. 2945–2956, Aug. 2011.
[11] 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.
[12] D. Iordache, J. Bioucas-Dias, and A. Plaza, "Hyperspectral unmixing with sparse group lasso," in *Proc. IEEE IGARSS*, Vancouver, BC, Canada, 2011, pp. 3586–3589.
[13] P. Sprechmann, I. Ramirez, G. Sapiro, and Y. Eldar, "C-Hilasso: A collaborative hierarchical sparse modeling framework," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4183–4198, Sep. 2011.
[14] D. Iordache, J. Bioucas-Dias, and A. Plaza, "Collaborative sparse regression for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 341–354, Jan. 2014.
[15] J. M. Kim, O. K. Lee, and J. C. Ye, "Compressive music with optimized partial support for joint sparse recovery," in *Proc. IEEE ISIT*, Aug. 2011, pp. 658–662.
[16] J. M. Kim, O. K. Lee, and J. C. Ye, "Compressive music: A missing link between compressive sensing and array signal processing," *arXiv:1004.4398v5*, Jun. 2011.
[17] Y. Eldar and H. Rauhut, "Average case analysis of multichannel sparse recovery using convex relaxation," *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 505–519, Jan. 2010.
[18] J. Chen and X. Huo, "Theoretical results on sparse representations of multiple measurement vectors," *IEEE Trans. Signal Process.*, vol. 54, no. 12, pp. 4634–4643, Dec. 2006.
[19] A. G. J. Tropp and M. Strauss, "Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit," *Signal Process.*, vol. 86, no. 3, pp. 572–588, Mar. 2006.
[20] S. Cotter, B. Rao, K. Engan, and K. Kreutz-Delgado, "Sparse solutions to linear inverse problems with multiple measurement vectors," *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2477–2488, Jul. 2005.
[21] D. Malioutov, M. Çetin, and A. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3010–3022, Aug. 2005.
[22] J. Tropp, "Algorithms for simultaneous sparse approximation. Part II: Convex relaxation," *Signal Process.*, vol. 86, no. 3, pp. 589–602, Mar. 2006.
[23] D. Wipf and B. Rao, "An empirical Bayesian strategy for solving the simultaneous sparse approximation problem," *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3704–3716, Jul. 2007.
[24] M. Mishali and Y. Eldar, "Reduce and boost: Recovering arbitrary sets of jointly sparse vectors," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 4692–4702, Oct. 2008.
[25] Y. Eldar, P. Kuppinger, and H. Bolcskei, "Block-sparse signals: Uncertainty relations and efficient recovery," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3042–3054, Jun. 2010.
[26] R. Baraniuk, V. Cevher, M. Duarte, and C. Hegde, "Model-based compressive sensing," *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1982–2001, Apr. 2010.
[27] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. AP-34, no. 3, pp. 276–280, Mar. 1986.
[28] G. Bienvenu and L. Kopp, "Adaptivity to background noise spatial coherence for high resolution passive methods," in *Proc. IEEE ICASSP*, 1980, vol. 5, pp. 307–310.
[29] P. Stoica and N. Arye, "MUSIC, maximum likelihood, and Cramer-Rao bound," *IEEE Trans. Acoust., Speech Signal Process.*, vol. 37, no. 5, pp. 720–741, May 1989.
[30] H. Krim and M. Viberg, "Two decades of array signal processing research: The parametric approach," *IEEE Signal Process. Mag.*, vol. 13, no. 4, pp. 67–94, Jul. 1996.
[31] J. Nascimento and J. Bioucas-Dias, "Hyperspectral subspace identification," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 10, pp. 2435–2445, Aug. 2008.
[32] M. Davies and Y. Eldar, "Rank awareness in joint sparse recovery," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 1135–1146, Feb. 2012.
[33] M. Pesavento and A. Gershman, "Maximum-likelihood direction-of-arrival estimation in the presence of unknown nonuniform noise," *IEEE Trans. Signal Process.*, vol. 49, no. 7, pp. 1310–1324, Jul. 2001.
[34] M. Li and Y. Lu, "Maximum likelihood DOA estimation in unknown colored noise fields," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 44, no. 3, pp. 1079–1090, Jul. 2008.
[35] R. Roger and J. Arnold, "Reliably estimating the noise in AVIRIS hyperspectral imagers," *Int. J. Remore Sens.*, vol. 17, no. 10, pp. 1951–1962, Jul. 1996.
[36] B. Turlach, W. Venables, and S. Wright, "Simultaneous variable selection," *Technometrics*, vol. 47, no. 3, pp. 349–363, Aug. 2004.
[37] J. Bioucas-Dias and M. Figueiredo, "Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing," in *Proc. 2nd WHISPERS*, 2010, pp. 1–4.
[38] M. Afonso, J. Bioucas-Dias, and M. Figueiredo, "An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems," *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 681–695, Mar. 2011.
[39] J. Eckstein and D. Bertsekas, "On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators," *Math. Programm.*, vol. 55, no. 3, pp. 293–318, Jun. 1992.
[40] E. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.
[41] D. Iordache, J. Bioucas-Dias, and A. Plaza, "Dictionary pruning in sparse unmixing of hyperspectral data," in *Proc. 4th IEEE GRSS WHISPERS*, Shanghai, China, 2012, pp. 1–4.
[42] J. Stuckens, S. Dzikiti, W. Verstraeten, S. Verreynne, R. Swennen, and P. Coppin, "Physiological interpretation of a hyperspectral time series in a citrus orchard," *Agr. Forest Meteorol.*, vol. 151, no. 7, pp. 1002–1015, Jul. 2011.
[43] 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.
[44] M. Zortea and A. Plaza, "Spatial preprocessing for endmember extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 8, pp. 2679–2693, Aug. 2009.
[45] M. Schaepman, S. Ustin, A. Plaza, T. Painter, J. Verrelst, and S. Liang, "Earth system science related imaging spectroscopy—An assessment," *Remote Sens. Environ.*, vol. 113, no. S1, pp. 123–137, Sep. 2009.
[46] B. Somers, K. Cools, S. Delalieux, J. Stuckens, D. V. der Zande, W. Verstraeten, and P. Coppin, "Nonlinear hyperspectral mixture analysis for tree cover estimates in orchards," *Remote Sens. Environ.*, vol. 113, no. 6, pp. 1183–1193, Jun. 2009.
[47] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.

Marian-Daniel Iordache received the M.Sc. degree from Politehnica University of Bucharest, Bucharest, Romania, in 2006 and the Ph.D. degree in electrical and computer engineering from Instituto Superior Técnico, Lisbon, Portugal, in 2011.

His research activity started in 2006 at the Electrical Engineering Research Center, Bucharest, where he worked on several national and European projects dedicated to microelectromechanical systems and high-frequency circuits. From 2008 to 2011, he was a Marie Curie Fellow with the Hyperspectral Imaging Network project funded by the European Commission under the Sixth Framework Programme. Since 2010, he has been a member of the Hyperspectral Computing Laboratory research group, University of Extremadura, Cáceres, Spain. He is currently carrying out his research as a Postdoctoral Researcher with the Flemish Institute for Technological Research, Center for Remote Sensing and Earth Observation Processes (TAP), Mol, Belgium. His research is focused on hyperspectral unmixing, with emphasis on the sparse characteristics of the solutions.

Dr. Iordache was awarded with The Romanian Student of the Year in Europe 2011 offered by the League of Romanian Students Abroad, for his research activity and social implication.

José M. Bioucas-Dias (S'87–M'95) received the E.E., M.Sc., Ph.D., and "Agregado" degrees in electrical and computer engineering from Instituto Superior Técnico (IST), Technical University of Lisbon, Lisbon, Portugal, in 1985, 1991, 1995, and 2007, respectively.

Since 1995, he has been with the Department of Electrical and Computer Engineering, IST, where he was an Assistant Professor from 1995 to 2007 and has been an Associate Professor since 2007. Since 1993, he has been a Senior Researcher with the Pattern and Image Analysis group, Instituto de Telecomunicações, which is a private non-profit research institution. His research interests include inverse problems, signal and image processing, pattern recognition, optimization, and remote sensing.

Dr. Bioucas-Dias was an Associate Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS from 1997 to 2000. He is an Associate Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING and the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING. He was a Guest Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING for the Special Issue on Spectral Unmixing of Remotely Sensed Data and of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING for the Special Issue on Hyperspectral Image and Signal Processing. He is a Guest Editor of the IEEE SIGNAL PROCESSING MAGAZINE for the Special Issue on Signal and Image Processing in Hyperspectral Remote Sensing. He was the General Cochair of the 3rd IEEE Geoscience and Remote Sensing Society Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS 2011). He has been a member of the program/technical committees of several international conferences.

Antonio Plaza (M'05–SM'07) is an Associate Professor (with accreditation for Full Professor) with the Department of Technology of Computers and Communications, University of Extremadura, Cáceres, Spain, where he is the Head of the Hyperspectral Computing Laboratory (HyperComp). From 2007 to 2011, he was the Coordinator of the Hyperspectral Imaging Network, a European project with total funding of €2.8 million. He authored more than 370 publications, including more than 100 JCR journal papers (60 in IEEE journals), 20 book chapters, and over 230 peer-reviewed conference proceeding papers (90 in IEEE conferences).

Mr. Plaza has guest edited seven special issues on JCR journals (three in IEEE journals). He has been a Chair for the IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (2011). He was a recipient of the recognition of Best Reviewers of the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS (2009) and Best Reviewers of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING (2010), a journal for which he has served as an Associate Editor from 2007 to 2012. He is also an Associate Editor for the IEEE GEOSCIENCE AND REMOTE SENSING MAGAZINE and Associate Editor for IEEE Access. He was a member of the Editorial Board of the IEEE GEOSCIENCE AND REMOTE SENSING NEWSLETTER from 2011 to 2012 and of the Steering Committee of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING in 2012. He was the Director of Education Activities for the IEEE Geoscience and Remote Sensing Society (GRSS) from 2011 to 2012 and has been the President of the Spanish Chapter of the IEEE GRSS since November 2012. He has been the Editor-in-Chief of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING since January 2013.

Ben Somers received the M.Sc. and Ph.D. degrees in bioscience engineering (land and forest management) from the Katholieke Universiteit Leuven (KULeuven), Leuven, Belgium, in 2005 and 2009, respectively.

In 2010, he was a Research Associate with the Geomatics Engineering group, KULeuven. Between 2011 and 2013, he was a Researcher with the Flemish Institute for Technological Research (VITO), Belgium. He is currently a Professor with the Forest, Nature and Landscape research group, Department of Earth and Environmental Sciences, KULeuven. He is experienced in the design of processing tools for hyperspectral remote sensing with a specific focus on spectral mixture analysis and its application in precision farming. His current research interest includes the implementation of advanced remote sensing tools in support of ecological research, including urban ecology.