Regularized Simultaneous Forward-Backward Greedy Algorithm for Sparse Unmixing of Hyperspectral Data

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

Abstract

Sparse unmixing assumes that each observed signature of a hyperspectral image is a linear combination of only a few spectra (endmembers) in an available spectral library. It then estimates the fractional abundances of these endmembers in the scene. Sparse unmixing problem still remains a great difficulty due to the usually high correlation of the spectral library. Under such circumstances, this paper presents a novel algorithm termed regularized simultaneous forward-backward greedy algorithm (RSFoBa) for sparse unmixing of hyperspectral data. RSFoBa has low computational complexity of getting an approximate solution for the l_0 problem directly and can exploit the joint sparsity among all the pixels in the hyperspectral data. Besides, the combination of forward greedy step and backward greedy step makes RSFoBa more stable and less likely to be trapped into the local optimum than the conventional greedy algorithms. Furthermore, when updating the solution in each iteration, a regularizer that enforces the spatial-contextual coherence within the hyperspectral image is considered to make the algorithm more effective. We also show that the sub-library obtained by RSFoBa can serve as input for any other sparse unmixing algorithms to make them more accurate and time-efficient. Experimental results on both synthetic data and real data demonstrate the effectiveness of the proposed algorithm.

Index Terms

Hyperspectral unmixing, sparse unmixing, dictionary pruning, multiple-measurement vector (MMV), greedy algorithm (GA).

Wei Tang and Zhenwei Shi (Corresponding Author) are with Image Processing Center, School of Astronautics, Beihang University, Beijing, P.R. China, 100191 (e-mail: tangwei@sa.buaa.edu.cn; shizhenwei@buaa.edu.cn).

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

I. INTRODUCTION

Hyperspectral unmixing which aims at identifying the constituent spectra (endmembers) and estimating their corresponding fractions (abundances) remains a great challenging task underlying many hyperspectral imagery applications [1–3]. As the linear unmixing model can be implemented easily and flexibly in different applications, it has been widely used for spectral unmixing [4, 5]. This model assumes that each mixed pixel is a linear combination of endmembers weighted by their corresponding abundance fractions. Early unmixing approaches based on geometry [6–9], statistics [10] and nonnegative matrix factorization (NMF) [11–14] extract endmembers merely from the hyperspectral data. However, some of these methods [9–14] could extract virtual endmembers [15] with no physical meaning and others [6–8] assume the presence in the data of at least one pure pixel per endmember, which is usually difficult to guarantee.

A recently developed approach to tackle the problem concerned with the generation of virtual endmembers and the unavailability of pure spectral signatures is to use only a few spectra in a spectral library, known a priori, to model each mixed pixel in the hyperspectral imagery [16–18]. As the number of spectral signatures in the spectral library is usually much larger than the number of endmembers present in the hyperspectral imagery, this approach often leads to a sparse solution. The newly formulated semi-supervised model (called sparse unmixing) needs us to solve a problem containing the non-smooth l_0 regularization which is combinatorial. Fortunately, sparse linear regression techniques can be used [19, 20] to solve it.

Three kinds of methods are usually adopted to tackle the sparse unmixing problem, namely greedy algorithms [16], convex relaxation methods [21–23] and sparse Bayesian methods [24]. The main idea of greedy algorithms[1] is to solve heuristic of making the locally optimal choice at each stage with the hope of finding a global optimum [25]. Specifically, greedy algorithms iteratively refine a sparse solution by successively identifying one or more potential endmembers from the spectral library that make the greatest improvement in reconstructing the mixed pixel [20, 26]. A major flaw of a greedy strategy is that it can never correct mistake made in the earlier steps, which means it could not produce an optimal solution for some problems. Nonetheless, it is able to yield locally optimal solutions that approximate a global optimal solution in reasonable time without smoothing the l_0 norm. Orthogonal matching pursuit (OMP) [27, 28], a typical greedy algorithm, has been utilized to solve the sparse unmixing problem in [16], but the high correlation of the spectral library blocks it from getting the optimal sparse solution. Convex relaxation methods replace the non-smooth l_0 norm with the l_1 norm, thus yielding convex optimization problems that admit tractable algorithm [20]. As the l_1 norm is the closet convex function to the l_0 norm, the replacement is quite natural. It is worth mentioning that there is also some work relaxing the l_0 norm to the l_p norm where $0 < p < 1$ [29]. Though the l_p norm is closer to the l_0 norm than the l_1 norm, it leads to a non-convex problem. As a result, the problem is computationally complex to solve and we can only obtain a suboptimal solution for it. Convex relaxation methods are more sophisticated than the greedy algorithms as they obtain the global solution of a well-defined optimization problem [19]. However, convex relaxation methods cannot control the sparsity of

---

[1]The greedy algorithms here denote forward greedy algorithms.

the solution directly (unlike the greedy algorithms) and are more complicated than the greedy algorithms. Several efficient convex relaxation methods [16, 21] have been proposed to solve the sparse unmixing problem based on the alternating direction method of multipliers (ADMM) which can decompose a difficult problem into a sequence of simpler ones [30, 31]. The sparse Bayesian methods [32, 33] have been applied to the sparse unmixing problem recently. In [24], sparse unmixing is formulated as a hierarchical Bayesian inference problem and priors for the model parameters and hyperparameters are selected to ensure the nonnegativity and sparsity of the abundances' vector. Then an iterative scheme termed Bayesian inference iterative conditional expectations (BI-ICE) is proposed to identify the maximum a posteriori estimation of the parameters. The BI-ICE algorithm can provide the sparse solution without necessarily tuning any parameters. However, it is much more computationally expensive than the greedy algorithms and convex relaxation methods. Besides, since the noise in the hyperspectral images is usually correlated [16], the assumption in the BI-ICE model that the additive noise is a zero-mean Gaussian distributed random vector, with independent and identically distributed (i.i.d.) elements, could be violated. Most recently, efforts have been paid to exploit the spatial-contextual information [22, 34] within the hyperspectral image and joint sparsity among all the pixels [23] and combine them into the convex relaxation methods to obtain a much better result.

In this paper, we propose a novel algorithm, regularized simultaneous forward-backward greedy algorithm (RSFoBa), for sparse unmixing of hyperspectral data. The proposed algorithm is inspired by the simultaneous orthogonal matching pursuit algorithm (SOMP) [35] and the adaptive forward-backward greedy algorithm (FoBa) [36]. Different with the existing greedy algorithms which solve a plain least squares problem when updating the solution in each iteration, the proposed algorithm solves a regularized least squares problem which enforces spatial-contextual coherence constraint on the solution. Thus the solution can be estimated more accurately in each iteration. To solve the regularized least squares problem, we need to solve a Lyapunov equation. There are mainly two contributions of this paper:

- In this paper, we formulate the RSFoBa algorithm based on the forward-backward algorithm proposed in [36] which has the advantages of both the forward and backward greedy algorithms. Specifically, as backward steps are carried out in each iteration to remove any errors caused by the earlier forward steps and avoid maintaining a large number of potential endmembers, RSFoBa can alleviate the drawback of classical greedy algorithms such as OMP that they cannot correct mistakes made in the earlier steps and thus are prone to be trapped into local optimum. Moreover, inspired by the SOMP algorithm [35], RSFoBa also deals with the whole hyperspectral data simultaneously rather than processes each pixel individually. Thus, it can exploit the joint sparsity among all the pixels in the hyperspectral data to alleviate the difficulty brought by the usually high correlation of the spectral library.

- As mentioned above, a regularizer that enforces spatial-contextual coherence within the hyperspectral image is considered when updating the solution in the proposed algorithm. Thus, RSFoBa can further use the spatial-contextual information in the hyperspectral image to select the actual endmembers more accurately. Theoretical results are presented to analyze the convergence and computational complexity of the proposed algorithm. We

also show that the sub-library obtained by RSFoBa can serve as input for any other sparse unmixing algorithms to make them more efficient, which means RSFoBa not only can be an effective sparse unmixing algorithm but also can be a preprocessing step to boost other sparse unmixing algorithms.

The rest of the paper is structured as follows. Section II introduces some related work concerned with the sparse unmixing model and the greedy algorithms. In Section III we present the regularized simultaneous forward-backward greedy algorithm. Experimental results are shown in Section IV. Finally, we conclude in Section V.

II. RELATED WORK

In this section, we begin with the introduction of the linear sparse unmixing model. Then the forward, backward and forward-backward greedy algorithms will be discussed, which lays a foundation for the proposed algorithm.

A. Sparse Unmixing Model

The linear sparse unmixing model assumes that the observed pixel is a linear combination of only a few spectral signatures in an available library. Let $\mathbf{A} \in R^{L \times m}$ denote the spectral library, where L is the number of bands and m is the number of spectral signatures in the library. Then, the sparse unmixing model is [16]:

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

where $\mathbf{y} \in R^L$ is the spectrum vector of a mixed pixel in a hyperspectral image with L bands, $\mathbf{x} \in R^m$ is the abundance vector with regard to the library $\mathbf{A}$ and $\mathbf{n} \in R^L$ is the vector of error term. Owing to physical constraints, the model has the following two constraints (x_i is the i-th element of $\mathbf{x}$):

$$\mathbf{x} \geq 0 \tag{2}$$

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

which are called abundance nonnegativity constraint and sum-to-one constraint, respectively [37]. But due to the strong signature variability in the real image, the sum-to-one constraint should be replaced with a so-called generalized sum-to-one constraint which is automatically imposed because of the nonnegativity of the sources [16, 38]. Thus we do not explicitly add the sum-to-one constraint to the sparse unmixing model in this paper.

Then the optimization problem of sparse unmixing can be written as follows:

$$\min_{\mathbf{x}} \mathbf{q}(\mathbf{x}) = \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 \quad \text{subject to} \quad \|\mathbf{x}\|_0 \leq k, \mathbf{x} \geq 0 \tag{4}$$

where $\|\mathbf{x}\|_0$ (called l_0 norm) denotes the number of nonzero components in $\mathbf{x}$, $k \geq 1$ is a given sparsity level.

The simultaneous sparse unmixing model (also called collaborative sparse regression model [23]) assumes that all the pixels in the hyperspectral image can be modeled using the same subset of endmembers selected from the spectral library. Considering the whole hyperspectral data, the sparse unmixing model in Eq. (1) becomes

$$\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{N} \tag{5}$$

where $\mathbf{Y} \in R^{L \times K}$ is the hyperspectral data matrix with L bands and K mixed pixels, $\mathbf{A} \in R^{L \times m}$ is the spectral library, $\mathbf{X} \in R^{m \times K}$ is the abundance matrix each column of which corresponds with the abundance fractions of a mixed pixel and $\mathbf{N} \in R^{L \times K}$ is the matrix of error term.

Under this model, the unmixing problem is:

$$\min_{\mathbf{X}} \mathbf{Q}(\mathbf{X}) = \frac{1}{2}\|\mathbf{Y} - \mathbf{A}\mathbf{X}\|_F^2$$

$$\text{subject to} \quad \|\mathbf{X}\|_{\text{row-0}} \leq k, \mathbf{X} \geq 0 \tag{6}$$

where $\|\mathbf{X}\|_F$ means the Frobenius norm of matrix $\mathbf{X}$, $\|\mathbf{X}\|_{\text{row-0}}$ (called row-l_0 quasi-norm [35]) is the number of nonzero rows in matrix $\mathbf{X}$, $k \geq 1$ is a given row-sparsity level. In [23], the problem in Eq. (6) is relaxed to a smooth convex problem and a collaborative sparse unmixing method is developed to solve it.

It is worth mentioning that the model in Eq. (6) is reasonable for sparse unmixing because there should be only a few nonzero rows in the abundance matrix $\mathbf{X}$ because there are always a small number of endmembers in a hyperspectral image, especially compared with the very large spectral library [23]. This model also has several advantages over the model in Eq. (4): i) as researched in the literature of multiple-measurement vectors (MMV) [39], the model in Eq. (6) is more probable to have a unique solution than the model in Eq. (4); ii) the constraint that all the pixels in a hyperspectral image are constituted by the same set of endmembers can alleviate the sparse regression limitations caused by the usually high correlation of the spectral library [23]; iii) most of the time, simultaneous sparse regression is superior to the non-simultaneous one [40]. Nevertheless, there are also some cases in which the nonzero rows of $\mathbf{X}$ are not sparse but each pixel adopts a sparse representation in $\mathbf{X}$, especially when the hyperspectral image is very complex. In this paper, we only focus on the situations in which the number of endmembers present in the hyperspectral image is much smaller than the number of spectral signatures in the spectral library.

B. Forward, Backward and Forward-backward Greedy Algorithms

In the literature of sparse signal reconstruction or sparse representation [19], there are mainly two kinds of greedy algorithms to solve the l_0 problem in Eq. (4) approximately, namely forward greedy algorithm and backward greedy algorithm.

The forward greedy algorithm begins with an all-zero solution and successively adds one or more nonzero elements to the solution that yield the greatest improvement in minimizing the objective function until the desired number of nonzero elements are obtained or the reconstruction error of the input signal is below the tolerance. Two typical forward greedy algorithms are the matching pursuit (MP) [26] and orthogonal matching pursuit (OMP) [27, 28]. A major flaw of the forward greedy algorithm is that it can never correct mistakes made in the earlier steps [36]. For example, consider the situation in Fig. 1. In this figure, suppose $\mathbf{y}$ is the spectral vector of a mixed pixel and it is constituted by two endmembers $\mathbf{a}_1$ and $\mathbf{a}_2$, but another spectral signature in the spectral library $\mathbf{a}_3$ is closer to $\mathbf{y}$ than the two actual endmembers. Thus, when the forward greedy algorithm such as OMP is used to unmix the mixed pixel, it will find $\mathbf{a}_3$ first and then $\mathbf{a}_1$ and $\mathbf{a}_2$. The problem is that we are not able to remove $\mathbf{a}_3$

selected in the first step, which will have a negative impact on the subsequent endmember selection and abundance estimation. This line of reasoning indicates that the forward greedy algorithm is inadequate for sparse unmixing. Indeed, the method only works well and stably when the spectral library is not correlated, which is not the case of sparse unmixing.
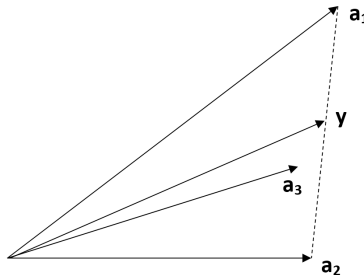

Fig. 1. Illustration of the flaw of the forward greedy algorithm.

Contrary to the forward greedy algorithm, the backward greedy algorithm begins with a dense solution estimated using the whole spectral library and one element that increases the reconstruction error least is set to zero at each step [41, 42]. Though it seems that the backward greedy algorithm has a broader perspective than the forward greedy algorithm and thus will not be trapped into the local optimum, the algorithm is computationally very expensive because it begins with the whole spectral library. Besides, when the spectral library matrix is overcomplete (especially when $m \gg L$), removing any spectral signature is helpless: the remaining sub-matrix of the spectral library is still overcomplete no matter which member is removed from the spectral library. In [16], a backward greedy algorithm, iterative spectral mixture analysis algorithm (ISMA) [43], is used for sparse unmixing. The experimental results [16] indicate that ISMA behaves worse than the convex relaxation algorithms and forward greedy algorithms and it is the most time-consuming algorithm.

In sum, the main strength of the forward greedy algorithm is its computational efficiency because it always deals with a sparse solution explicitly [36]. But its major flaw is that it cannot correct any error made in the earlier steps. However, the backward greedy step can potentially correct such an error but needs to start with a small subset of the spectral library that is non-overcomplete. The comparison of the two kinds of greedy algorithms leads us to consider seeking a combination of them to overcome the flaws of both methods. In the past few years, there have been some work concerned with such a combination (called forward-backward greedy algorithm) [36, 44]. But to the best of our knowledge, such algorithms have never been used in the field of hyperspectral image processing. In this paper, we consider using the adaptive forward-backward greedy algorithm (FoBa) proposed in [36] for sparse unmixing. FoBa has been proved to be effective in learning sparse representations both experimentally and theoretically [36].

As a preprocessing step, we successively apply the zero-mean normalization and l_2 unit-length normalization to both the input pixel vector and the columns of spectral library matrix. For convenience, we use $\mathbf{z}$ instead of $\mathbf{x}$ to update the solution in each iteration. $\mathbf{z}$ could be considered as the elements of $\mathbf{x}$ that have the indices of

the support of selected indices $\mathbf{S}$, which means $\mathbf{z}$ is not necessarily sparse. The whole process of using FoBa for sparse unmixing of hyperspectral data is summarized in Algorithm 1. The algorithm includes two main parts: endmember selection and abundance estimation. In the first part, the normalized hyperspectral signature vector and the normalized spectral library will be used while the original ones are used for estimating the abundances. In the endmember selection part, steps 7-12 are the forward step and steps 13-21 are the backward step. Step 10 computes the decrease of $\mathbf{q}(\mathbf{z}; \mathbf{S}) = \frac{1}{2}\|\mathbf{A_S z} - \mathbf{y}\|_2^2$ after selecting an additional member from the spectral library. If the decrease of $\mathbf{q}$ is smaller than the preset tolerance, the endmember selection part will stop. Otherwise, the backward step will be implemented. In step 14, the member that increases the function $\mathbf{q}$ least will be taken as a candidate to be removed. If the increase of $\mathbf{q}$ (d^-) after removing the candidate member is no more than half of the decrease of $\mathbf{q}$ in the earlier corresponding forward step (d^+), the candidate member will be removed. In this way, we can expect the FoBa algorithm to have the following two properties [36]: i) it can take reasonably aggressive backward steps to remove any errors made in the earlier forward steps and to avoid maintaining too many potential endmembers; ii) it can take backward steps adaptively and guarantee that any backward greedy step does not erase the gain made in the forward steps, which means we are always making progress. Finally, the abundances are estimated using the selected potential endmembers under the constraint of nonnegativity.

However, when the spectral library is very correlated, any algorithm (including the FoBa algorithm) that mimics l_0 regularization is unstable [36]. Thus, additional prior knowledge and regularization that can exploit the characteristics of the hyperspectral data are needed to make the forward-backward greedy algorithm more effective.

III. SIMULTANEOUS FORWARD-BACKWARD GREEDY ALGORITHM WITH SPATIAL-CONTEXTUAL COHERENCE REGULARIZATION

Here, we will present the proposed algorithm by incorporating the joint sparsity among all the pixels and spatial-contextual information of the hyperspectral image into the forward-backward greedy algorithm. Then, we give a convergence theorem for the proposed algorithm.

A. Simultaneous Forward-backward Greedy Algorithm

Some prior knowledge concerned with the hyperspectral data is considered in the proposed algorithm. The first prior knowledge we consider is the joint sparsity among all the pixels in the hyperspectral data. Hyperspectral remote sensing has been evolved to be capable of acquiring simultaneously in 100 to 200 contiguous spectral bands [45]. However, researchers have shown that the hyperspectral data cannot fill up the whole high-dimensional Euclidean space [14]. Instead, they are more likely to lie in a low-dimensional submanifold of a high-dimensional ambient space [14, 46]. For the problem of sparse unmixing, the high-dimensional hyperspectral data live in a lower dimensional subspace of the spectral library. Thus only a small set of the spectral library is active in constructing all the pixels in the hyperspectral data [47], which is called the joint sparsity among all the pixels in the hyperspectral data [23]. Thus, we can utilize this characteristic of the hyperspectral data for sparse unmixing. Then we need to solve the problem in Eq. (6). A natural way to solve this problem is to use the simultaneous greedy algorithm

---

Algorithm 1 Forward-backward greedy algorithm for sparse unmixing of hyperspectral data

---

Preprocessing:

1: Apply zero-mean normalization and l_2 unit-length normalization successively to the vector of spectral signature $\mathbf{y}$ and spectral library matrix $\mathbf{A}$

Initialization:

2: Iteration: $t = 0$

3: Initial residual: $\mathbf{r}^0 = \mathbf{y}$

4: Initial support of selected indices: $\mathbf{S} = \phi$

5: Initialize $\mathbf{z}^0$ as an empty vector

Endmember selection:

6: **while** t is smaller than the maximum iteration number
    //forward step

7:      $j = \arg\max_{i \notin \mathbf{S}} \mathbf{a}_i^T \mathbf{r}^t$; $\{\mathbf{a}_i$ is the i-th column of $\mathbf{A}\}$

8:      Set $\mathbf{S} = \mathbf{S} \bigcup j$

9:      Update solution: $\mathbf{z}^{t+1} \leftarrow (\mathbf{A_S}^T \mathbf{A_S})^{-1} \mathbf{A_S}^T \mathbf{y}$; $\{\mathbf{A_S}$ means the matrix containing the columns of $\mathbf{A}$ having the indices from $\mathbf{S}\}$

10:     Set $\delta^{t+1} = \mathbf{q}(\mathbf{z}^t; \mathbf{S}/j) - \mathbf{q}(\mathbf{z}^{t+1}; \mathbf{S})$; $\{\mathbf{q}(\mathbf{z}; \mathbf{S}) = \frac{1}{2}\|\mathbf{A_S z} - \mathbf{y}\|_2^2\}$

11:     **if** $\delta^{t+1} \leq \epsilon$, set $\mathbf{S} = \mathbf{S}/j$ and **break**; $\{\epsilon$ is the predefined error tolerance$\}$

12:     Set $t = t + 1$
        //backward step

13:     **while (true)**

14:         $j = \arg\min_{i \in \mathbf{S}} \mathbf{q}(\widetilde{\mathbf{z}}^i; \mathbf{S}/i)$; $\{\widetilde{\mathbf{z}}^i = \arg\min_{\widetilde{\mathbf{z}} \in R^{|\mathbf{S}/i|}} \mathbf{q}(\widetilde{\mathbf{z}}; \mathbf{S}/i)\}$

15:         $d^- = \mathbf{q}(\widetilde{\mathbf{z}}^j; \mathbf{S}/j) - \mathbf{q}(\mathbf{z}^t; \mathbf{S})$

16:         $d^+ = \delta^t$

17:         **if** $d^- > \frac{1}{2} d^+$, **break**

18:         Set $t = t - 1$

19:         Set $\mathbf{S} = \mathbf{S}/j$

20:         Update solution: $\mathbf{z}^t \leftarrow (\mathbf{A_S}^T \mathbf{A_S})^{-1} \mathbf{A_S}^T \mathbf{y}$

21:     **end while**

22: **end while**

Inversion (abundance estimation):

23: Compute the abundance fractions for the scene using the original hyperspectral data and spectral library:
    $\mathbf{z} \leftarrow \arg\min_{\mathbf{Z}} \|\mathbf{A_S z} - \mathbf{y}\|_2$, subject to $\mathbf{z} \geq 0$.

Algorithm 2 Simultaneous forward-backward greedy algorithm for sparse unmixing of hyperspectral data

---

Preprocessing:

1:    Apply zero-mean normalization and l_2 unit-length normalization successively to the columns of hyperspectral data matrix $\mathbf{Y}$ and spectral library matrix $\mathbf{A}$

Initialization:

2:    Iteration: $t = 0$

3:    Initial residual: $\mathbf{R}^0 = \mathbf{Y}$

4:    Initial support of selected indices: $\mathbf{S} = \phi$

5:    Initialize $\mathbf{Z}^0$ as an empty matrix

Endmember selection:

6:    **while** t is smaller than the maximum iteration number
        //forward step

7:        $j = \arg\max_{i \notin \mathbf{S}} \|(\mathbf{R}^t)^T \mathbf{a}_i\|_p$; $\{p$ could be 2 or $\infty\}$

8:        Set $\mathbf{S} = \mathbf{S} \bigcup j$

9:        Update solution: $\mathbf{Z}^{t+1} \leftarrow (\mathbf{A_S}^T \mathbf{A_S})^{-1} \mathbf{A_S}^T \mathbf{Y}$

10:       Set $\delta^{t+1} = \mathbf{Q}(\mathbf{Z}^t; \mathbf{S}/j) - \mathbf{Q}(\mathbf{Z}^{t+1}; \mathbf{S})$;
          $\{\mathbf{Q}(\mathbf{Z}; \mathbf{S}) = \frac{1}{2}\|\mathbf{A_S}\mathbf{Z} - \mathbf{Y}\|_F^2\}$

11:       **if** $\delta^{t+1} \leq \epsilon$, set $\mathbf{S} = \mathbf{S}/j$ and **break**; $\{\epsilon$ is the predefined error tolerance$\}$

12:       Set $t = t + 1$
          //backward step

13:       **while (true)**

14:           $j = \arg\min_{i \in \mathbf{S}} \mathbf{Q}(\widetilde{\mathbf{Z}}^i; \mathbf{S}/i)$;
              $\{\widetilde{\mathbf{Z}}^i = \arg\min_{\widetilde{\mathbf{Z}} \in R^{|\mathbf{S}/i| \times K}} \mathbf{Q}(\widetilde{\mathbf{Z}}; \mathbf{S}/i)\}$

15:           $d^- = \mathbf{Q}(\widetilde{\mathbf{Z}}^j; \mathbf{S}/j) - \mathbf{Q}(\mathbf{Z}^t; \mathbf{S})$

16:           $d^+ = \delta^t$

17:           **if** $d^- > \frac{1}{2}d^+$, **break**

18:           Set $t = t - 1$

19:           Set $\mathbf{S} = \mathbf{S}/j$

20:           Update solution: $\mathbf{Z}^t \leftarrow (\mathbf{A_S}^T \mathbf{A_S})^{-1} \mathbf{A_S}^T \mathbf{Y}$

21:       **end while**

22:   **end while**

Inversion (abundance estimation):

23:   Compute the abundance fractions for the scene using the original hyperspectral data and spectral library:
      $\mathbf{Z} \leftarrow \arg\min_{\mathbf{Z}} \|\mathbf{A_S}\mathbf{Z} - \mathbf{Y}\|_F$, subject to $\mathbf{Z} \geq 0$.

---

[35] or relax it to a convex optimization problem [23]. Here, we extend FoBa displayed in Algorithm 1 to the simultaneous forward-backward greedy algorithm (SFoBa), which is shown in Algorithm 2. Similarly, we use $\mathbf{Z}$ instead of $\mathbf{X}$ to update the solution. $\mathbf{Z}$ could be considered as the rows of $\mathbf{X}$ that have the indices of $\mathbf{S}$. The difference is that $\mathbf{X}$ is row-sparse but $\mathbf{Z}$ is not.

The difference between FoBa and SFoBa is that FoBa processes each pixel in the hyperspectral data separately but

SFoBa deals with the whole hyperspectral data simultaneously. Step 7 is the greedy selection step by maximizing the absolute correlations ($p = 2$ or ∞) between the hyperspectral data and the selected member. Small value of p promotes the selection of potential endmember that has a strong overall contribution to the hyperspectral data. In contrast, larger value of p will favor the member most correlated to one mixed pixel. In our implementations, we consider both cases. Note when estimating the abundances, though the matrix $\mathbf{Z}$ is not row-sparse, each column of $\mathbf{Z}$ can still be sparse as the number of endmembers present in a mixed pixel is much less than the number of endmembers present in the whole image. Thus, we can add the l_1 norm regularizer to the object function in step 23 and obtain a sparser solution.

A drawback of the simultaneous greedy algorithms such as SOMP and SFoBa is that when the number of endmembers present in the hyperspectral data is very large, they become unstable and tend to be trapped into the local minima. Here we use a block-processing strategy to sidestep this obstacle. Specifically, we divide the whole hyperspectral image into several blocks. Then in each block the greedy algorithm will pick several potential endmembers from the spectral library. Finally, the abundances are computed using the whole hyperspectral data and potential endmembers picked associated with all the blocks.

The intuition behind this strategy is threefold: i) the number of endmembers present in one block is likely to be smaller than that in the whole image and the simultaneous greedy algorithm is more effective when the number of members that construct the input signals is not large [35]; ii) the greedy algorithm can further utilize the spatial-contextual information within the block; iii) one endmember present in one mixed pixel is probably also present in the others (either neighborly or non-neighborly) of the whole image, thus we can use the potential endmembers picked from all the blocks to reconstruct the hyperspectral image.

B. Regularized Simultaneous Forward-backward Greedy Algorithm

The second prior knowledge we consider is the spatial-contextual information. Till now, the hyperspectral data are not treated as an image, but as unordered listings of sampled spectral measurements collected in a matrix (namely the hyperspectral data matrix $\mathbf{Y}$). Indeed, there exists spatial correlation between image features [22]. A total variation (TV) spatial regularizer is added into the sparse unmixing model in [22] and optimistic improvement has been obtained in the abundance estimation. In this paper, we consider incorporating the spatial-contextual information into the simultaneous forward-backward greedy algorithm, but in a different way. In the t-th iteration of the SOMP algorithm [35] and SFoBa, after an additional member is selected from the spectral library, we need to solve a least squares problem to update the solution:

$$\mathbf{Z}^t \leftarrow \arg \min_{\mathbf{Z} \in R^{|\mathbf{S}| \times K}} \mathbf{Q}(\mathbf{Z}; \mathbf{S}) = \frac{1}{2} \|\mathbf{A_S Z} - \mathbf{Y}\|_F^2 \tag{7}$$

where $\mathbf{S}$ is a set of indices of selected members in the spectral library, $\mathbf{A_S}$ means the matrix containing the columns of $\mathbf{A}$ having the indices from $\mathbf{S}$. Such an updating procedure has two advantages when used in the field of sparse signal reconstruction: on one hand the input signals ($\mathbf{Y}$) will be reconstructed as well as possible using the selected members from the library; on the other hand, the residual $\mathbf{R}^t = \mathbf{Y} - \mathbf{A_S Z}^t$ will be orthogonal to the column space

of $\mathbf{A_S}$ and thus a member from the library will be selected at most once. However, when used in sparse unmixing, this kind of updating does not consider the spatial-contextual coherence of the abundance matrix and the structure of the residual matrix could be destroyed. Specifically, consider the situation in Fig. 2. Suppose $\mathbf{y}$ is the spectral vector of a mixed pixel and it is constituted by two endmembers $\mathbf{a}_1$ and $\mathbf{a}_2$. $\mathbf{a}_3$ is a member of the spectral library that is absent in the mixed pixel. In the first iteration, as $\mathbf{a}_1$ is closest to $\mathbf{y}$, it will be selected. Then the residual vector $\mathbf{r}$, as illustrated in Fig. 2, will be computed: $\mathbf{r} = \mathbf{y} - \mathbf{p}$ where the vector $\mathbf{p}$ is the orthogonal projection of $\mathbf{y}$ on $\mathbf{a}_1$. However, in the next iteration, $\mathbf{a}_3$ rather than $\mathbf{a}_2$ will be selected. In this case, the orthogonality is not beneficial to the sparse unmixing. Indeed, we would like the estimated residual $\mathbf{r}$ to be as close to the *actual residual* $\mathbf{a}_2$ as possible. Thus we need to update the solution more accurately in each iteration.
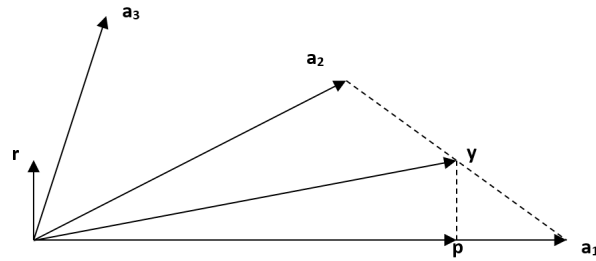

Fig. 2. Illustration of why orthogonality is not always good.

To tackle this problem, we consider the spatial-contextual coherence in the hyperspectral data which will depend on the spectral similarities between a pixel and its neighbouring pixels. This can be implemented by defining a quadratic function $\mathbf{G}$ with respect to the abundances:

$$\mathbf{G}(\mathbf{X}) = \frac{1}{2}\sum_{i=1}^{K}\sum_{j\in\eta}\|\mathbf{x}_i - \mathbf{x}_j\|_2^2 \mathbf{W}_{ij} \tag{8}$$

where η denotes the neighborhood of the i-th pixel, $\mathbf{x}_i$ is the i-th column of $\mathbf{X}$, $\mathbf{W}_{ij}$ is the weight matrix measuring the similarity between the i-th pixel and the j-th pixel. Specifically, if the j-th pixel is in the neighborhood of the i-th pixel, then the weight is assigned as

$$\mathbf{W}_{ij} = e^{-\frac{\|\mathbf{y}_i - \mathbf{y}_j\|_2^2}{\sigma}} \tag{9}$$

where σ is the density parameter controlling the width of the Guassian. It is obvious that if $\mathbf{y}_i$ and $\mathbf{y}_j$ (in a neighbourhood) are similar, $\mathbf{W}_{ij}$ will be relatively big. In this paper, the neighborhood η in Eq. (8) is set to be a $(2r + 1) \times (2r + 1)$ square region ($r = 2$) and the σ in Eq. (9) is set to 2/3. Indeed, we have tested these two parameters on synthetic data set (the same one used in the synthetic experiment in Section IV) in situations of different endmembers and different noise levels using different values of r (i.e. 1, 2, 3, 4) and σ (i.e. 1/3, 2/3, 1, 3, 5) and find that the proposed algorithm attains constantly good performance when $r \geq 2$ and gets the overall

best performance when $r = 2$ and $\sigma = 2/3$. For convenience of deduction, we transform the $\mathbf{G}(\mathbf{X})$ in Eq. (8) as

$$
\begin{aligned}
\mathbf{G}(\mathbf{X}) &= \frac{1}{2} \sum_{i,j=1}^{K} \|\mathbf{x}_i - \mathbf{x}_j\|_2^2 \mathbf{W}_{ij} \\
&= \sum_{i=1}^{K} \mathbf{x}_i^T \mathbf{x}_i \mathbf{D}_{ii} - \sum_{i,j=1}^{K} \mathbf{x}_i^T \mathbf{x}_j \mathbf{W}_{ij} \\
&= \mathrm{Tr}(\mathbf{X}\mathbf{D}\mathbf{X}^T) - \mathrm{Tr}(\mathbf{X}\mathbf{W}\mathbf{X}^T) \\
&= \mathrm{Tr}(\mathbf{X}\mathbf{L}\mathbf{X}^T)
\end{aligned}
\tag{10}
$$

where $\mathrm{Tr}(\cdot)$ means the trace of a matrix, $\mathbf{D}$ is a diagonal matrix and $\mathbf{D}_{ii} = \sum_j \mathbf{W}_{ij}$, $\mathbf{L} = \mathbf{D} - \mathbf{W}$ is a positive semidefinite matrix. The first equality holds because $\mathbf{W}_{ij} = 0$ if $j \notin \eta$. Thus, when updating the solution in each iteration, instead of solving the plain least squares problem in Eq. (7), we solve the following quadratic problem:

$$
\begin{aligned}
\mathbf{Z}^t \leftarrow \arg \min_{\mathbf{Z} \in R^{|\mathbf{S}| \times K}} \mathbf{Q_r}(\mathbf{Z}; \mathbf{S}) &= \mathbf{Q}(\mathbf{Z}; \mathbf{S}) + \frac{\lambda}{2} \mathbf{G}(\mathbf{Z}) \\
&= \frac{1}{2} \|\mathbf{A_S}\mathbf{Z} - \mathbf{Y}\|_F^2 + \frac{\lambda}{2} \mathrm{Tr}(\mathbf{Z}\mathbf{L}\mathbf{Z}^T)
\end{aligned}
\tag{11}
$$

where $\lambda \geq 0$ is the regularization parameter.

Note $\mathbf{Q_r}(\mathbf{Z}; \mathbf{S})$ is a convex quadratic function, thus any local minimum will be its global minimum. Then, to get the global optimum of the problem in Eq. (11) we need to take the gradient of $\mathbf{Q_r}(\mathbf{Z}; \mathbf{S})$ with respect to $\mathbf{Z}$ and set it to 0:

$$
\nabla \mathbf{Q_r}(\mathbf{Z}; \mathbf{S}) = \mathbf{A_S}^T \mathbf{A_S}\mathbf{Z} - \mathbf{A_S}^T \mathbf{Y} + \lambda \mathbf{Z}\mathbf{L} = 0.
\tag{12}
$$

Eq. (12) is actually a Lyapunov equation [48] which has the following form:

$$
\mathbf{C}_1 \mathbf{Z} + \mathbf{Z}\mathbf{C}_2 = \mathbf{C}_3
\tag{13}
$$

where $\mathbf{C}_1 = \mathbf{A_S}^T \mathbf{A_S}$, $\mathbf{C}_2 = \lambda \mathbf{L}$ and $\mathbf{C}_3 = \mathbf{A_S}^T \mathbf{Y}$. Considering the computational complexity and space complexity, we choose the method proposed in [48] to solve the Lyapunov equation. The eigenvalue decompositions of symmetric matrices $\mathbf{C}_1$ and $\mathbf{C}_2$ can be obtained as:

$$
\mathbf{U}^{-1}\mathbf{C}_1\mathbf{U} = \begin{bmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_{k'} \end{bmatrix},
$$

$$
\mathbf{V}^{-1}\mathbf{C}_2\mathbf{V} = \begin{bmatrix} \beta_1 & & \\ & \ddots & \\ & & \beta_K \end{bmatrix}
$$

where $k' \leq k$ is the cardinality of set $\mathbf{S}$, α_i $(i = 1, \ldots, k')$ are the eigenvalues of $\mathbf{C}_1$, β_i $(i = 1, \ldots, K)$ are the eigenvalues of $\mathbf{C}_2$, both $\mathbf{U}$ and $\mathbf{V}$ are orthogonal matrices. Then the solution [48] of Eq. (13) is

$$
\mathbf{Z} = \mathbf{U}\widetilde{\mathbf{Z}}\mathbf{V}^{-1}
\tag{14}
$$

where $\widetilde{\mathbf{Z}}_{ij} = \frac{\widetilde{\mathbf{C}}_{ij}}{\alpha_i + \beta_j}$ $(i = 1, \ldots, k'; j = 1, \ldots, K)$ and $\widetilde{\mathbf{C}} = \mathbf{U}^{-1}\mathbf{C}_3\mathbf{V}$. As $\mathbf{U}$ and $\mathbf{V}$ are orthogonal matrices, computation of their inversions can be replaced by computing their transposes when programming. When the λ in Eq. (12) is set to 0, the Lyapunov equation degrades to the normal equation and a plain least squares solution is obtained.

Finally we can get our regularized simultaneous forward-backward greedy algorithm (RSFoBa) which is summarized in Algorithm 3. Indeed, the relationship among the Algorithms 1-3 is that SFoBa is a special case of RSFoBa and FoBa is a special case of SFoBa. The proposed algorithm is inspired by SOMP [35] and FoBa [36], nevertheless they are quite different. Firstly, SOMP is a forward greedy algorithm but RSFoBa has backward step. Secondly, SOMP only considers the case of $p = 2$ in the member selection step but we also consider $p = \infty$. Thirdly, FoBa processes each pixel individually but RSFoBa deals with the whole image simultaneously. Fourthly and most importantly, RSFoBa exploits the spatial-contextual information in the hyperspectral image but SOMP and FoBa do not. Thus, RSFoBa will be more suitable for sparse unmixing because it not only has the advantages of SOMP and FoBa, but also exploits the characteristics of the hyperspectral image for sparse unmixing.

C. Theoretical Analysis

Here, we give a convergence theorem for the RSFoBa algorithm.

Lemma 3.1: In Algorithm 3, $d^+ \geq 0$ and $d^- \geq 0$ hold.

Proof: For $\forall j \in \mathbf{S}$, the following inequality holds:

$$\min_{\mathbf{Z} \in R^{|\mathbf{S}| \times K}} \mathbf{Q_r}(\mathbf{Z}; \mathbf{S}) \leq \min_{\mathbf{Z} \in R^{|\mathbf{S}| \times K}, \mathbf{Z}_{j,:}=0} \mathbf{Q_r}(\mathbf{Z}; \mathbf{S})$$

$$= \min_{\mathbf{Z} \in R^{|\mathbf{S}/j| \times K}} \mathbf{Q_r}(\mathbf{Z}; \mathbf{S}/j)$$

where $\mathbf{Z}_{j,:}$ denotes the j-th row of matrix $\mathbf{Z}$. Thus, according to the definitions of d^+ and d^-, they will be no less than 0. □

Theorem 3.1: The total number of main iterations for the RSFoBa algorithm to terminate is no more than $1 + 2\mathbf{Q_r}(\mathbf{0})/\epsilon$.

Proof: Suppose RSFoBa will take q' backward steps after taking q $(q' \leq q)$ forward steps. After q forward steps, the decrease of $\mathbf{Q_r}$ is

$$D^+ = \sum_{i=1}^{q} d_i^+ \tag{15}$$

where $d_i^+ = \delta^i$, $i = 1, \ldots, q$. Suppose in the backward step corresponding to the i-th forward step, the increase of $\mathbf{Q_r}$ is d_i^-, $i = q, q-1, \ldots, q-q'+1$. After q' backward steps, the increase of $\mathbf{Q_r}$ is

$$D^- = \sum_{i=q-q'+1}^{q} d_i^- \tag{16}$$

Algorithm 3 Regularized simultaneous forward-backward greedy algorithm for sparse unmixing of hyperspectral data

---

Preprocessing:

1: Apply zero-mean normalization and l_2 unit-length normalization successively to the columns of hyperspectral data matrix $\mathbf{Y}$ and spectral library matrix $\mathbf{A}$

Initialization:

2: Iteration: $t = 0$

3: Initial residual: $\mathbf{R}^0 = \mathbf{Y}$

4: Initial support of selected indices: $\mathbf{S} = \phi$

5: Initialize $\mathbf{Z}^0$ as an empty matrix

Endmember selection:

6: **while** t is smaller than the maximum iteration number
   //forward step

7: $j = \arg\max_{i \notin \mathbf{S}} \|(\mathbf{R}^t)^T \mathbf{a}_i\|_p$; {$p$ could be 2 or ∞}

8: Set $\mathbf{S} = \mathbf{S} \bigcup j$

9: Compute $\mathbf{Z}^{t+1}$ via solving Eq. (12)

10: Set $\delta^{t+1} = \mathbf{Q_r}(\mathbf{Z}^t; \mathbf{S}/j) - \mathbf{Q_r}(\mathbf{Z}^{t+1}; \mathbf{S})$

11: **if** $\delta^{t+1} \leq \epsilon$, set $\mathbf{S} = \mathbf{S}/j$ and **break**; {ϵ is the predefined error tolerance}

12: Set $t = t + 1$
   //backward step

13: **while (true)**

14:   $j = \arg\min_{i \in \mathbf{S}} \mathbf{Q_r}(\widetilde{\mathbf{Z}}^i; \mathbf{S}/i)$;
    $\{\widetilde{\mathbf{Z}}^i = \arg\min_{\widetilde{\mathbf{Z}} \in R^{|\mathbf{S}/i| \times K}} \mathbf{Q_r}(\widetilde{\mathbf{Z}}; \mathbf{S}/i)\}$

15:   $d^- = \mathbf{Q_r}(\widetilde{\mathbf{Z}}^j; \mathbf{S}/j) - \mathbf{Q_r}(\mathbf{Z}^t; \mathbf{S})$

16:   $d^+ = \delta^t$

17:   **if** $d^- > \frac{1}{2}d^+$, **break**

18:   Set $t = t - 1$

19:   Set $\mathbf{S} = \mathbf{S}/j$

20:   Compute $\mathbf{Z}^t$ via solving Eq. (12)

21: **end while**

22: **end while**

Inversion (abundance estimation):

23: Compute the abundance fractions for the scene using the original hyperspectral data and spectral library:
  $\mathbf{Z} \leftarrow \arg\min_{\mathbf{Z}} \|\mathbf{A_S}\mathbf{Z} - \mathbf{Y}\|_F$, subject to $\mathbf{Z} \geq 0$.

Then the actual decrease of $\mathbf{Q_r}$ after q forward steps and q' backward steps is

$$
\begin{aligned}
D &= D^+ - D^- \\
&= \sum_{i=1}^{q} d_i^+ - \sum_{i=q-q'+1}^{q} d_i^- \\
&\geq \sum_{i=1}^{q} \frac{1}{2} d_i^+ \\
&> \frac{q\epsilon}{2}.
\end{aligned}
\tag{17}
$$

The first inequality holds because $d_i^- \leq \frac{1}{2}d_i^+$; the second inequality holds because $d_i^+ = \delta^i > \epsilon$. Thus, if we take q forward steps, then no matter how many backward steps are performed, $\mathbf{Q_r}$ will decrease by at least an amount of $q\epsilon/2$. It follows that the algorithm terminates after no more than $2\mathbf{Q_r}(\mathbf{0})/\epsilon$ iterations. $\square$

It is worth mentioning that the number of iterations present in Theorem 3.1 is not tight. However, we can see that the RSFoBa can make progress in every main iteration.

IV. EXPERIMENT

In this section, we use both synthetic data and real-world data to test the performance of the proposed algorithm. The RSFoBa algorithm is compared with three greedy algorithms (i.e. OMP [27], SOMP [35] and SSP [49, 50]) and two convex relaxation methods (i.e. SUnSAL [16] and SUnSAL-TV [22]). We also explore whether the backward step, joint sparsity and spatial-contextual coherence regularizer are beneficial to the proposed algorithm in the synthetic experiment by comparing RSFoBa with its special cases. As mentioned earlier, when the spectral library is over-complete, the backward greedy algorithm behaves worse and implements much slower than the forward greedy algorithm and convex relaxation method, thus we do not include it in our experiment. In step 7 of the RSFoBa algorithm, both the l_2 norm and the l_∞ norm are considered in our experiments. To distinguish the two cases, we denote them by RSFoBa ($p = 2$) and RSFoBa ($p = \infty$), respectively. We also consider combining the RSFoBa with the SUnSAL, which means that the SUnSAL rather than the nonnegative least squares technique is used to estimate the abundances in the final step of RSFoBa. We denote this combination as RSFoBa-SUnSAL. The maximum iteration number and the error tolerance ϵ of the RSFoBa algorithm will determine when the main iteration of the RSFoBa algorithm terminates. In all the experiments, the maximum iteration number is set to 20, which is enough for the RSFoBa algorithm to terminate by satisfying the error tolerance. As the value of $\mathbf{Q_r}$ for a block is proportional to the number of pixels in this block, the error tolerance ϵ should be set to $\epsilon' \times K_b$, where ϵ' is the scaled error tolerance unrelated to the block size, K_b is the number of pixels in a block and $K_b = K$ if the block-processing strategy is not adopted. It is obvious that smaller ϵ' tends to make RSFoBa retain more spectral signatures from the spectral library. However, if the ϵ' is set too small, the RSFoBa algorithm will retain too many non-existing spectral signatures, which will have a negative effect on the abundance estimation; if the ϵ' is set too large, the RSFoBa algorithm could miss some actual endmembers. Thus, an appropriate ϵ' is necessary. We have tested this parameter on synthetic data set (the same one used in synthetic experiment) in situations of different endmembers and different noise levels using different values of ϵ': 0.05, 0.01, 0.005 and 0.001. We find that when $\epsilon' = 0.05$, RSFoBa will usually miss the actual endmembers, which indicates this value of ϵ' is too large; when $\epsilon' = 0.001$, RSFoBa will select too many non-existing spectral signatures; $\epsilon' = 0.01$ leads to a slightly better performance of RSFoBa than $\epsilon' = 0.005$ in most cases. Thus, in all the experiments, the error tolerance ϵ in step 11 of the RSFoBa algorithm is set to $0.01 \times K_b$. Then, the parameters for the RSFoBa algorithm to be tuned are the block size and the regularization parameter λ in Eq. (12), both of which will be discussed in Section IV-D. All the considered algorithms have taken into account the abundance nonnegativity constraint. The preprocessing procure adopted by RSFoBa is also applied to OMP, SOMP and SSP to improve their performances. The block-processing

strategy is applied to the SOMP and SSP to avoid missing the actual endmembers when the number of endmembers in the hyperspectral data is large. The nonnegative least squares problem in the abundance estimation step of the four greedy algorithms can be solved using the active set method. Specifically, we use the *lsqnonneg* function in MATLAB to solve it. The TV regularizer used in SUnSAL-TV is the non-isotropic one. All the parameters of the considered algorithms are tuned to their best performances. Complexity analysis comes at the end of this section.

A. Evaluation with Synthetic Data

In the synthetic data experiment, we evaluate the performances of the sparse unmixing algorithms in situations of different noise types, different signal-to-noise ratios (SNR $\equiv 10 \log_{10} \frac{\|\mathbf{A}\mathbf{x}\|_2^2}{\|\mathbf{n}\|_2^2}$) of noise and different endmember numbers. The root mean square error (RMSE) is used to evaluate the abundance estimations. For the i-th endmember, RMSE is defined as

$$\mathrm{RMSE}_i = \sqrt{\frac{1}{K} \sum_{j=1}^{K} (\mathbf{X}_{ij} - \hat{\mathbf{X}}_{ij})^2}. \tag{18}$$

Here, $\mathbf{X}$ denotes the true abundances and $\hat{\mathbf{X}}$ represents the estimated ones. The mean value of all the endmembers' RMSEs will be computed. Generally speaking, the smaller the RMSE is, the more the estimation approximates the truth. The synthetic hyperspectral image we used is homogeneous with rich spatial-contextual information. The Chapter 1 of the United States Geological Survey (USGS) [51] digital spectral library (splib06a) is used in this experiment. The reflectance values of 498 materials are measured for 224 spectral bands distributed uniformly in the interval $0.4 - 2.5 \ \mu m$ ($\mathbf{A} \in R^{224 \times 498}$). Fifteen spectral signatures are chosen from the library to generate our synthetic data. Fig. 3 shows five example endmember signatures used for all the following experiments. The other ten endmembers that are not displayed in the figure include Neodymium Oxide GDS34, Monazite HS255.3B, Samarium Oxide GDS36, Pigeonite HS199.3B, Meionite WS700.HLsep, Spodumene HS210.3B, Labradorite HS17.3B, Grossular WS484, Zoisite HS347.3B and Wollastonite HS348.3B, which are used to evaluate the performances of the sparse unmixing algorithms in situations of different endmember numbers.

The synthetic data are created as follows:

1) Divide the scene, whose size is $z^2 \times z^2$ ($z = 8$), into $z \times z$ regions. Initialize each region with the same type of ground cover, randomly selected from the endmember class. The endmember number is P. The size of spectral signatures matrix $\mathbf{W}$ is $L \times P$ ($L = 224$).

2) Generate mixed pixels through a simple $(z + 1) \times (z + 1)$ spatial low-pass filter.

3) Replace all the pixels in which the abundance of a single endmember is larger than 70% with a mixture made up of this endmember and its next endmember (the abundances of these two endmembers both equal 50%) so as to further remove pure pixels and represent the sparseness of abundances at the same time; After these three steps, we obtain the distribution of P endmembers in the scene and abundance values are stored in $\mathbf{H}$ with a size of $P \times K$ ($K = z^2 \times z^2$).

4) Use linear spectral mixing model $\mathbf{Y} = \mathbf{W} \times \mathbf{H}$ to generate hyperspectral data, add Gaussian white noise or
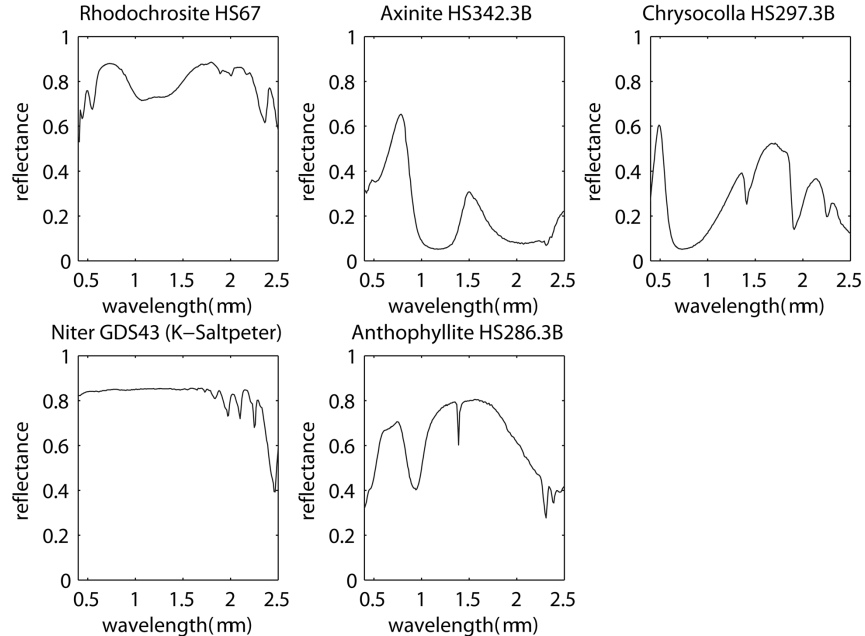
Fig. 3. Five example spectral signatures from the USGS used in our synthetic data experiments. The title of each subimage denotes the mineral corresponding to the signature.

correlated noise[2] with specific SNR at the same time. The size of hyperspectral data $\mathbf{Y}$ is $L \times K$.

Note that no pure pixel exists in this synthetic data set. Because this data set is homogeneous with rich spatial-contextual information, it can simulate the real scenario.

We first analyse without considering the RSFoBa-SUnSAL. Fig. 4 shows results obtained on the synthetic data with white noise as function of SNR when the endmember number is 5. The RMSEs of all the algorithms decrease as the SNR increases. We can find that RSFoBa, SOMP, SSP and SUnSAL-TV behave better than SUnSAL and OMP, which indicates that the use of spatial information can improve the performances of the sparse unmixing algorithms. Besides, it can be seen that RSFoBa outperforms the other algorithms. Fig. 5 shows the results obtained on this data set with 30 dB white noise as function of endmember number. As shown, the performances of all the algorithms tend to degrade as the endmember number increases. This indicates that the sparsity of the solution can alleviate the sparse regression limitations linked to the usually high correlation of the spectral library. In most cases, RSFoBa behaves best among all the algorithms. It is worth mentioning that in both Fig. 4 and Fig. 5, the performances of RSFoBa ($p = 2$) and RSFoBa ($p = \infty$) are different: RSFoBa ($p = \infty$) behaves better than RSFoBa ($p = 2$) in most cases. Thus, when using RSFoBa for sparse unmixing, the l_∞ norm could be a better choice than the l_2 norm.

[2]The Gaussian white noise is generated using the *awgn* function in MATLAB. The correlated noise is generated using the *correlatedGaussianNoise* function that is available online: http://www.mathworks.com/matlabcentral/fileexchange/21156-correlated-gaussian-noise/content/correlatedGaussianNoise.m. The correlation matrix is set as default.
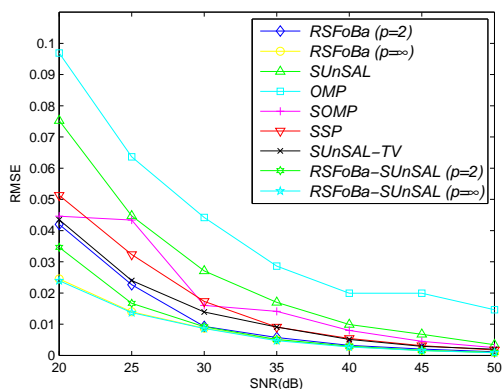
Fig. 4. Results on the synthetic data with white noise when the endmember number is 5: RMSE as function of SNR.
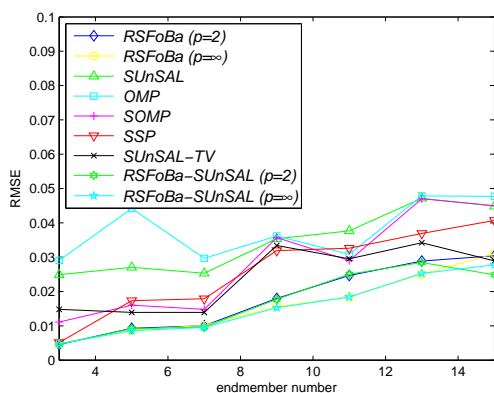

Fig. 5. Results on the synthetic data with 30 dB white noise: RMSE as function of endmember number.

Fig. 6 and Fig. 7 show the results obtained on the synthetic data affected by correlated noise as functions of SNR and endmember number, respectively. Since the noise in the real hyperspectral images is usually correlated, this case is closer to the practical ones. As shown in Fig. 6, the performances of RSFoBa ($p = \infty$), SOMP and SSP are similar to those in the white noise cases. But RSFoBa ($p = 2$), SUnSAL, OMP and SUnSAL-TV behave worse when the data are corrupted by the correlated noise. The results in Fig. 7 are similar with those in Fig. 5. We can find that the RSFoBa has an overall better performance than the other considered algorithms. Besides, we can also observe that RSFoBa ($p = \infty$) behaves better than RSFoBa ($p = 2$) in most cases, which is in accordance with the results when the noise is white.

In order to validate whether the proposed algorithm significantly outperforms the others, we also conduct significant tests on the results obtained above. Specifically, there are totally 28 RMSEs for each sparse unmixing algorithm obtained in situations of different endmember numbers, different noise types and different noise levels. As the samples (RMSEs) for an algorithm are obtained in correlated situations rather than randomly and independently,
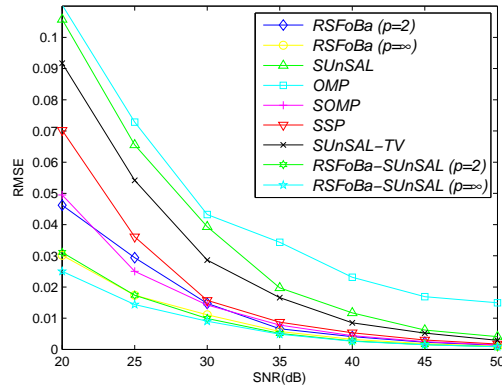
Fig. 6. Results on the synthetic data with correlated noise when the endmember number is 5: RMSE as function of SNR.
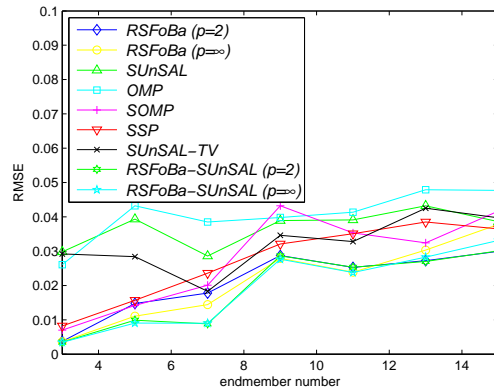

Fig. 7. Results on the synthetic data with 30 dB correlated noise: RMSE as function of endmember number.

it is inappropriate to use the independent Student t test. Now consider two arbitrary sparse unmixing algorithms, their RMSEs can be divided into 28 pairs corresponding to the 28 different situations. Then the two RMSEs in a pair are obtained in the same situation by the two algorithms. This line of reasoning leads us to consider using the paired-samples t test (or matched-samples t test) [52]. Indeed, researchers typically use the paired-samples t test to analyze paired data, which essentially is a one-sample Student t test performed on difference scores [53]. The paired-samples t test can reduce the effects of confounders which in our case are the endmember number, noise type and noise level. Take the comparison between RSFoBa ($p = 2$) and OMP as an example, hypotheses for a one-sided paired-samples t test take the following form:

- H_0: there is no significant difference between the performances of RSFoBa ($p = 2$) and OMP.
- H_1: the performance of RSFoBa ($p = 2$) is significantly better than that of OMP.

Suppose μ is the mean of the differences between the paired RMSEs (subtract the RMSE obtained by RSFoBa ($p = 2$) from the RMSE obtained by OMP), the hypotheses above are equivalent to the following hypotheses:

TABLE I

RESULTS (AT $\alpha = 0.01$ SIGNIFICANCE LEVEL) AND p-VALUES OBTAINED BY CONDUCTING PAIRED-SAMPLES t TESTS BETWEEN RSFoBA ($p = 2$) AND THE OTHER ALGORITHMS

Algorithm	OMP	SOMP	SSP	SUnSAL	SUnSAL-TV
Result	reject H_0	reject H_0	reject H_0	reject H_0	reject H_0
p-value	1.61E-10	3.80E-05	1.42E-06	3.61E-08	8.40E-05

TABLE II

RESULTS (AT $\alpha = 0.01$ SIGNIFICANCE LEVEL) AND p-VALUES OBTAINED BY CONDUCTING PAIRED-SAMPLES t TESTS BETWEEN RSFoBA ($p = \infty$) AND THE OTHER ALGORITHMS

Algorithm	OMP	SOMP	SSP	SUnSAL	SUnSAL-TV
Result	reject H_0	reject H_0	reject H_0	reject H_0	reject H_0
p-value	4.19E-09	1.51E-06	1.43E-05	4.93E-07	5.89E-05

- H_0: $\mu = 0$
- H_1: $\mu > 0$

Similarly, we can formulate the hypotheses for the comparison of RSFoBA and other algorithms. Using the *ttest* function in MATLAB, we can get the results of the paired-samples t test easily, which are displayed in Tabs. I and II. It can be found that at the $\alpha = 0.01$ significance level, the results give evidence that the performances of RSFoBA ($p = 2$) and RSFoBA ($p = \infty$) are better than the other sparse unmixing algorithms.

Fig. 8 shows the number of potential endmembers retained from the original library by RSFoBA, SOMP and SSP. In all the cases, all the considered algorithms succeed to find all the actual endmembers. We can observe that RSFoBA can select the actual endmembers more accurately than SOMP and SSP, which can explain why RSFoBA behaves better than SOMP and SSP. Besides, RSFoBA ($p = \infty$) has an overall better performance than RSFoBA ($p = 2$).

Now we tend to the RSFoBa-SUnSAL. Note in the final step of the RSFoBa, a constrained least squares technique is applied to obtain the abundances of the selected potential endmembers. By doing so, we do not get the optimal sparse solution because i) the number of endmembers present in each mixed pixel is smaller than the number of endmembers in the whole hyperspectral image; ii) observed from Fig. 8, the numbers of endmembers extracted by the greedy algorithms are still larger than the actual number. The abundances obtained by RSFoBA is sparse because of the nonnegative constraint and the small subset of the library. Indeed, we can use the sub-library obtained by RSFoBA as input for the other sparse unmixing algorithms. In this way, we can expect that this kind of combination will be better than both methods. From Figs. 4-7, we can find that the combination of RSFoBA and SUnSAL outperforms both of them, which is expected.

In order to explore whether the incorporation of the backward step, spatial-contextual information and joint sparsity among all the pixels is beneficial to the proposed algorithm, we compare the RSFoBA algorithm with its
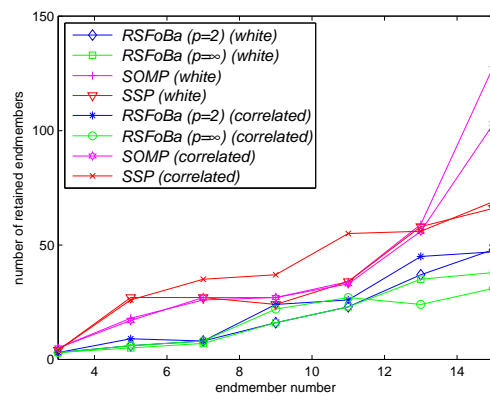
Fig. 8. Results obtained by RSFoBa, SOMP and SSP on the synthetic data with 30 dB white noise or correlated noise: the number of retained endmembers as function of endmember number.

special cases. Specifically, SFoBa is the special case of RSFoBa when the λ in Eq. (11) is set to 0; RSOMP is the special case of RSFoBa that removes the backward step and it can be also considered as incorporating the spatial-contextual coherence regularizer into the SOMP algorithm; the SOMP [35] and FoBa [36], which can be respectively considered as the special cases of RSOMP and SFoBa, are also considered. In order to make the differences among the RSFoBa algorithm and its special cases clearer, we display the comparison of them concerned with several properties in Tab. III. Fig. 9 and Fig. 10 show the results obtained by RSFoBa, SFoBa, RSOMP, FoBa and SOMP on the synthetic data with 30 dB white noise and correlated noise, respectively. We can see that in most cases, RSFoBa behaves better than SFoBa and RSOMP behaves better than SOMP, which indicates the spatial-contextual coherence regularizer can make the greedy algorithms more effective. The advantage of RSFoBa over SFoBa becomes more obvious as the endmember number gets larger, in which case the hyperspectral image becomes more difficult to unmix for all the algorithms. Besides, we can also find that SFoBa outperforms FoBa much, which suggests the importance of incorporating the joint sparsity among all the pixels into the greedy algorithm. Furthermore, in most cases, SFoBa behaves better than SOMP and RSFoBa behaves better than RSOMP, which indicates the necessity of combining the forward greedy algorithm and backward greedy algorithm in sparse unmixing.

B. Evaluation with Real Data 1

Now, we test the validity of the proposed algorithm on the real data. As the true abundance maps of the real-world data are not available, we resort to the classification maps of the real hyperspectral image to make a qualitative analysis. For visual comparison, we display the abundances estimated by SOMP, SSP, RSFoBa and SUnSAL-TV. Besides, we also report the sparsity of solution and the reconstruction error of the hyperspectral image obtained by each algorithm. Specifically, the sparsity is obtained by calculating the average number of nonzero entries in the abundance vectors estimated by each algorithm. The same as [23], we declare the fractional abundances

TABLE III

THE COMPARISON OF RSFOBA AND ITS SPECIAL CASES CONCERNED WITH WHETHER THE BACKWARD STEP, JOINT SPARSITY AND

CONTEXTUAL INFORMATION ARE CONSIDERED

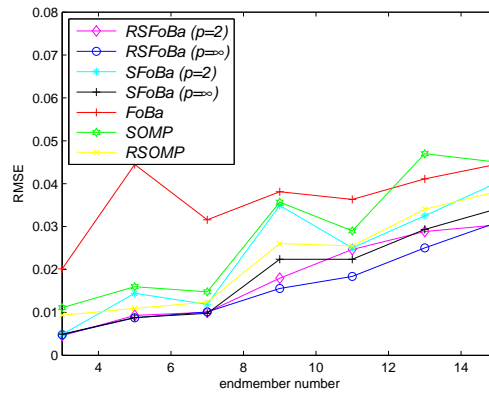| property	backward step	joint sparsity	contextual information
RSFoBa	✓	✓	✓
SFoBa	✓	✓	
RSOMP		✓	✓
FoBa	✓		
SOMP		✓	

Fig. 9. Results obtained by RSFoBa, SFoBa, FoBa, SOMP and RSOMP on the synthetic data with 30 dB white noise: RMSE as function of endmember number.
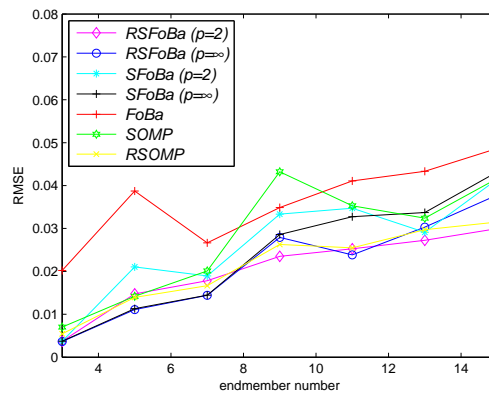

Fig. 10. Results obtained by RSFoBa, SFoBa, FoBa, SOMP and RSOMP on the synthetic data with 30 dB correlated noise: RMSE as function of endmember number.

larger than 0.001 as *nonzero abundances* to avoid counting negligible values. The reconstruction error is calculated using the original hyperspectral image and the one reconstructed by the actual endmembers and their corresponding abundances. The actual endmembers in the hyperspectral image can be known from the literatures. The reconstruction error calculated in this way can roughly evaluate the abundance estimation for the actual endmembers when no ground-truth is available [54]. It can be quantified by the RMSE between the original hyperspectral image and the reconstructed one [54], which is denoted by $\hat{\mathbf{Y}}$. For the i-th band image, RMSE is defined as

$$\text{RMSE}_i = \sqrt{\frac{1}{K}\sum_{j=1}^{K}(\mathbf{Y}_{ij} - \hat{\mathbf{Y}}_{ij})^2}. \tag{19}$$

The mean value of the RMSEs corresponding to all the bands will be computed. All the algorithms have been tuned to their best performances. If a sparse unmixing algorithm can get a sparse solution as well as low reconstruction error, it means that this algorithm can use the actual endmembers to explain the data effectively.

The first data set we used is the well-known AVIRIS Cuprite data set[3]. In our experiment, we use a 350×350 pixels subset with 188 spectral bands (low-SNR bands are removed). The spectral library used here is still the Chapter 1 of the USGS spectral library (contains 498 minerals) with the corresponding bands removed. The minerals map[4] which was produced by a Tricorder 3.3 software product is shown in Fig. 11. Note that the Tricorder map was produced in 1995, but the publicly available AVIRIS Cuprite data were collected in 1997. Thus, we only adopt the minerals map as a reference to make a qualitative analysis of the performances of different sparse unmixing methods.

All the greedy algorithms have adopted the block-processing strategy to get better performances. The block sizes of SOMP and SSP are both set empirically to 8. The block size set for RSFoBa is 15. The λ's in the RSFoBa algorithm are set to 0.1 ($p = 2$) and 0.05 ($p = \infty$), respectively. Fig. 12 shows a qualitative comparison among the fractional abundance maps of 3 highly materials (i.e. Alunite, Buddingtonite and Chalcedony) in the considered scene estimated by RSFoBa, SUnSAL-TV, SOMP and SSP. The classification maps of these materials produced by Tricorder software are also displayed. The abundance maps estimated by SUnSAL-TV shown in Fig. 12 are from [22]. Note that the Tricorder maps and the abundance maps estimated by the sparse unmixing algorithms are not exactly the same. This distinction results from the difference between the classification maps and the abundance maps: as classification maps, Tricorder maps consider each pixel in the hyperspectral image pure and classify it as member of a specific class correlated to the representative mineral in the pixel; unmixing is indeed a subpixel level classification and the abundances for a mixed pixel depend on the degree of presence of the mineral in the pixel. But, as shown in Fig. 12, the highest abundances estimated by the sparse unmixing algorithms generally correspond with those pixels classified as members of the respective class of materials. We can also find that the abundances estimated by RSFoBa are comparable or higher in the regions assigned to the respective minerals in comparison to the other considered algorithms. The sparsity of the abundance vectors and the reconstruction error obtained by

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

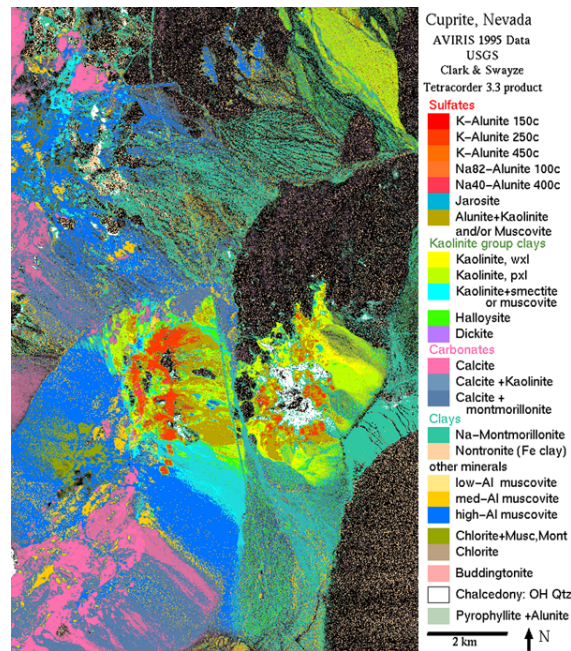[4]http://speclab.cr.usgs.gov/PAPERS/tetracorder/

Fig. 11. USGS map showing the distribution of different minerals in the Cuprite mining district in Nevada.

TABLE IV

EXPERIMENTAL RESULTS OBTAINED ON THE FIRST REAL DATA SET: SPARSITY OF THE ABUNDANCE VECTORS AND THE RECONSTRUCTION

ERRORS

Algorithms	RSFoBa ($p = 2$)	RSFoBa ($p = \infty$)	SOMP	SSP	SUnSAL-TV
Sparsity	13.65	13.11	14.98	14.85	17.13
RMSE	0.2305	0.2312	0.2394	0.2485	0.2359

each algorithm are displayed in Tab. IV. As the Cuprite site has been well understood mineralogically, the actual endmembers in the scene are determined by referring to the Tricorder map and the related literatures [55]. From Tab. IV, we can observe that RSFoBa can obtain the sparsest solutions and achieve the lowest reconstruction errors, which demonstrates the effectiveness of the proposed algorithm.

C. Evaluation with Real Data 2

The second data set we used is a subscene from the Washington, DC, mall data set[5] (shown in Fig. 13), which contains 150×150 pixels. After low-SNR bands are removed, only 191 bands remain. There are five targets of interest: water/shadows, grass, roof, tree and road. The spectral library used here contains two parts. The first part is the Chapter 6 of the USGS spectral library (splib06a) which contains 200 different spectral signatures of plants, vegetation communities, mixtures with vegetation, and microorganisms with the corresponding bands removed.

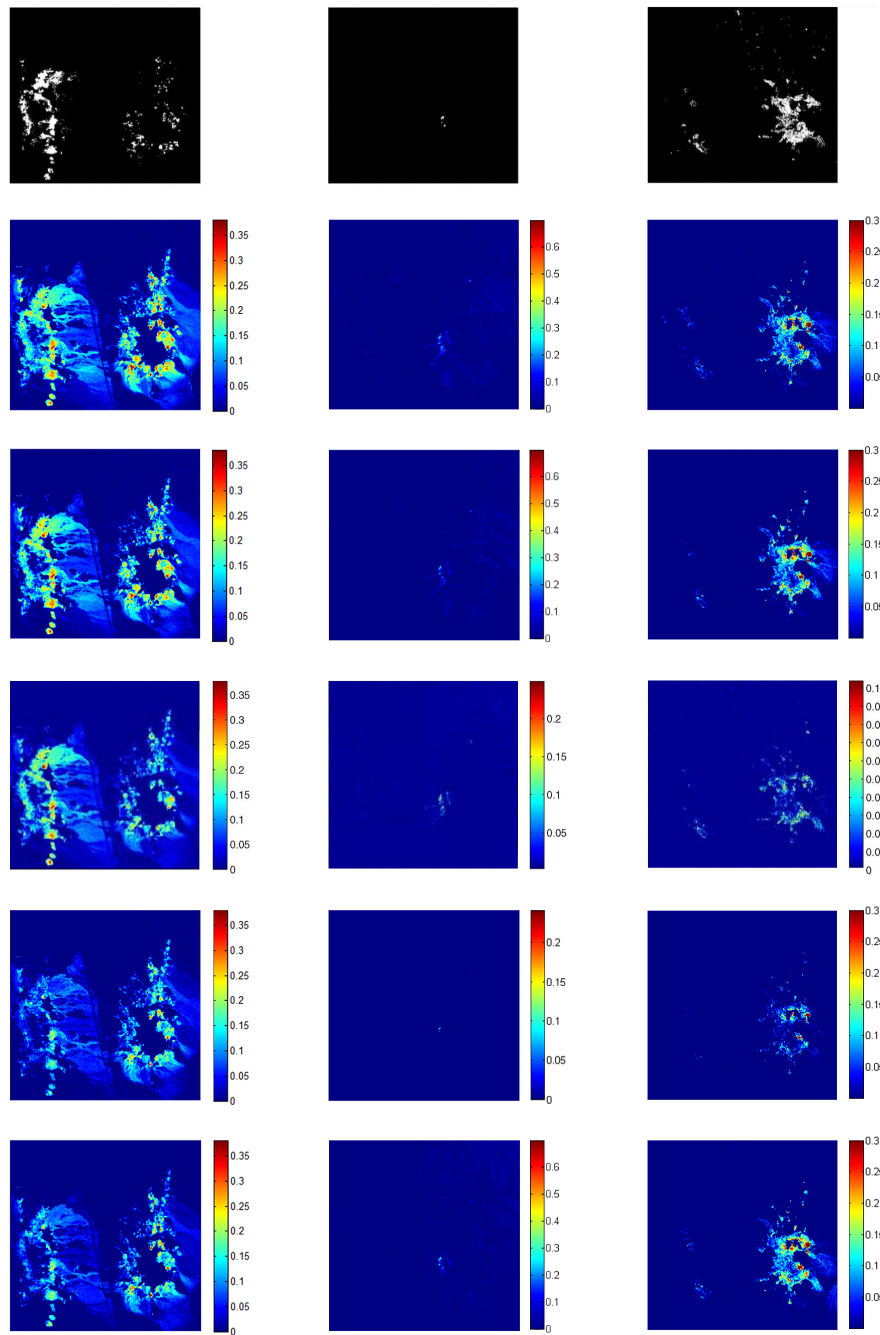[5]https://engineering.purdue.edu/ biehl/MultiSpec/hyperspectral.html

Fig. 12. The fractional abundance maps estimated by RSFoBa, SUnSAL-TV, SOMP, SSP and the distribution maps produced by Tricorder software for the 350×350 pixels subset of the AVIRIS Cuprite scene. From top row to bottom row are the maps produced or estimated by Tricorder software, RSFoBa ($p = \infty$), RSFoBa ($p = 2$), SUnSAL-TV, SOMP and SSP respectively. From left column to right column are the maps corresponding to Alunite, Buddingtonite and Chalcedony, respectively.

Fig. 13. Subscene extracted from Washington, DC, mall data set, band 80.

However, the spectral signatures in the USGS spectral library and the pixel signatures in the Washigton data were collected in the different regions, different time and different weather which will significantly affect the spectral signatures of the plant and vegetation. Besides, the USGS spectral library does not contain the signatures for the water, roof and road. Thus, we manually select several pixels in the hyperspectral scene for each single target of interest and compute the average signature of these signatures to represent the pure spectral signature of the target. Then, the second part of the spectral library we used is the 5 spectral signatures of the targets. The final spectral library contains 205 spectral signatures.

The block sizes of SOMP and SSP are both set empirically to 20. The block size set for RSFoBa is 30. The λ's in the RSFoBa algorithm are set to 3 ($p = 2$) and 0.3 ($p = \infty$), respectively. The λ and λ_{TV} in the SUnSAL-TV algorithm are set to 0.01 and 0.001, respectively. The abundance fractions for the five targets estimated by the different algorithms are displayed in Fig. 14. To make a qualitative evaluation of the performances of the sparse unmixing algorithms, we resort to the classification map [56] shown in Fig. 15. It can be found that the estimated abundance maps are very close to the distribution of the five materials in the classification map. All the algorithms can return the highest abundances for those pixels classified as members of the respective class of materials. It is worth noting that the fractional abundances estimated by RSFoBa are generally comparable or higher in the regions assigned to the respective target of interest in comparison to the SOMP, SSP and SUnSAL-TV. This observation is especially obvious for the abundances corresponding to the roof. Tab. V shows the sparsity of the abundance vectors and the reconstruction error obtained by each algorithm. Similarly, we can find that RSFoBa can obtain the sparsest solutions and lowest reconstruction errors among all the considered algorithms. Note that the reconstruction errors in this data set are much smaller than those in the first real data experiment. This is because the first real hyperspectral image is much more complex than the second one and the spectral library used in the first real data experiment is much larger and more correlated than that used in the second real data experiment, which means the first real data set is much more difficult to unmix for the sparse unmixing algorithms than the second real data set. Nevertheless, we can find that the RSFoBa is more effective for sparse unmixing than the other algorithms in both
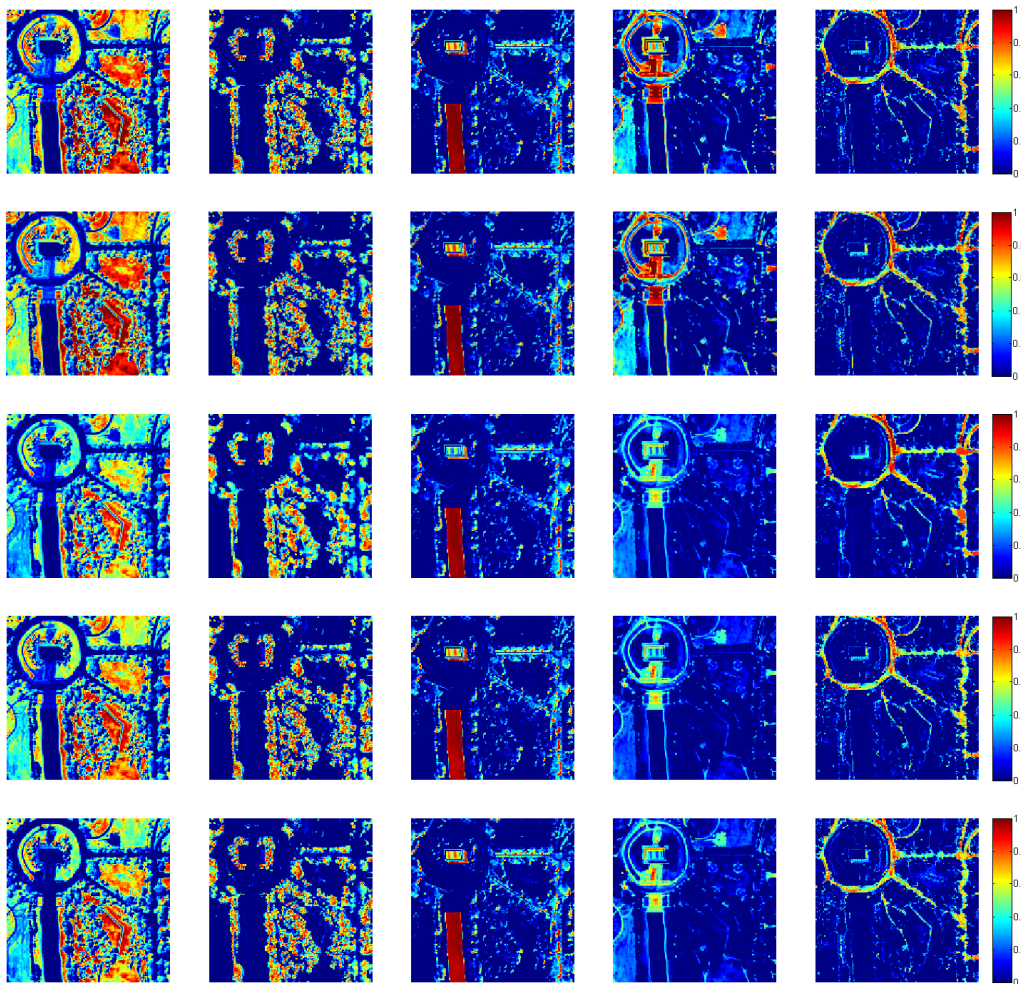
Fig. 14. Abundance maps estimated by RSFoBa, SUnSAL-TV, SOMP and SSP for the five targets in the Washington, DC, mall data set. From left column to right column are abundance maps of grass, tree, water/shadows, roof and road, respectively. From top row to bottom row are the abundance maps estimated by RSFoBa ($p = \infty$), RSFoBa ($p = 2$), SUnSAL-TV, SOMP and SSP respectively.

real experiments.

D. Discussion of Parameters Setting

As mentioned before, there remain two parameters to be set for the RSFoBa algorithm, i.e. the λ and the block size. The λ controls how much the spatial-contextual coherence regularization will affect the update of the solution in the RSFoBa algorithm. When $\lambda = 0$, the problem in Eq. (11) degrades to a plain least squares problem and we call the degraded algorithm SFoBa. As shown in Fig. 9 and Fig. 10, RSFoBa behaves better than SFoBa, especially when the endmember number becomes large. In the synthetic data experiment, the RSFoBa is tested using different values of λ when the block size is set properly: 0.0001, 0.0005, 0.001, 0.005, 0.01, 0.05, 0.1, 0.3, 0.5, 1, 3, 5, 10. We display the optimal λ's in situations of different endmember numbers in Tab. VI. Note that the optimal λ may
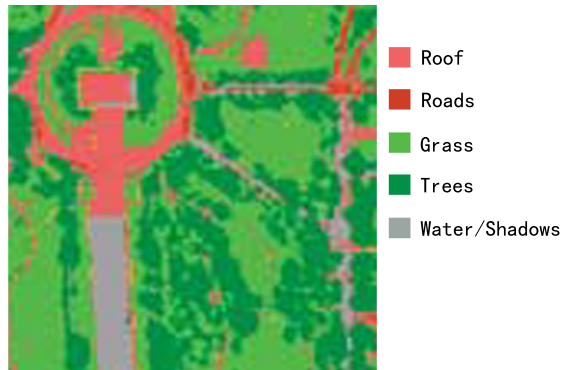
Fig. 15. Classification map of the subscene of Washington, DC, mall data set.

TABLE V

EXPERIMENTAL RESULTS OBTAINED ON THE SECOND REAL DATA SET: SPARSITY OF THE ABUNDANCE VECTORS AND THE

RECONSTRUCTION ERRORS

Algorithms	RSFoBa ($p = 2$)	RSFoBa ($p = \infty$)	SOMP	SSP	SUnSAL-TV
Sparsity	2.46	3.12	3.99	5.78	6.45
RMSE	0.0128	0.0139	0.0156	0.0190	0.0185

not be unique. Generally, we can find that the larger the endmember number, the smaller the optimal λ. This is because in the synthetic data we used, the larger endmember number means more complex hyperspectral scene and thus less spatial-contextual information, which leads to smaller optimal λ. As we observed in the synthetic data experiment, when the endmember number is fixed and the block size is set properly, the optimal λ is relatively stable and varies slightly with the SNR of the noise. For the real hyperspectral image, we divide it into several 64×64 blocks and compute the average endmember number in each block to roughly estimate the spatial complexity of the hyperspectral image. Then we refer to Tab. VI to choose the suitable λ. Note that the spatial distribution in synthetic data is ideal but that in real data is more complex. However, we find in the real experiments that by choosing λ according to Tab. VI after dividing the image, RSFoBa can always get good results and the optimal λ is very close to the one estimated by using Tab. VI, which also indicates the proposed algorithm behaves stably with this parameter. There are several algorithms that can estimate the number of endmembers in the hyperspectral image efficiently [57, 58]. In the real data experiment, we use the virtual dimensionality [57] to estimate the number of endmembers in each block. In the first real data we used, the endmember number in the scene is known to be about 14 [8] and the average number of endmembers in each 64×64 block is estimated to be about 11. Besides, the noise in the real hyperspectral data is usually correlated. Thus, referring to Tab. VI, we set the λ's in RSFoBa (p=2) and RSFoBa ($p = \infty$) in the first real data experiment to be 0.1 and 0.05 respectively, which can lead to their optimal performances. As there are mainly 5 targets of interest in the second real data experiment and the average number of endmembers in each block is less than 5, the λ's for RSFoBa (p=2) and RSFoBa ($p = \infty$) in the second

TABLE VI

THE OPTIMAL λ'S OF THE RSFoBa ALGORITHM OBTAINED ON THE SYNTHETIC DATA CORRUPTED BY 30 dB WHITE OR CORRELATED

NOISE IN SITUATIONS OF DIFFERENT ENDMEMBER NUMBERS

endmember number	3	5	7	9	11	13	15
RSFoBa ($p = 2$) (Correlated)	1	1	0.3	0.1	0.1	0.1	0.05
RSFoBa ($p = \infty$) (Correlated)	1	0.3	0.1	0.01	0.05	0.05	0.01
RSFoBa ($p = 2$) (White)	1	1	0.1	0.05	0.1	0.01	0.01
RSFoBa ($p = \infty$) (White)	1	0.3	0.3	0.05	0.01	0.01	0.01

real data experiment should be set to 1 and 0.3 respectively by referring to Tab. VI. However, as $\lambda = 3$ will lead to a slightly better performance of RSFoBa ($p = 2$) than $\lambda = 1$, the λ is finally set to 3 for RSFoBa ($p = 2$) to achieve its optimal performance. Nevertheless, we can still find that the optimal λ (i.e. 3) is very close to the one (i.e. 1) estimated by using Tab. VI.

The block size is also an important parameter for the RSFoBa algorithm. As mentioned earlier, when the number of endmembers present in the hyperspectral image is large, the simultaneous greedy algorithm becomes unstable and could miss the actual endmembers. In this paper, we adopt a block-processing strategy with proper block size to sidestep this obstacle. Here, we discuss for the RSFoBa algorithm when the block-processing strategy is necessary and how to properly set the block size. In the synthetic data experiment, we have tested the block size from 4 to the image size with the step size of 4. We find that when the number of endmembers present in the block or hyperspectral image is no more than 4, the RSFoBa algorithm can find the actual endmembers efficiently. Thus, the strategy is necessary when the endmember is larger than 4. Then a natural answer to the second question is that the block size should be chosen to make the number of endmembers in each block no more than 4. But rather than find the actual endmembers present in each block exactly, we only need to select all the actual endmembers from all the blocks. Thus, the choosing of block size can be relaxed to make the average number of endmembers in each block no more than 4. Besides, given a hyperspectral image, smaller block size leads to larger block number, which also helps preventing RSFoBa from missing the actual endmembers. Furthermore, we can also use the overlapping blocks to increase the number of blocks when the block size is fixed. In this way, we can avoid missing the actual endmembers further, but the number of endmembers selected by RSFoBa that is not present in the hyperspectral image will increase. In the real data experiment, we use the virtual dimensionality [57] to estimate the number of endmembers in each block. The average number of endmembers in each block will decrease as the block size decreases. For the first real data used in our experiment, when the block size is 20, the average number of endmembers in all the blocks is 4.5. Finally, the block size we set for RSFoBa in this real experiment is a little smaller than 20, i.e. 15. Similarly, following the above procedure and considering the smaller block size will make the RSFoBa faster (which will be seen in the next subsection), we set the block size in the second real experiment to 30. Note that the block sizes set for SOMP and SSP are different with those for RSFoBa in the experiments. This is because compared with RSFoBa, SOMP and SSP are more sensitive to this parameter. We also test different

TABLE VII

PROCESSING TIMES MEASURED AFTER APPLYING THE ALGORITHMS TO THE SYNTHETIC DATA WITH 30 dB WHITE NOISE CONSTRUCTED
BY 5 ENDMEMBERS

Algorithms	RSFoBa ($p = 2$)	RSFoBa-SUnSAL ($p = 2$)	SUnSAL	OMP	SOMP
Time (s)	17.50	14.50	34.79	8.82	4.10
Algorithms	RSFoBa ($p = \infty$)	RSFoBa-SUnSAL ($p = \infty$)	SUnSAL-TV	SSP	
Time (s)	10.13	8.25	346.92	5.93	

block sizes for SOMP and SSP on the synthetic data and find that when the number of endmembers present in the block or hyperspectral image is no more than 3, the SOMP and SSP are less likely to miss the actual endmembers. Thus, the block sizes set for SOMP and SSP are always smaller than those for RSFoBa to achieve optimal results.

E. Complexity Analysis

In the following analysis, suppose L and K are the band number and pixel number in the hyperspectral image, respectively; m is the number of spectral signatures in the spectral library. Suppose the row-sparsity level k is far smaller than m, L and K. We first analyze the computational complexity of the proposed RSFoBa algorithm. The complexity of step 7 in Algorithm 3 is $O(mLK)$. In step 9, step 14 and step 20 we need to solve the Lyapunov equation in Eq. (13). As $\mathbf{C}_2 = \lambda \mathbf{L}$ is a sparse matrix, its eigenvalue decomposition is much more computationally efficient than a $K \times K$ arbitrary matrix's worst case (i.e. $O(K^3)$). Indeed, if $\mathbf{C}_2$ is sufficiently sparse, its eigenvalue decomposition is linear complexity and thus it will never be the most time-consuming. Then, the complexity of solving the Lyapunov equation is dominated by implementing the matrix multiplication in Eq. (14) whose complexity is $O(K^2)$. But when the block-processing strategy is adopted, this complexity decreases to $O(K^2/n)$, where n is the block number in the hyperspectral image. Finally, the complexity of RSFoBa is given by $O(\max\{mLK, K^2/n\})$. Then we give out the computational complexities of the other considered algorithms. For the whole hyperspectral data, SOMP, SSP and OMP have complexity $O(mLK)$; SUnSAL and SUnSAL-TV have complexity $O(L^2K)$. When the hyperspectral image is very large, RSFoBa will be more time-consuming than the other considered sparse unmixing algorithms. But, the block-processing strategy can mitigate this difficulty.

Tab. VII reports the processing time (the average time after running each algorithm 10 times) measured after applying the algorithms to the synthetic data with 30 dB white noise constructed by 5 endmembers. A desktop PC equipped with an Intel Core 2 Duo CPU (2.66 GHz) and 4 GB of RAM memory was used to implement all the algorithms on MATLAB R2012b. The results of RSFoBa-SUnSAL are also included. We can find that using the sub-library obtained by RSFoBa as input, the speed of SUnSAL increases much.

V. CONCLUSION AND DISCUSSION

In this paper, we propose a novel algorithm termed regularized simultaneous forward-backward greedy algorithm (RSFoBa) for sparse unmixing of hyperspectral data. The main idea of the proposed algorithm is the combination

of forward greedy algorithm and backward greedy algorithm as well as considering several characteristics of the hyperspectral data. Experiments on both synthetic data and real data demonstrate the effectiveness of the RSFoBa algorithm. Besides, we also demonstrate that the sub-library obtained by RSFoBa can serve as input for any other sparse unmixing algorithms, making them more accurate and time-efficient.

The first topic we want to discuss is why RSFoBa is effective for sparse unmixing. Indeed, the main difference among RSFoBa, SOMP and SSP is the endmember selection procedure. The final abundances for each endmember are estimated using the constrained least squares technique. Two factors lead to the sparsity of the abundance vectors estimated by these three algorithms: on one hand the numbers of the potential endmembers selected by them are much smaller than the number of spectral signatures in the library; on the other hand, based on the relatively small library, the constrained least squares technique makes the abundance vectors sparse further. It is obvious that the more accurately the actual endmembers are selected, the better the abundances will be estimated. As RSFoBa can select the actual endmembers more accurately than SOMP and SSP, it is more effective for sparse unmixing. However, from Fig. 8, we can observe that the number of endmembers selected by the proposed algorithm is still larger than the actual number due to the high correlation of the spectral library. Thus, abundances could be generated for the non-existing endmembers in the successive step, which will have a negative effect on the estimation of the abundances corresponding to the actual endmembers. In this paper, we also use the sub-library obtained by RSFoBa as input for the other sparse unmixing algorithms to get the optimal sparse solution. Experimental results indicate that this combination can improve the unmixing results of RSFoBa. But we need to tune more regularization parameters to achieve the optimal results. In conclusion, RSFoBa not only can be an effective sparse unmixing algorithm but also can be a preprocessing step to boost other sparse unmixing algorithms.

Another topic we want to discuss is the obstacles that will be faced by sparse unmixing algorithms in the real situation. The first obstacle is that the spectral signatures from the spectral library could differ to image signatures because they were acquired from different region and background, which would significantly degrade abundance estimated accuracy. This obstacle has been considered in [16]. Specifically, advanced atmospheric correction algorithms which convert the input hyperspectral data from at-sensor radiance to reflectance units have been proposed to solve this problem. For more details, we would like to refer to [16]. The second obstacle is that the image signatures and spectral signatures from the library may have different spectral channels. As the hyperspectral signatures are usually very smooth, this obstacle can be overcome by using the interpolation techniques [59]. The third obstacle is that how and what target spectra to be selected manually in practice if we do not have a priori knowledge. In this case, the sparse unmixing algorithms can still be applied by resorting to an artificially generated spectral library constructed using image-derived endmembers [16]. These endmembers can be generated using the endmember extraction algorithms [6, 8] or dictionary learning methods [19] which have been applied to hyperspectral data in [60].

ACKNOWLEDGMENT

REFERENCES

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

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

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

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

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

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

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

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

[9] J. Li and J. Bioucas-Dias, "Minimum volume simplex analysis: a fast algorithm to unmix hyperspectral data," in *IEEE Geoscience and Remote Sensing Symposium- IGARSS'08*, Boston, 2008.

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

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

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

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

[14] X. Lu, H. Wu, Y. Yuan, P. Yan, and X. Li, "Manifold regularized sparse nmf for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, pp. 1–12, Oct. 2012.

[15] 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," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 49, no. 8, pp. 2945–2956, 2011.

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

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

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

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

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

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

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

[23] M.-D. Iordache, J. M. B. Dias, and A. Plaza, "Collaborative sparse regression for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, article in press.

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

[25] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to algorithms*. MIT press, 2001.

[26] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *Signal Processing, IEEE Transactions on*, vol. 41, no. 12, pp. 3397–3415, 1993.

[27] Y. C. Pati, R. Rezaiifar, and P. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Signals, Systems and Computers, 1993. 1993 Conference Record of The Twenty-Seventh Asilomar Conference on*. IEEE, 1993, pp. 40–44.

[28] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *Information Theory, IEEE Transactions on*, vol. 53, no. 12, pp. 4655–4666, 2007.

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

[30] M. V. Afonso, J. M. Bioucas-Dias, and M. A. Figueiredo, "An augmented lagrangian approach to the

constrained optimization formulation of imaging inverse problems," *Image Processing, IEEE Transactions on*, vol. 20, no. 3, pp. 681–695, 2011.

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

[32] S. Ji, Y. Xue, and L. Carin, "Bayesian compressive sensing," *Signal Processing, IEEE Transactions on*, vol. 56, no. 6, pp. 2346–2356, 2008.

[33] S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Bayesian compressive sensing using laplace priors," *Image Processing, IEEE Transactions on*, vol. 19, no. 1, pp. 53–63, 2010.

[34] X.-L. Zhao, F. Wang, T.-Z. Huang, M. K. Ng, and R. Plemmons, "Deblurring and sparse unmixing for hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, article in press.

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

[36] T. Zhang, "Adaptive forward-backward greedy algorithm for learning sparse representations," *Information Theory, IEEE Transactions on*, vol. 57, no. 7, pp. 4689–4708, 2011.

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

[38] C. A. Bateson, G. P. Asner, and C. A. Wessman, "Endmember bundles: a new approach to incorporating endmember variability into spectral mixture analysis," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 38, no. 2, pp. 1083–1094, 2000.

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

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

[41] C. Couvreur and Y. Bresler, "On the optimality of the backward greedy algorithm for the subset selection problem," *SIAM Journal on Matrix Analysis and Applications*, vol. 21, no. 3, pp. 797–808, 2000.

[42] S. J. Reeves, "An efficient implementation of the backward greedy algorithm for sparse signal reconstruction," *Signal Processing Letters, IEEE*, vol. 6, no. 10, pp. 266–268, 1999.

[43] D. M. Rogge, B. Rivard, J. Zhang, and J. Feng, "Iterative spectral unmixing for optimizing per-pixel endmember sets," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 44, no. 12, pp. 3725–3736, 2006.

[44] N. Rao, P. Shah, S. Wright, and R. Nowak, "A greedy forward-backward algorithm for atomic norm constrained minimization," in *ICASSP*, 2013.

[45] A. GOETZ, "Imaging spectrometry for earth remote sensing," *Science*, vol. 228, pp. 1147–1153, 1985.

[46] C. Bachmann, T. Ainsworth, and R. Fusina, "Exploiting manifold geometry in hyperspectral imagery," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 43, no. 3, pp. 441–454, 2005.

[47] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Dictionary pruning in sparse unmixing of hyperspectral

data," in *2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, 2010, pp. 1–4.

[48] A. Jameson, "Solution of the equation ax+xb=c by inversion of an m*m or n*n matrix," *SIAM Journal on Applied Mathematics*, vol. 16, no. 5, pp. 1020–1023, 1968.

[49] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE Transactions on information theory*, vol. 55, no. 5, pp. 2230–2249, May 2009.

[50] Y. Chen, N. M. Nasrabadi, and T. D. Tran, "Hyperspectral image classification using dictionary-based sparse representation," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 49, no. 10, pp. 3973–3985, 2011.

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

[52] D. C. Howell, *Statistical methods for psychology*.   Cengage Learning, 2011.

[53] D. W. Zimmerman, "Teachers corner: A note on interpretation of the paired-samples t test," *Journal of Educational and Behavioral Statistics*, vol. 22, no. 3, pp. 349–360, 1997.

[54] J. Plaza, A. Plaza, P. Martínez, and R. Pérez, "H-comp: A tool for quantitative and comparative analysis of endmember identification algorithms," in *Geoscience and Remote Sensing Symposium, 2003. IGARSS'03. Proceedings. 2003 IEEE International*, vol. 1.   IEEE, 2003, pp. 291–293.

[55] R. N. Clark, G. A. Swayze, K. E. Livo, R. F. Kokaly, S. J. Sutley, J. B. Dalton, R. R. McDougal, and C. A. Gent, "Imaging spectroscopy: Earth and planetary remote sensing with the usgs tetracorder and expert systems," *Journal of Geophysical Research*, vol. 108, no. E12, p. 5131, 2003.

[56] J. A. Benediktsson, J. A. Palmason, and J. R. Sveinsson, "Classification of hyperspectral data from urban areas based on extended morphological profiles," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 43, no. 3, pp. 480–491, 2005.

[57] C.-I. Chang and Q. Du, "Estimation of number of spectrally distinct signal sources in hyperspectral imagery," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 42, no. 3, pp. 608–619, 2004.

[58] J. M. Bioucas-Dias and J. M. Nascimento, "Hyperspectral subspace identification," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 46, no. 8, pp. 2435–2445, 2008.

[59] P. Geladi, J. Burger, and T. Lestander, "Hyperspectral imaging: calibration problems and solutions," *Chemometrics and Intelligent Laboratory Systems*, vol. 72, no. 2, pp. 209–217, 2004.

[60] A. Castrodad, Z. Xing, J. B. Greer, E. Bosch, L. Carin, and G. Sapiro, "Learning discriminative sparse representations for modeling, source separation, and mapping of hyperspectral imagery," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 49, no. 11, pp. 4263–4281, 2011.