HALS-based NMF with Flexible Constraints for Hyperspectral Unmixing

Weishi Chen, Mireille Guillaume

HALS-based NMF with Flexible Constraints for Hyperspectral Unmixing

Weishi Chen[1, 2] and Mireille Guillaume[1]

1. Institut Fresnel, 13397 Marseille, France
2. School of Electronics and Information Engineering, Beihang University, 100191 Beijing, China

***Abstract*-In this paper, the hyperspectral unmixing problem is solved with the nonnegative matrix factorization (NMF) algorithm. The regularized criterion is minimized with a hierarchical alternating least squares (HALS) scheme. Under the HALS framework, four constraints are introduced to improve the unmixing accuracy, including the sum-to-unity (STU) constraint, the constraints for minimum spectral dispersion and maximum spatial dispersion, and the minimum volume constraint. The derived algorithm is called F-NMF, for NMF with flexible constraints. We experimently compare F-NMF with different constraints and combined ones. We test the sensitivity and robustness of F-NMF to many parameters such as the purity level of endmembers, the number of endmembers and pixels, the SNR, the sparsity level of abundances, and the overestimation of endmembers. The proposed algorithm improves the results estimated by** vertex component analysis **(VCA). A comparative analysis on real data is included. The unmixing results given by a geometrical method, the simplex identification via split augmented Lagrangian (SISAL) and the F-NMF algorithms with combined constraints are compared, which shows the relative stability of F-NMF.**

***Index Terms*- Hyperspectral unmixing, nonnegative matrix factorization (NMF), hierarchical alternating least squares (HALS), constraint.**

I. Introduction

Airborne hyperspectral sensors collect images in hundreds of narrow and contiguous spectral bands. Due to the limited spatial resolution of hyperspectral image (HSI), each observed pixel generally contains more than one material spectral signature. Hence, the hyperspectral unmixing, which decomposes a mixed pixel into a combination of pure material spectra known as endmembers, weighted by their corresponding abundance coefficients, is a challenging task.

Let $\mathbf{R}$ ($L\times I$) be the matrix unfolded HSI, whose I columns are the spectral pixels and the L rows are the vectorial spectral band images. As N is the related noise matrix, the linear spectral mixing model (LSMM) can be written as

$$\mathbf{R} = \mathbf{AS} + \mathbf{N} = \mathbf{X} + \mathbf{N} \tag{1}$$

The rows of $\mathbf{S}$ ($J\times I$) are the abundance maps corresponding to the respective endmembers, whose spectra are located in the columns of $\mathbf{A}$ ($L\times J$). J denotes the number of endmembers.

Basically, hyperspectral unmixing is a problem of blind source separation (BSS). However, compared with most BSS applications, the endmembers of HSI data are dependent and the elements in $\mathbf{A}$ and $\mathbf{S}$ are nonnegative, so the hyperspectral unmixing is beyond the reach of many BSS algorithms (e.g. independent component analysis (ICA) [1]). To fulfill these constraints, numerous special algorithms have been proposed to solve the hyperspectral unmixing problem under the LSMM assumption, including the approaches of convex geometry, Bayesian source separation (BSS), and nonnegative matrix factorization (NMF). The geometrical approaches first determine the endmembers and estimate the abundances in a second step, while the BSS and NMF-based approaches find the endmembers and the abundances simultaneously.

Geometrical approaches try to determine the vertices of the J-simplex enclosing the observed pixels, such as pixel purity index (PPI) [2], N-FINDR [3], vertex component analysis (VCA) [4]. The PPI algorithm projects every spectral vector onto skewers (large number of random vectors). The points corresponding to extremes, for each skewer direction, are stored and cumulated. The pixels with the highest scores are the purest ones. N-FINDR finds the set of pixels defining the largest volume within the data. VCA iteratively projects data onto a direction orthogonal and the endmembers correspond to the extreme of the projections. The issue of these approaches is to find extreme points within the data with the assumption of pure pixel of each endmember, which is always unsatisfactory for real hyperspectral data. Recently, the state-of-art reference algorithms MVSA [5], MVES [6] and the simplex identification via split augmented Lagrangian (SISAL) [7] have proposed various ways to find a minimum volume simplex, showing very good performances in the estimation of endmembers. Particularly, SISAL is able to unmix HSI data in the case of no pure pixel.

The geometrical approaches do not work well when the observed data are highly mixed, because there are not enough vectors in simplex facets. In these cases, the separation problem can be addressed in a Bayesian framework. Several Bayesian Positive Source Separation (BPSS) algorithms under positivity and sum-to-one constraints have been recently developed [8-10]. In [10], a discussion on the effectiveness of the sum-to-one constraint is done, showing that full constrained BPSS2 gives better results than BPSS for simulated data, while it is the contrary for the real OMEGA data, "due to non-linearity in the radiative transfer and noise in the dataset in contradiction with the full additivity constraint". We think that it would be not the same with the proposed NMF-based algorithm, firstly because the full additivity is not a hard but a soft constraint, and second because the residual error RQE is able to represent measurement noise or model noise, and then the algorithm is quite robust for real data, which can contain non-linear mixed terms. This can be seen by comparing the results of a geometrical algorithm like SISAL, very performant on simulated data, with the results obtained on Cuprite real data, which drops dramatically, while the NMF-based algorithms keep performing.

In the last decade, NMF has been a popular algorithm since Lee and Seung [11] investigated the properties of the algorithm and published some simple and useful algorithms for two types of factorizations. The NMF algorithm has been broadly used in text mining, image analysis, speech processing and automatic control. The basic NMF problem consists of finding two nonnegative data matrices whose product approximates the mixed data in a chosen measure sense [e.g., the reconstruction quadratic error (RQE)]. However, the solution to NMF is not unique so various regularizations with prior knowledge should be taken into account to reduce the number of solutions. The sum-to-unity (STU) constraint is proposed in [12], which regularizes the RQE with a function of $\mathbf{S}$ to normalize the columns of it. The authors of [13] propose constraints based on two inherent characteristics of hyperspectral data: the spectral piecewise smoothness and spatial sparseness. In [14], a minimum volume constrained NMF (MVC-NMF) based on projected gradient (PG) optimization method is proposed, whose regularization term minimizes the simplex volume spanned by the endmembers.

Other authors [15], propose a minimum distance constrained NMF (MDC-NMF), which consider the endmember distance instead of the volume of the estimated simplex. MDC-NMF makes a slight modification of the optimized algorithm used for MVC-NMF. MiniDisCo algorithm makes the assumption of minimum spectral dispersion for NMF regularization [16], and MDMD-NMF regularizes with minimum spectral dispersion and maximum spatial dispersion [17]. A new step-size estimation technique is proposed for the two algorithms to hasten the PG convergence.

The optimization algorithms and constraints on **A** and **S** are two main techniques for NMF-based hyperspectral unmixing. The authors of [18] propose a flexible hierarchical alternating least squares (HALS) algorithm with a set of local cost functions called alpha and beta divergences. The word "flexible" means the variation of the optimization algorithm. In this paper, we propose an improved NMF algorithm with four constraints due to the characteristics of HSI, called the flexible NMF (F-NMF). The word "flexible" means the variation of constraints on **A** and **S**. F-NMF also uses the HALS update rules, significantly outperfoming the PG update rules in convergence speed. Actually, the novelty is both the combination of the constraints and the development of these constraints under HALS-based algorithm.

The paper is organized as follows: Section II presents the basic NMF algorithm and the HALS update rules. In Section III, we introduce four contraint functions and integrate them into the F-NMF algorithm. In Section IV, the comparison and analysis of the F-NMF with different constraints are given by processing various simulated HSIs. The algorithms are applied to real data in Section V. The F-NMF algorithms are compared with SISAL, for the two algorithms are both able to unmix hyperspectral data in which the pure pixel assumption is violated. Finally, some conclusion closes the paper.

II. NMF FOR HYPERSPECTRAL UNMIXING

In this section, we first present the NMF problem and then the optimization algorithm used to solve it in this paper.

A. NMF problem

The aim of basis NMF methods is to find two estimated matrices $\hat{\mathbf{A}}$ and $\hat{\mathbf{S}}$ such that

$$\mathbf{X} \simeq \hat{\mathbf{A}}\hat{\mathbf{S}} \tag{2}$$

A commonly used theoretical solution is to find nonnegative matrices minimizing the RQE

$$\mathrm{RQE}(\mathbf{A},\mathbf{S}) = \left\|\mathbf{X}-\mathbf{AS}\right\|_F^2 \tag{3}$$

where $\|\cdot\|_F$ is the Frobenius (e.g., quadratic) norm.

B. HALS algorithms

In [19], the authors show that the HALS scheme works remarkably well in practice, outperforming, in most cases, the other optimization algorithms for NMF. In particular, it is proved to be locally more efficient [20] and shown to converge to a stationary point under some mild assumptions [21]. For these reasons, we choose HALS as the optimization technique.

The basic idea of HALS is to define residues as

$$\mathbf{X}^{(k)} = \mathbf{X} - \sum_{i\neq k}\mathbf{A}_i\mathbf{S}_i = \mathbf{X} - \mathbf{AS} + \mathbf{A}_k\mathbf{S}_k \tag{4}$$

for k = 1, 2,…, J. $\mathbf{A}_k$ (L×1) is one endmember spectrum and $\mathbf{S}_k$ (1× I) corresponds to its abundance fraction.

By substituting equation (4) into (3), the new RQE function is

$$\mathrm{RQE}(\mathbf{A}_k,\mathbf{S}_k) = \left\|\mathbf{X}^{(k)} - \mathbf{A}_k\mathbf{S}_k\right\|_F^2 \tag{5}$$

The gradients of the above function are expressed by

$$\begin{aligned}\frac{\partial \mathrm{RQE}(\mathbf{A}_k,\mathbf{S}_k)}{\partial \mathbf{A}_k} &= 2(\mathbf{A}_k\left\|\mathbf{S}_k\right\|^2 - \mathbf{X}^{(k)}\mathbf{S}_k^T)\\ \frac{\partial \mathrm{RQE}(\mathbf{A}_k,\mathbf{S}_k)}{\partial \mathbf{S}_k} &= 2(\left\|\mathbf{A}_k\right\|^2\mathbf{S}_k - \mathbf{A}_k^T\mathbf{X}^{(k)})\end{aligned} \tag{6}$$

By setting the above equation to zero, the updating rules are obtained:

$$\mathbf{A}_k \leftarrow \left[\frac{\mathbf{X}^{(k)}\mathbf{S}_k^T}{\left\|\mathbf{S}_k\right\|^2}\right]_{[0,1]}, \qquad \mathbf{S}_k \leftarrow \left[\frac{\mathbf{A}_k^T\mathbf{X}^{(k)}}{\left\|\mathbf{A}_k\right\|^2}\right]_{[0,1]}. \tag{7}$$

For k = 1, 2, …, J, where $[\delta]_{[0,1]}$ is to enforce every element δ_{ij} lies in [0, 1], so

$$\left([\boldsymbol{\delta}]_{[0,1]}\right)_{ij} = \begin{cases}0, & \text{if } \delta_{ij}<0\\ \delta_{ij}, & \text{if } 0<\delta_{ij}<1\\ 1, & \text{if } \delta_{ij}>1\end{cases}$$

Clearly, the HALS algorithm is bound-constrained. It is also shown that the optimal value of each entry of **A** ($\mathbf{A}_k$) does not depend on the other entries of the same column. By symmetry, the same property holds for each row of **S** ($\mathbf{S}_k$). Thus, the detailed HALS algorithm is summarized as follows:

1) Initialize **A** and **S** with the VCA algorithm;
2) **for** i = 1, 2, …, do
   for k = 1, 2, …, J
      Update $\mathbf{A}_k$ and $\mathbf{S}_k$ with the HALS update rules;
   end
   until the stop criteria is reached

The simplest update rules are given in eq. (7), and the regularized f with all constraints will be proposed in eq. (18). The maximum number of iterations is always set high (e.g. 2000) to obtain accuate estimations. However, the overestimation of the iteration number induces time waste. Indeed, the RQE value slightly increases from certain iteration whereas the regularized f keeps decreasing. Thus, the algorithm is stopped at this iteration when the RQE value goes to a minimum although the highest iteration number is not reached. The stop criteria is expressed as

$$\mathrm{RQE}^{k-50} < \min_{k'=0,\ldots,50}\mathrm{RQE}^{k-k'}.$$

III. NMF WITH FLEXIBLE CONSTRAINTS

The basic NMF optimized function ensures that the two constraints **A** and **S** are both nonnegative. Since the NMF solution is not unique, some prior knowledge on HSIs can be introduced to regularize the problem. A generic expression for the optimized function is

$$f(\mathbf{A},\mathbf{S}) = \sum_e \sigma_e D_e(\mathbf{A},\mathbf{S}) + \sum_i \alpha_i D_i(\mathbf{S}) + \sum_j \beta_j D_j(\mathbf{A}) \tag{8}$$

where σ, α and β are regularized parameters for the estimation error and the spectral and abundance constraints. D(**A**,**S**) measures the difference between **X** and **AS** with respect to some norms. By substituting equation (4) into (8) and using the RQE norm, the new optimized function f is

$$f(\mathbf{A}_k,\mathbf{S}_k) = \mathrm{RQE}(\mathbf{A}_k,\mathbf{S}_k) + \sum_i \alpha_i D_i(\mathbf{S}_k) + \sum_j \beta_j D_j(\mathbf{A}_k) \tag{9}$$

In this section, we add four constraints for **A** and **S** to the function to improve the unmixing result. With all these constraints, the algorithm is called flexible NMF (F-NMF), based on HALS update rules.

A. STU constraint

The STU constraint makes the sums of the columns of **S** equal to 1. The STU constraint is defined as follows:

$$D_1(\mathbf{S}_k)=\left\|\mathbf{S}_k+\sum_{i\neq k}\mathbf{S}_i-\mathbf{1}_{1I}\right\|_F^2 \tag{10}$$

where $\mathbf{1}_{1I}$ is an (1×I) vector of ones. The gradient derivation of D_1 with respect to $\mathbf{S}_k$ is

$$\frac{\partial D_1(\mathbf{S}_k)}{\partial \mathbf{S}_k}=2\left(\mathbf{S}_k+\sum_{i\neq k}\mathbf{S}_i-\mathbf{1}_{1I}\right) \tag{11}$$

B. Maximum spatial dispersion constraint

In real situations, abundance matrix is often very sparse because the materials are mostly grouped in separate regions even if the pure pixels. We note that reducing the data enclosing simplex volume is equivalent to increase the dispersion of the abundances fractions in the sum-to-one constrained subspace enclosing the abundances. Actually, the most impossible situation is the uniformly-mixed data. Therefore, as the mean value of abundances is 1/*J*, we defined the maximum spatial dispersion constraint as follows:

$$D_2(\mathbf{S}_k)=-\left\|\mathbf{S}_k-\frac{1}{J}\mathbf{1}_{1I}\right\|_F^2 \tag{12}$$

This constraint encourages null abundance pixels and full pixels, as in real scenes, not all the endmembers are present in all pixels, and in contrast some pixels contain only one material. The gradient derivation of D_2 with respect to $\mathbf{S}_k$ is

$$\frac{\partial D_2(\mathbf{S}_k)}{\partial \mathbf{S}_k}=-2\left(\mathbf{S}_k-\frac{1}{J}\mathbf{1}_{1I}\right) \tag{13}$$

C. Minimum spectral dispersion constraint

This constrained function depends on **A**, encouraging the variance of each endmember spectrum to be as low as possible. This dispersion constraint is to improve the shape estimation of flat endmember spectra. Consequently, if the estimation of some spectra is improved, the estimation of the other spectra involved in the mixture will also indirectly be improved due to the parameter interdependences. We define the minimum spectral dispersion constraint as

$$D_1(\mathbf{A}_k)=\left\|\left(\mathbf{I}_{LL}-\frac{1}{L}\mathbf{1}_{LL}\right)\mathbf{A}_k\right\|_F^2 \tag{14}$$

The gradient derivation of D_1 with respect to $\mathbf{A}_k$ is

$$\frac{\partial D_1(\mathbf{A}_k)}{\partial \mathbf{A}_k}=2\left(\mathbf{I}_{LL}-\frac{1}{L}\mathbf{1}_{LL}\right)\mathbf{A}_k \tag{15}$$

D. Minimum distance constraint

In MVC-NMF [9], the volume of **A** is calculated as the constraint, which suffers from numerical instabilities [11]. Here, we choose the minimum distance constraint as a substitute in order to shrink the volume of the data enclosing the simplex. The distance is measured and summed up from every endmember to their centroid. This constraint is defined as

$$D_2(\mathbf{A}_k)=\left\|\left(\mathbf{I}_{LL}-\frac{1}{L}\mathbf{1}_{LL}\right)\left(\mathbf{A}_k-\frac{1}{J}(\mathbf{A}_k+\sum_{i\neq k}\mathbf{A}_i)\right)\right\|_F^2 \tag{16}$$

The gradient derivation of D_2 with respect to $\mathbf{A}_k$ is

$$\frac{\partial D_2(\mathbf{A}_k)}{\partial \mathbf{A}_k}=2(\mathbf{I}_{LL}-\frac{1}{L}\mathbf{1}_{LL})(1-\frac{1}{J})\left(\mathbf{A}_k-\frac{1}{J}(\mathbf{A}_k+\sum_{i\neq k}\mathbf{A}_i)\right) \tag{17}$$

The final F-NMF update rules to minimize f with all these considerations are derived from (6), (9), (11), (13), (15) and (17). Thus

$$\mathbf{S}_k\leftarrow\left[\frac{\mathbf{A}_k^T\mathbf{X}^{(k)}+\alpha_1(\mathbf{1}_{1I}-\sum_{i\neq k}\mathbf{S}_i)-\frac{\alpha_2}{J}\mathbf{1}_{1I}}{\|\mathbf{A}_k\|^2+\alpha_1-\alpha_2}\right]_{[0,1]}$$

$$\mathbf{A}_k\leftarrow\left[\frac{\mathbf{X}^{(k)}\mathbf{S}_k^T+\beta_2\frac{1}{J}(1-\frac{1}{J})(\mathbf{I}_{LL}-\frac{1}{L}\mathbf{1}_{LL})\sum_{i\neq k}\mathbf{A}_i}{\|\mathbf{S}_k\|^2\mathbf{I}_{LL}+(\mathbf{I}_{LL}-\frac{1}{L}\mathbf{1}_{LL})\cdot\left(\beta_1+\beta_2(1-\frac{1}{J})^2\right)}\right]_{[0,1]} \tag{18}$$

IV. Simulations on Synthetic Data

In this section, we present a batch of simulations to quantitively compare the F-NMF algorithms with different constraints. First, we present the used evaluation metrics. Then, we present the way we build simulated data. Finally, the experimental results of five F-NMF algorithms are given.

A. Evaluation metrics

1) To evaluate the abundance estimation, we define the abundance mean squared error (AME) as

$$\mathrm{AME}(\hat{\mathbf{S}},\mathbf{S})=\frac{1}{JI}\left\|\hat{\mathbf{S}}-\mathbf{S}\right\|_F^2 \tag{19}$$

2) To evaluate the endmember spectra estimation, we define the spectral mean squared error (SME) as

$$\mathrm{SME}(\hat{\mathbf{A}},\mathbf{A})=\frac{1}{LJ}\left\|\hat{\mathbf{A}}-\mathbf{A}\right\|_F^2 \tag{20}$$

3) To consider the global shape of the spectra, the spectral angle distance (SAD) is defined as

$$\mathrm{SAD}(\hat{\mathbf{a}},\mathbf{a})=\cos^{-1}\left(\frac{\mathbf{a}^T\hat{\mathbf{a}}}{\sqrt{\hat{\mathbf{a}}^T\hat{\mathbf{a}}}\sqrt{\mathbf{a}^T\mathbf{a}}}\right) \tag{21}$$

where **a** is the true spectral vector and $\hat{\mathbf{a}}$ is its estimate.

B. Synthetic data

The HSI synthesis process is in three steps corresponding to the matrices **A**, **S** and the noise matrix **N**.

First, the *J* endmember spectra are randomly selected among the U.S. Geological Survey (USGS) spectral library. The selected 224-channel spectra constitute the columns of the matrix **A**.

Then, the *J*-element column vector in **S** is generated following a Dirichlet pdf, with parameters equal to 1. The element maximal value of each column is controlled by a threshold ξ ($0<\xi\leq 1$). This operation allows one to control the mixing or purity level of the data. In particular, the image can contain "pure" pixels when $\xi = 1$. We also introduce a sparsity parameter ι ($\iota > 0$), which controls the sparsity of **S**. If ι is set at 0.8, 20% of the $J\times I$ elements in **S** are selected randomly and set to zeros at first, and then the non-zero elements in each column vector of **S** are generated following the Dirichlet pdf with the STU constraint and the maximal threshold ξ.

Finally, we add a noise matix **N**, assumed to be zero-mean white Gaussian. The noise is characterized by the SNR

$$\mathrm{SNR}=10\cdot\log\left(\frac{\|\mathbf{X}\|_F^2}{LI\sigma^2}\right)$$

where σ^2 is its variance.

Therefore, a synthetic HSI is characterized by *J*, the randomly selected endmember spectra, *I*, ξ, ι, and the SNR. The default configuration is given in Table 1.

TABLE 1

Parameter	Default value
J	4
I	1000
ξ	0.8
ι	0.8
SNR	Inf

C. Compared algorithms

In our simulations, we compare the F-NMF algorithms with different constraints and the typical geometrical and Bayesian algorithms. All the algorithms are used with the same initial and stop conditions.

1) F1-NMF: the basic HALS-NMF with no extra constraint based on HALS optimization algorithm. Only the nonnegative constraints are guaranteed.
2) F2-NMF: the HALS-NMF is improved with the STU constraint.
3) F3-NMF: the HALS-NMF with the STU and maximum spatial dispersion constraints.
4) F4-NMF: the HALS-NMF with the STU and minimum spectral dispersion constraints.
5) F5-NMF: the HALS-NMF with the STU and minimum distance constraints.
6) F35-NMF: the HALS-NMF with the combined constraints of F3 and F5.
7) VCA: a popular geometrical algorithm proposed in [4].
8) BPSS2: an improved Bayesian algorithm addressed in [9] under non-negativity and full additivity constraints.
9) MiniDisCo: a novel NMF-based algorithm with spectral constraint given in [16].

Note that the initializations of **A** and **S** for all the algorithms are chosen from a uniform distribution on the interval [0, 1].

D. Simulations

The first simulation shows the behavior of the objective f function along the interations of two optimization algorithms. Experiments 2-7 present statistical simulations to compare the average behaviors of the five F-NMF algorithms while varying the parameters given in Table 1, and robustness to an overestimation of the endmember number J.

1) The first experiment is to assess the choice of the optimization algorithm. We compare the convergence efficiency between the PG, which is widely used for NMF optimization, and the HALS algorithm. Here, the PG and HALS algorithms are regularized with the minimum spectral dispersion constraint. The PG-based algorithm is named MiniDisCo in [11], and HALS-based algorithm in this experiment is also called F4-NMF as above. The f value is calculated with the corresponding constraints and the perfomances of the two estimators are presented in Fig. 1. Note that both the curves result from the same HSI, with the same random initial conditions; thus, the only variablity is the optimization method. We note that the final value of f is almost the same with both algorithms, whereas the convergence speed of HALS is faster.

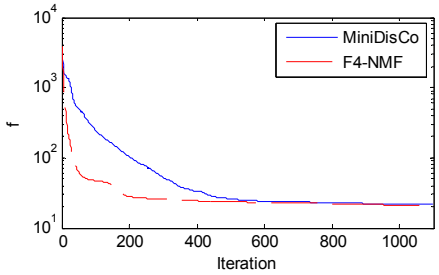


Fig. 1. The f value along the iterations.

The following experiments 2-7 present the behaviors of the F-NMF algorithms with different constraints while varying the parameters summed up in Table 1. The unmixing results are evaluated by AME, SME and SAD. The presented results are averages (bars) and standard deviations (error bars) resulting from 20 experiments. Note that all the considered algorithms are compared on the same sets of 20 HSIs. We perform preliminary Monte Carlo simulations to find relevant values for the regularized parameters. The retained values are $\alpha_1 = 1$, $\alpha_2 = 0.1$, $\beta_1 = 0.1$ and $\beta_2 = 0.1$, which are chosen to minimize the average evaluation errors for synthetic data.

2) In the second experiment, the algorithms are compared when the number of endmembers J varies.

In this experiment, we first test the efficiency of the algorithms. The F-NMF algorithms are compared with the PG-NMF with no contraint except positivity as F1-NMF, and J is set from 3 to 10 as the experiment in [11]. The performance metrics of SME are shown in Fig. 2. Note that the considered statistics do not necessarily include each of the results. Here, SME values higher than 0.5 are not included. In particular, the PG-NMF results are never considered while the F_NMF results are all included, because the SME values of PG-NMF are always greater than 0.5. With NMF algorithms, only a local minimum can be attained in general. In the case of random initializations and no constraints, HALS is able to obtain a better solution than PG.

Then, we set J higher to 20 to test the performance of F-NMF. The performance metrics are shown in Fig. 3. Note that the F1-NMF without constraints performs worse as the number of endmembers increases. In the case of the constrained F-NMF (F2, F3, F4, F5, F35), the results are much better. Fig. 3 puts forward the high robustness of the constrained F-NMF algorithms, when the basic F-NMF is sensitive to the number of endmembers. The combination of constraints F3 and F5, F35, gives good results.

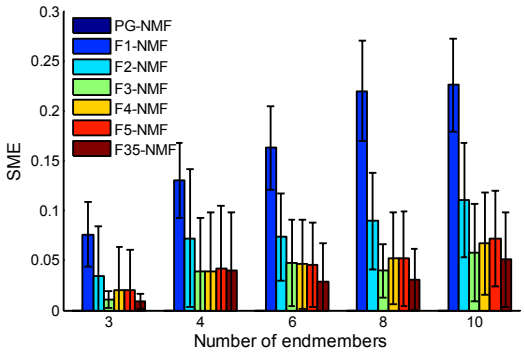
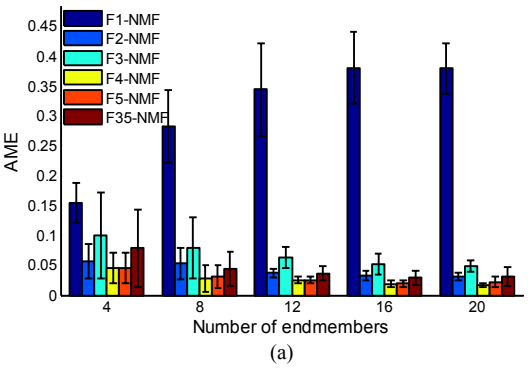
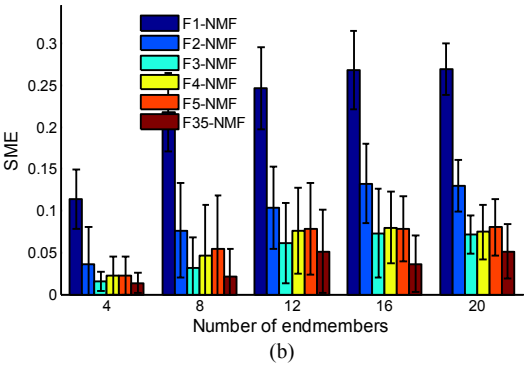


Fig. 2. Algorithm performances for different values of J.



(a)



(b)

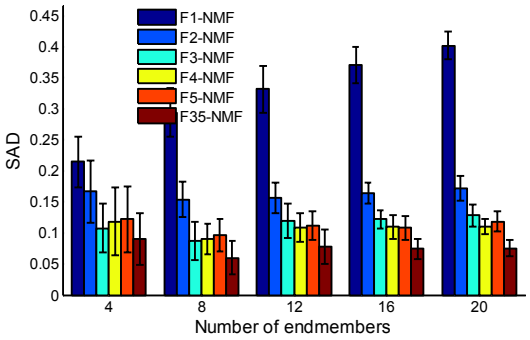


Fig. 3. Algorithm performance for different values of J. (a) AME; (b) SME; (c) SAD.

3) The purity level ξ is the topic of the third experiment. None of the considered algorithms are based on the hypothesis of one pure pixel for each endmember, but the un-mixing performance may vary with the purity. The obtained performance metrics are presented in Fig. 4. F3-NMF is particularly worse for AME when $\xi = 0.6$, because the low purity level make the maximum spatial dispersion constraint ineffective. F35-NMF also performs worse in term of AME due to the maximum spatial dispersion constraint.

Fig. 5 shows a comparison of the proposed F35-NMF algorithm with the geometrical method (e.g. VCA), a BSS algorithm (e.g. BPSS2) and another NMF-based algorithm (e.g. MiniDisCo), with the variation of ξ. The two NMF-based algorithms and BPSS2 are each initialized with VCA. The parameter of the spectral constraint is 0.1 for MiniDisCo. VCA performs better with higher purity level due to its assumption of pure pixels. MiniDisCo and F35-NMF both improve the unmixing results of VCA. Specifically, MiniDisCo and BPSS2 outperform F35-NMF in the sense of AME but the result is quite the reverse in the sense of SME, which is caused by different constraints in MiniDisCo and F35-NMF. In the sense of AME, F35-NMF performs worse as the purity level decreases, because the algorithm is regularized by the maximum spatial dispersion constraint, which improves the values of AME for the mixing data with high purity level. This could be verified by the results shown in Fig. 4(a). The algorithms with the maximum spatial dispersion constraint (F3-NMF and F35-NMF) give worse results than the other algorithms (F4-NMF and F5-NMF). We choose F35-NMF for comparison due to its better performances in SME and SAD. In the sense of SAD, MiniDisCo is better than F35-NMF with lower purity level, but F35-NMF performs better with purity level lower than 0.7. The performance of BPSS2 is always worse. This may be resulted by the minimum distance constraint in F35-NMF, which plays an important role in the unmixing of highly-mixed data.

As it can be noted, the results can vary for the various metrics, i.e., some algorithms can be efficient for spectral estimation, and not for abundances, and vice-versa. We have chosen to keep the three metrics for the complement of information they bring. A small SAD indicates very similar spectral shapes, and is not sensitive to a scale factor, while SME also depends on the values and is sensitive to a scale factor. For abundances, only the values are relevant, because the sum-to-one constraint sets the scale factor. In one sense, SAD is the more meaningful metric, used to identify endmembers from spectral libraries.

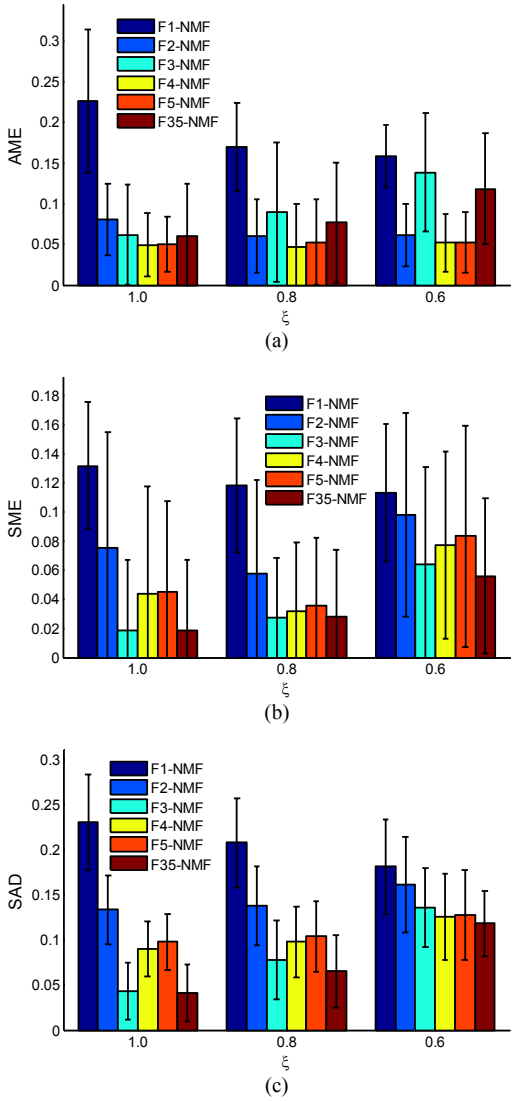
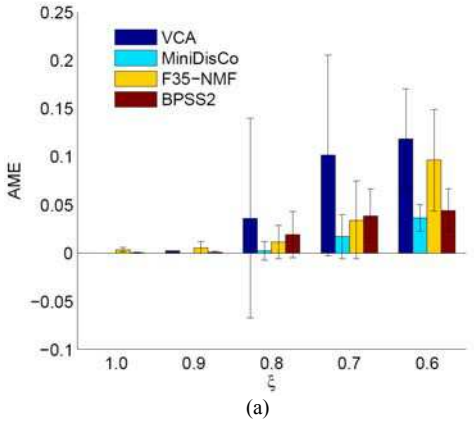
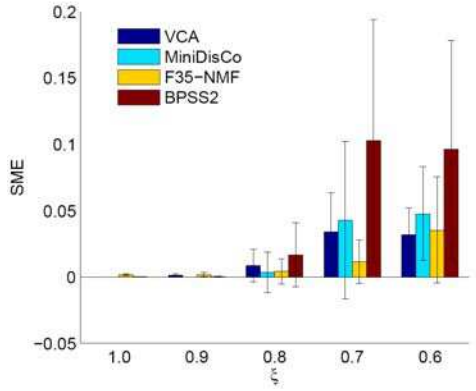
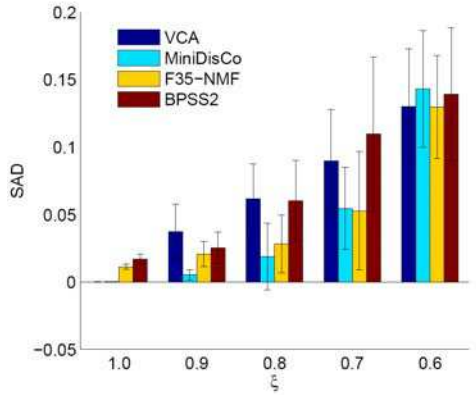


Fig. 4. Performance metrics for different mixing levels ξ. (a) AME; (b) SME; (c) SAD.





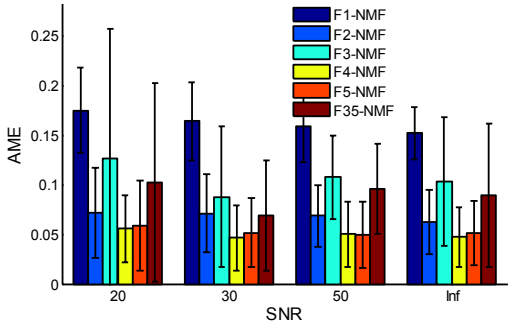
(b)



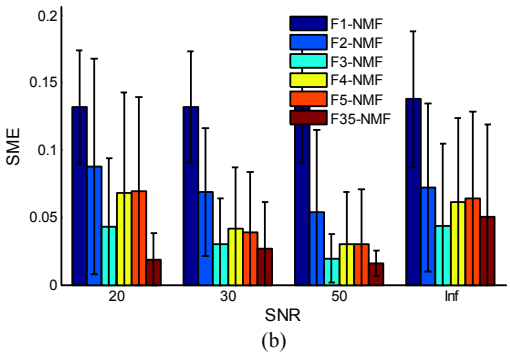
(c)

Fig. 5. Performance comparison for different mixing levels ξ. (a) AME; (b) SME; (c) SAD.

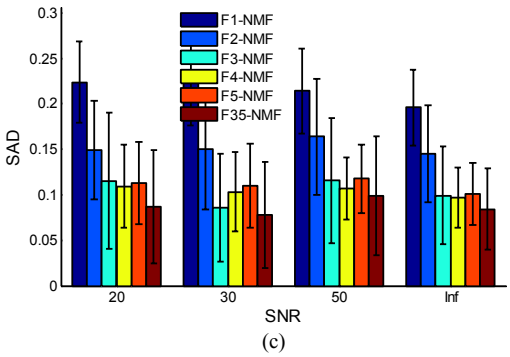
4) The fourth experiment studies the robustness to noise of the considered algorithms. The metric values obtained for various SNR are shown in Fig. 6. The F-NMF algorithms are all based on the RQE minimization, which is optimal for white Gaussian noise. Thus, the performances do not significantly depend on the noise. In accordance with the experiment 3, the F3-NMF and F35-NMF results are not good in AME, but better in the terms of SME and SAD.



(a)



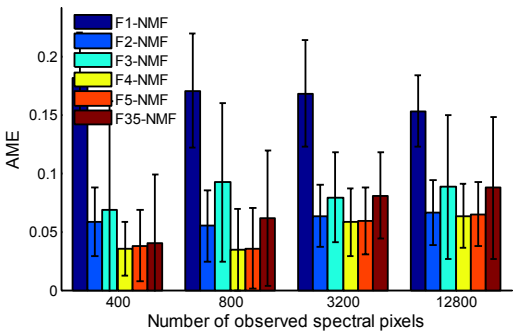
(b)



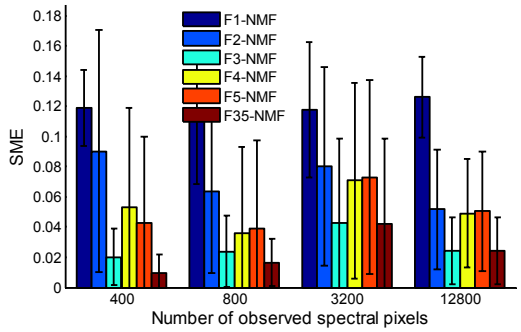
(c)

Fig. 6. Unmixing performances for various noise levels. (a) AME; (b) SME; (c) SAD.

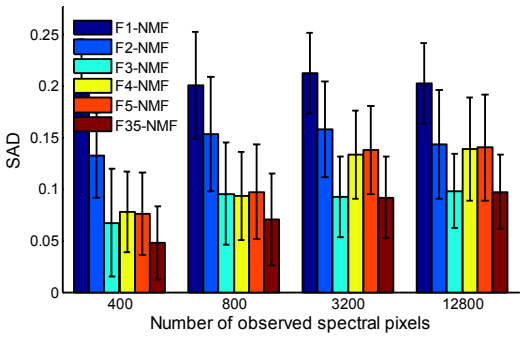
5) It is interesting to study the estimation quality in terms of the data spatial dimensions. Fig. 7 presents the influence of the number of observed spectral pixels. The F-NMF algorithms are both robust to a small number of spectral pixels and a large amount of data. It is interesting to see that a small number of spectral pixels globally improve the performances of the regularized NMF. The F4-NMF and F5-NMF outperform the other algorithms in AME, but the results of F3-NMF and F35-NMF algorithms are better in the terms of SME and SAD. In general, a large data set does not improve the results, so it is more efficient to use a small set of data (400 pixels).



(a)



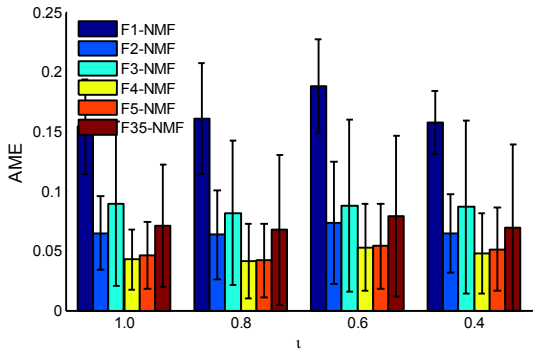
(b)



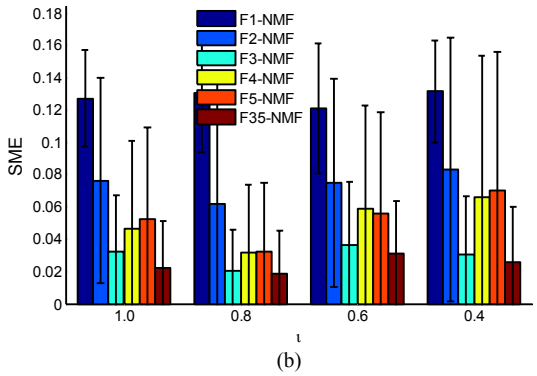
(c)

Fig. 7. Performance of algorithms for different numbers of spectral pixels. (a) AME; (b) SME; (c) SAD.

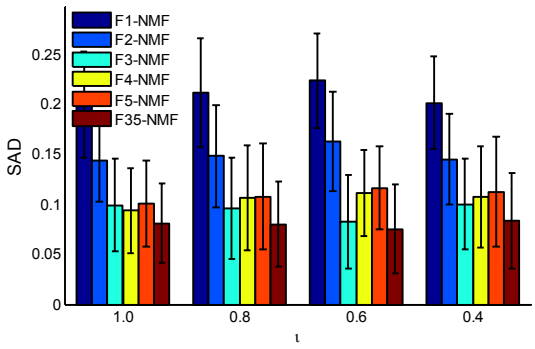
6) This experiment tests the influence of the sparsity parameter ι. The results are presented in Fig. 8. All the algorithms are not very sensitive to the sparsity parameter. The F4-NMF and F5-NMF outperform the other algorithms in AME, and the maximum spatial dispersion constraint brings improvement in SME and SAD.



(a)



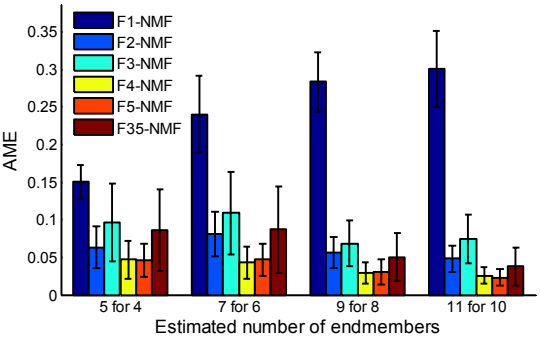
(b)



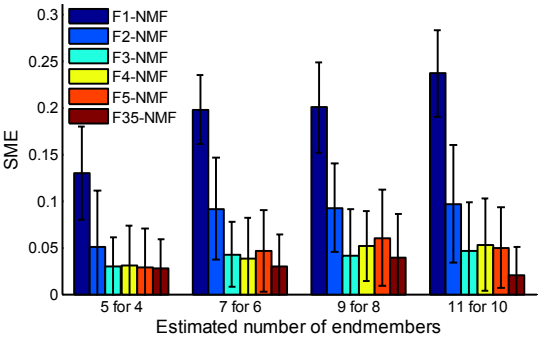
(c)

Fig. 8. Algorithm performance as a function of ι. (a) AME; (b) SME; (c) SAD.

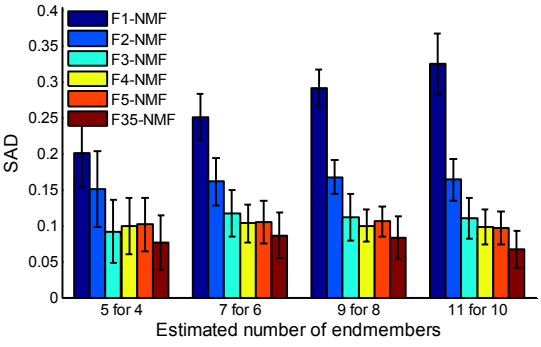
7) Estimating the endmember number J is the first issue of the HSI analysis. On real data, existing methods to estimate J generally overestimate the number [22]. Thus, we study the robustness of the algorithms to an overestimation of J (Fig. 9). Here, we overestimate J by 1. The estimation errors show that constrained F-NMF algorithms are robust to an overestimation of J, while the basic F-NMF is sensitive to the number of endmembers.



(a)



(b)



(c)

Fig. 9. Algorithm performances when J is overestimated by 1. (a) AME; (b) SME; (c) SAD.

The following conclusions can be drawn from these experiments:

1) The optimized algorithm of HALS outperforms PG in convergence speed and efficiency. In [11], poor estimations due to local minimum affect the basic PG-NMF, so the estimated values of SME higher than 0.5 are not included in the statistics. In F-NMF, the estimation performance is much better so all the results of the experiments are considered.
2) The performances of constrained F-NMF are better than the basic NMF, according to all the parameters (J, ξ, SNR, I and ι) and the different performance metrics.
3) The NMF algorithms with minimum spectral dispersion constraint (F4-NMF) and minimum distance constraint (F5-NMF) performs better in AME, while the algorithms with maximum spatial dispersion constraint (F3-NMF and F35-NMF) outperfom the other algorithms in the terms of SME and SAD.
4) The NMF algorithm with combined contraints (F35-NMF) performs better than the algorithm with one constraint (F3-NMF).
5) F-NMF algorithm can improve the unmixing results initialized by VCA.

V. Application on Real Hyperspectral Data

We have applied the five F-NMF algorithms on a hyperspectral scene captured by the AVIRIS sensor. This sensor has a 20m spatial resolutions and a 10nm spectral resolution and acquires 224 spectral bands between 0.4 and 2.5μm. The analyzed reflectance image is a 99×99 pixel selection of the Cuprite geological data. A RGB representation of the scene is shown in Fig. 10. Some spectral bands have been removed due to noise corruption and atmosphere absorption, and only the data of the remaining 188 bands have been used. In this section, we choose SISAL as the compared algorithm because it is able to deal with the unmixing problem without the pure pixel assumption as the NMF algorithms.

It is required to estimate the number of endmembers J before unmixing the image. In this paper, the number of endmembers is determined from the final RQE obtained after convergence for many preliminary experiences, and is set to $J = 11$; however, this value is only an approximation.

Fig. 10. Analyzed scene, a 99×99-pixel region of the Cuprite data.

To improve the algorithm performances, we run the five F-NMF algorithms with the VCA initializations [4] to obtain better local solutions. The estimated endmembers are associated with the closest ones contained in the USGS library in the SAD sense. To evaluate the stability of the algorithms and the ability to find a unique solution, we make 50 runs for each F-NMF algorithm and keep the 11 estimated endmembers at each run. In each run, a new HSI synthetic data is generated with the same parameters (J, I, ξ, ι, SNR), when the endmembers are selected randomly from the library. We should obtain the same 11 identified references in each experiment. However, the results vary in the 50 experiments.

In order to compare the results with a minimum volume based algorithm, we choose SISAL for its good performances on simulated data and its high speed. Note that the F-NMF and SISAL algorithms are all based on the assumption that the endmembers, or at least some of them, are not in the data set. The references identified by F-NMF are presented in Table 2-7 and the results by SISAL in Table 8. The estimated endmembers are identified as the closest library spectra in the sense of SAD. It can be seen from the tables that F3-NMF gives 77 names for a total of 550 possible different answers, whereas the other four F-NMFs give much more references. The top 11 responses of F3-NMF and F35-NMF represent 66.7% and 68% respectively of all the answers. All these results show the stability of F3-NMF, due to the maximum spatial dispersion constraint. From Table 8, we can see that the SISAL identifies 146 names from 550 possible answers, which shows its serious instability. Therefore, the F-NMF algorithms are more stable than SISAL. Otherwise, the mean SAD between the estimated endmembers and the closest references in the library is significantly lower with F-NMF than SISAL, so we can conclude that the F-NMF algorithms are more efficient in endmember identification for difficult real cases.

Fig. 11 and 12 give the estimated endmember spectra and abundance maps by F-35 in one experiment. The endmember spectra resulting from the F35-NMF analysis in one experiment are shown in Fig. 11(a). In this figure, the y-coordinate tick (from $j = 1$ to J) corresponds to zero reflectance of the j^{th} endmember. The associated spectral endmembers are the closest library spectra (Fig. 11(b)) in the sense of SAD. Note that 3, 7, 8 and 10 spectra are all identified as Kaolin/Smect KLF508 85%K, whose proportion is 26% the first in Table 7. This is the reason of the low identification dispersion of F35-NMF in 50 runs. The estimated abundance maps are given in Fig. 12, where the maximum abundance value ξ_j of each endmember j is high due to the maximum spatial dispersion constraint.

TABLE 2
References identified by F1-NMF

USGS reference name	Percent (%)	SAD
Kaolin/Smect KLF508 85%K	21.273	5.0403
Andradite HS111.3B	7.6364	4.8941
Andradite WS488	3.8182	4.5956
Lepidolite NMNH105543	3.4545	6.9935
Goethite WS219 (limonite)	3.0909	6.8237
Sphene HS189.3B	3.0909	4.3982
Kaolin/Smect H89-FR-5 30K	3.0909	3.8384
Barite HS79.3B	2.5455	4.0984
Rectorite RAr-1	2.5455	3.5563
Kaolin/Smect KLF511 12%K	2.3636	3.1599
Almandine WS477	2.3636	5.8007
Sum of the top 11	55.273	
83 names	100	5.1017

TABLE 3
References identified by F2-NMF

USGS reference name	Percent (%)	SAD
Kaolin/Smect KLF508 85%K	15.455	5.1444
Andradite HS111.3B	6.7273	4.5232
Andalusite NMNHR17898	5.4545	3.5174
Richterite HS336.3B	3.6364	3.6131
Andradite WS487	3.2727	3.8255
Kaolin/Smect H89-FR-5 30K	3.0909	3.7755
Kaolin/Smect KLF511 12%K	2.5455	3.3961
Lepidolite NMNH105543	2.3636	7.2303
Rectorite RAr-1	2.3636	3.5788
Nontronite NG-1.a	2.3636	4.3183
Goethite WS219 (limonite)	2.1818	6.8177
Sum of the top 11	49.455	
98 names	100	5.9825

TABLE 4
REFERENCES IDENTIFIED BY F3-NMF

USGS reference name	Percent (%)	SAD
Kaolin/Smect KLF508 85%K	26.545	4.5345
Andradite HS111.3B	9.0909	4.2029
Andradite WS487	6	3.6531
Andalusite NMNHR17898	5.0909	3.6367
Richterite HS336.3B	4.1818	3.2513
Kaolin/Smect H89-FR-5 30K	3.8182	3.295
Illite IL105 (1Md)	2.9091	3.8568
Montmorillonite+Illi CM37	2.5455	3.1543
Barite HS79.3B	2.5455	3.2624
Kaolin/Smect KLF511 12%K	2.1818	2.719
Rectorite RAr-1	1.8182	3.877
Sum of the top 11	66.727	
77 names	100	5.1216

TABLE 5
REFERENCES IDENTIFIED BY F4-NMF

USGS reference name	Percent (%)	SAD
Kaolin/Smect KLF508 85%K	16.364	5.1932
Andradite HS111.3B	6.7273	4.513
Andalusite NMNHR17898	5.8182	3.5355
Richterite HS336.3B	4	3.674
Kaolin/Smect H89-FR-5 30K	3.4545	3.829
Andradite WS487	3.2727	3.8056
Kaolin/Smect KLF511 12%K	2.7273	3.3183
Lepidolite NMNH105543	2.5455	7.6882
Rectorite RAr-1	2.3636	3.518
Goethite WS219 (limonite)	2.1818	6.8385
Nontronite NG-1.a	2.1818	4.2891
Sum of the top 11	51.636	
97 names	100	5.1216

TABLE 6
REFERENCES IDENTIFIED BY F5-NMF

USGS reference name	Percent (%)	SAD
Kaolin/Smect KLF508 85%K	16	5.1622
Andradite HS111.3B	6.7273	4.5837
Andalusite NMNHR17898	5.8182	3.524
Richterite HS336.3B	4.1818	3.6714
Andradite WS487	3.2727	3.806
Kaolin/Smect H89-FR-5 30K	3.2727	3.7812
Lepidolite NMNH105543	2.5455	7.4252
Rectorite RAr-1	2.5455	3.6228
Nontronite NG-1.a	2.3636	4.3196
Kaolin/Smect KLF511 12%K	2.3636	3.3328
Goethite WS219 (limonite)	2.1818	6.828
Sum of the top 11	51.273	
97 names	100	5.8524

TABLE 7
REFERENCES IDENTIFIED BY F35-NMF

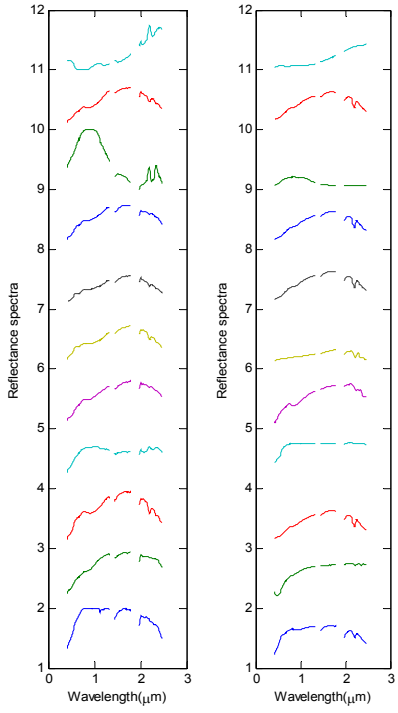
USGS reference name	Percent (%)	SAD
Kaolin/Smect KLF508 85%K	26.545	4.437
Andradite HS111.3B	11.818	4.3172
Andradite WS487	6.7273	3.6406
Andalusite NMNHR17898	5.6364	3.6947
Kaolin/Smect KLF511 12%K	3.2727	2.8078
Montmorillonite+Illi CM37	2.9091	3.2787
Richterite HS336.3B	2.7273	3.4012
Illite IL105 (1Md)	2.3636	3.784
Barite HS79.3B	2	3.0921
Perthite HS415.3B	2	2.774
Sphene HS189.3B	2	3.6569
Sum of the top 11	68	
80 names	100	4.9158

TABLE 8
REFERENCES IDENTIFIED BY SISAL

USGS reference name	Percent (%)	SAD
Kaolin/Smect KLF508 85%K	5.8182	7.1877
Acmite NMNH133746	4.9091	87.33
Hornblende_Fe HS115.3B	4	24.591
Desert_Varnish GDS141	3.6364	6.5823
Limonite HS41.3	3.4545	29.357
Hematite FE2602	3.4545	4.6658
Lepidolite NMNH105543	3.2727	8.6374
Mordenite+Clinopt. GDS151	2.7273	3.1797
Almandine WS477	2.3636	5.2999
Rutile HS137.3B	2.1818	5.5928
Almandine HS114.3B	2.1818	27.247
Sum of the top 11	38	
146 names	100	15.808

We compare the references identified by F-NMF and SISAL with the available ground truth of the Cuprite scene from the website [23]. In Tables 2-8, the identified results, which appear in the ground-truth list, are highlighted. Each of the considered algorithms only can identify two or three ground-truth minerals. In particular, F-NMF and SISAL algorithms all detect Kaolin, and Goethite is detected by F1-NMF, F2-NMF, F4-NMF and F5-NMF. In addition, Nontronite is detected by F2-NMF, F4-NMF and F5-NMF, and F35-NMF detects Montmorillonite and SISAL detects Hematite. The identified results illustrate the difficulty of the unmixing problem for real data. Three reasons can be explained as follows:

1) The analyzed Cuprite data is only a selection of the whole scene, which holds 18 endmembers; thus, the unmixing results are also incomplete.
2) It is difficult to find the right spectra in the considered library with a huge amount of references (500). Some priori knowledge should be used to reduce the number of references before the comparison.
3) We use a linear mixing model in this paper, but the radiative transfer is always non-linear in real scene [10].
4) It is subjective to identify the endmembers with SAD. A more robust method for identification should make the decision jointly with several criteria. Moreover, the variability of real spectra has made their identification from library more difficult.



(a) (b)

Fig. 11. References identified in one experiment. (a) Endmember spectra estimated with F35-NMF; (b) Associated USGS references.

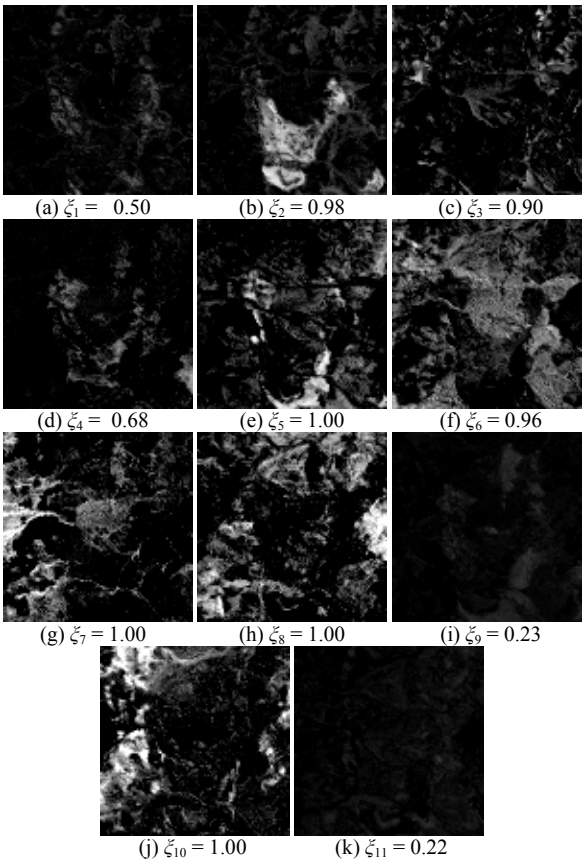


Fig. 12. Estimated abundance maps by F35-NMF. The maximum abundance value ξ_j of each endmember j is presented with the corresponding map.

Finally, it is important to analyze the computation time of the F-NMF algorithms. Under the Matlab environment and 3GHz CPU, the computation times for an iteration of the F-NMF algorithms with the real data (99×99 pixels) are shown in Table 9. It is clear that the algorithms with spectral constraints (F4, F5, F35) are more time-consuming due to the computation of matrix inversion. If the number of iteration is more than a thousand, the running of any F-NMF algorithm will cost a few minutes. In the case of computation cost, geometrical methods (e.g. VCA and SISAL) are fast and efficient, while the NMF-based methods are always slow.

TABLE 9
Computation time of F-NMF algorithms (s)

F1	F2	F3	F4	F5	F35
0.23	0.24	0.24	0.40	0.40	0.40

VI. Conclusion

In this paper, we have proposed a NMF-based hyperspectral unmixing algorithm with flexible constraints, including the sum-to-unity contraint, the maximum spatial dispersion constraint, the minimum spectral dispersion constraint and the minimum distance constraint. The optimization scheme is based on the HALS, whose convergence speed outperforms that of PG. The resulting algorithm, called F-NMF, is experimentally tested with different constraints. The estimation accuracy shows that the F-NMF works stably in all experiments, overcoming the estimation instability of PG-NMF. In particular, the F-NMF algorithms are robust to high number of endmembers, low SNR, low number of observed pixels and overestimation of the number of endmembers.

The F-NMF algorithms seem to be effective in the estimation of abundance maps, since they consider the sum-to-unity and maximum spatial dispersion constraints. The identified references of real data by F-NMF seem more stable and reliable than geometrical method like SISAL. However, the identified results of real data are unsatisfied so the identification method needs further investigation to impove the results.

Acknowledgment

The authors would like to thank S. Moussaoui for letting us do the benchmarks with the code of his BPSS algorithms.

References

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

[2] J. Theiler, D. Lavenier, N. Harvey, S. Perkins, and J. Szymanski, "Using blocks of skewers for faster computation of pixel purity index," presented at the SPIE Int. Conf. Optical Science and Technology, 2000.

[3] M. E. Winter, "N-findr: An algorithm for fast autonomous spectral end-member determination in hyperspectral data," in Proc. SPIE Conf. Imaging Spectrometry V, 1999, pp. 266–275.

[4] J. Nascimento and J. Dias, "Vertex Component Analysis: A Fast Algorithm to Unmix Hyperspectral Data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43. no. 4, pp.898-908, Apr. 2005.

[5] J. Li and J. M. Bioucas-Dias, "Minimum Volume Simplex Analysis: A Fast Algorithm to Unmix Hyperspectral Data," in *Proc. IEEE IGARSS*, Boston, 2008.

[6] T. H. Chan, C. Y. Chi, Y. M. Huang and W. K. Ma, "A Convex Analysis-Based Minimum-Volume Enclosing Simplex Algorithm for Hyperspectral Unmixing", *IEEE Trans. Signal Process.,* vol. 57. no. 11, pp.4418-4432, Nov. 2009.

[7] J. M. Bioucas-Dias, "A variable splitting augmented Lagrangian approach to linear spectral unmixing," in *Proc. WHISPERS,* Grenoble, France, Aug. 2009.

[8] S. Moussaoui, D. Brie, A. Mohammad-Djafari and C. Carteret, "Separation of Non-negative Mixture of Non-negative Sources using a Bayesian Approach and MCMC Sampling". IEEE Trans. Signal Processing, vol. 54(11), pp. 4133-4145, 2006.

[9] N. Dobigeon, S. Moussaoui, J.-Y. Tourneret, and C. Carteret, "Bayesian separation of spectral sources under non-negativity and full additivity constraints," Signal Processing, vol. 89, no. 12, pp. 2657–2669, Dec. 2009.

[10] Schmidt F., Schmidt A., Tréguier E., et al. "Implementation Strategies for Hyperspectral Unmixing Using Bayesian Source Separation." *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 11, p. 4003-4013, Mar. 2010.

[11] Daniel D. Lee and H. Sebastian Seung. "Learning the parts of objects by non-negative matrix factorization." *Nature*, vol. 401 (6755): 788–791, 1999.

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

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

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

[15] Y. Yu, S. Guo and W. D. Sun, "Minimum distance constrained nonnegative matrix factorization for the endmember extraction of hyperspectral images," in Proc. SPIE, vol. 6790, 679015, MIPPR, Wuhan, China, Nov., 2007.

[16] A. Huck, M. Guillaume and J. Blanc-Talon, "Minimum Dispersion

Constrained Nonnegative Matrix Factorization to Unmix Hyperspectral Data," *IEEE Trans. Geosci. Remote Sens.*, vol. 48. no. 6, pp.2590-2602, May. 2010.
[17] A. Huck and M. Guillaume, "Robust hyperspectral data unmixing with spatial and spectral regularized NMF," in *Proc. WHISPERS,* Reykjavik, Iceland, Aug. 2010.
[18] Cichocki, A., Phan, A.H. and Caiafa, C. "Flexible HALS algorithms for sparse non-negative matrix/tensor factorization." *IEEE Workshop on Machine Learning for Signal Processing*, vol. 4, no. 4, p. 73-78, 2008.
[19] N. Gillis and F. Glineur, "Using Underapproximations for Sparse Nonnegative Matrix Factorization," *Pattern Recognition*, vol. 43 no. 4, pp. 1676-1687, Jun. 2010.
[20] N. Gillis and F. Glineur, "Nonnegative Factorization and The Maximum Edge Biclique Problem," CORE Discussion paper 2008/64, 2008.
[21] N.-D. Ho, "Nonnegative Matrix Factorization - Algorithms and Applications," PhD thesis, Universite catholique de Louvain, 2008.
[22] C.-I Chang and Q. Du, "Estimation of number of spectrally distinct signal sources in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 3, p. 608-619, Mar. 2004.
[23] [Oneline].Available:http://speclab.cr.usgs.gov/PAPERS/cuprite.gr.truth.1992/swayze.1992.html

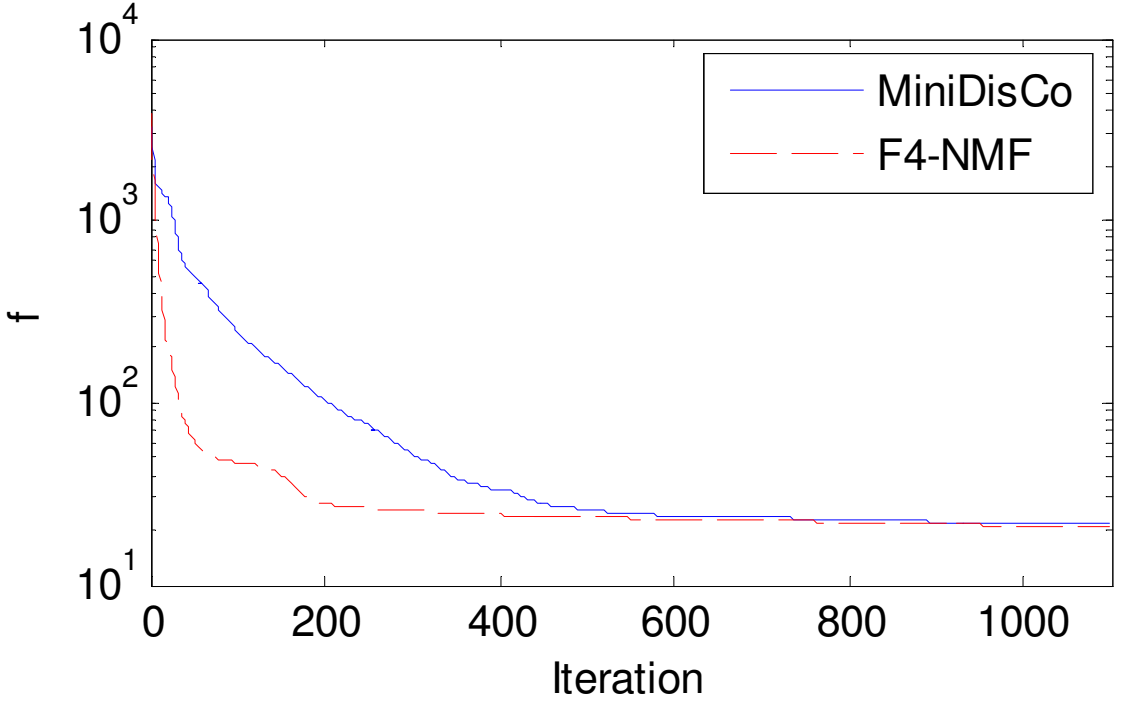

MiniDisCo
F4-NMF
f
Iteration
10^4
10^3
10^2
10^1
0
200
400
600
800
1000

Figure 1

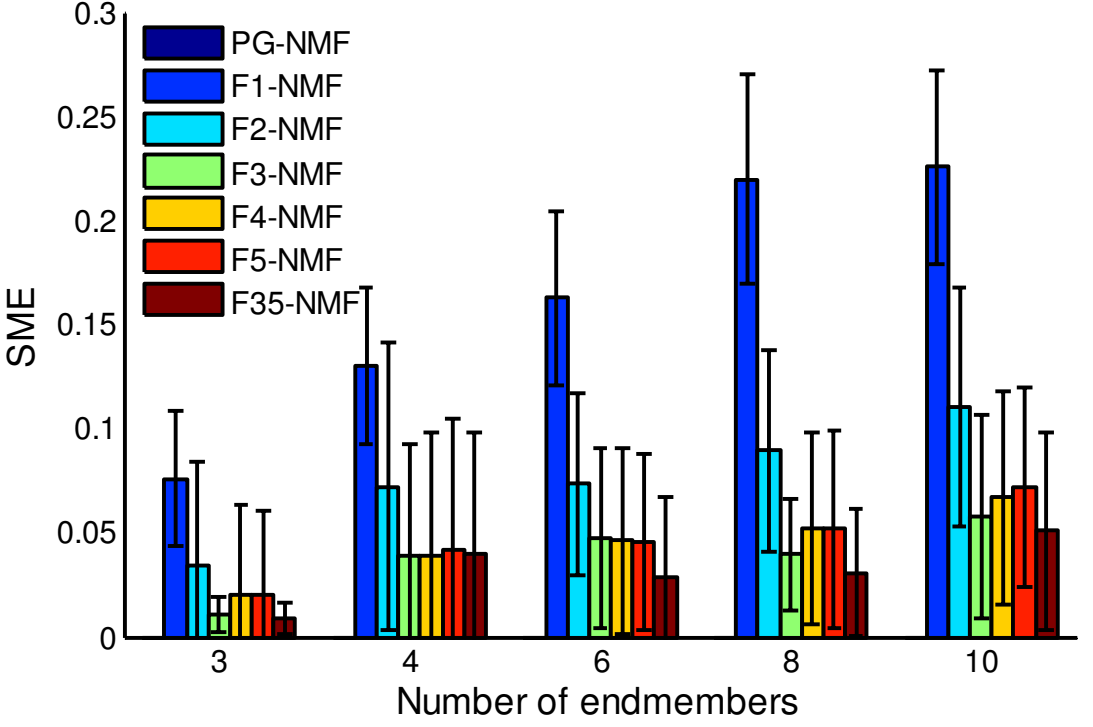

PG-NMF
F1-NMF
F2-NMF
F3-NMF
F4-NMF
F5-NMF
F35-NMF
SME
0.3
0.25
0.2
0.15
0.1
0.05
0
3
4
6
8
10
Number of endmembers

Figure 2

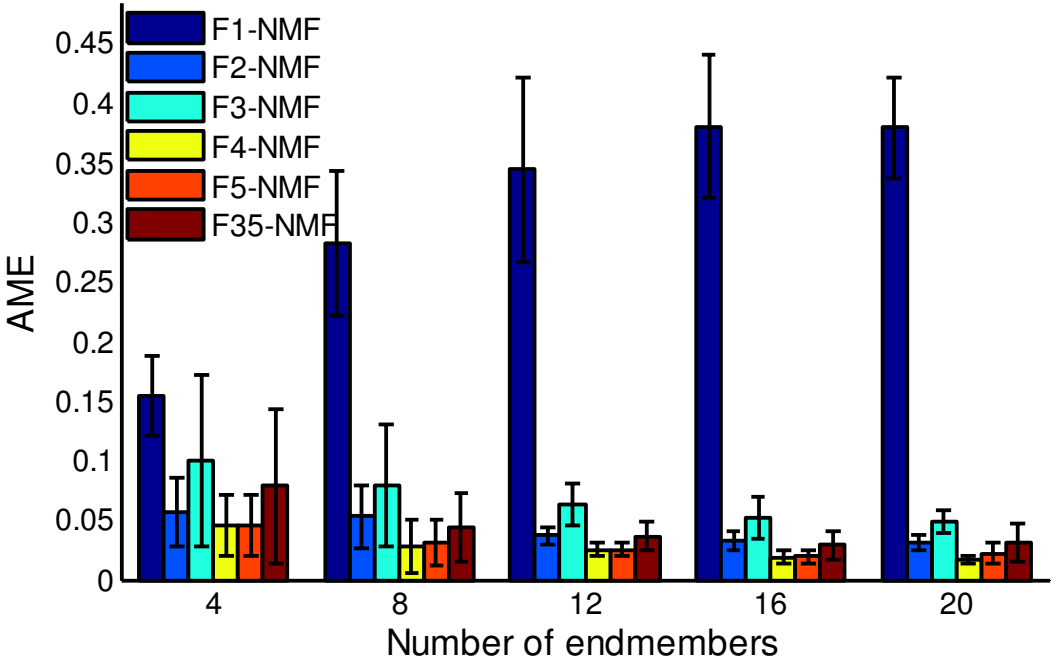

F1-NMF
F2-NMF
F3-NMF
F4-NMF
F5-NMF
F35-NMF
AME
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
4
8
12
16
20
Number of endmembers

Figure 3

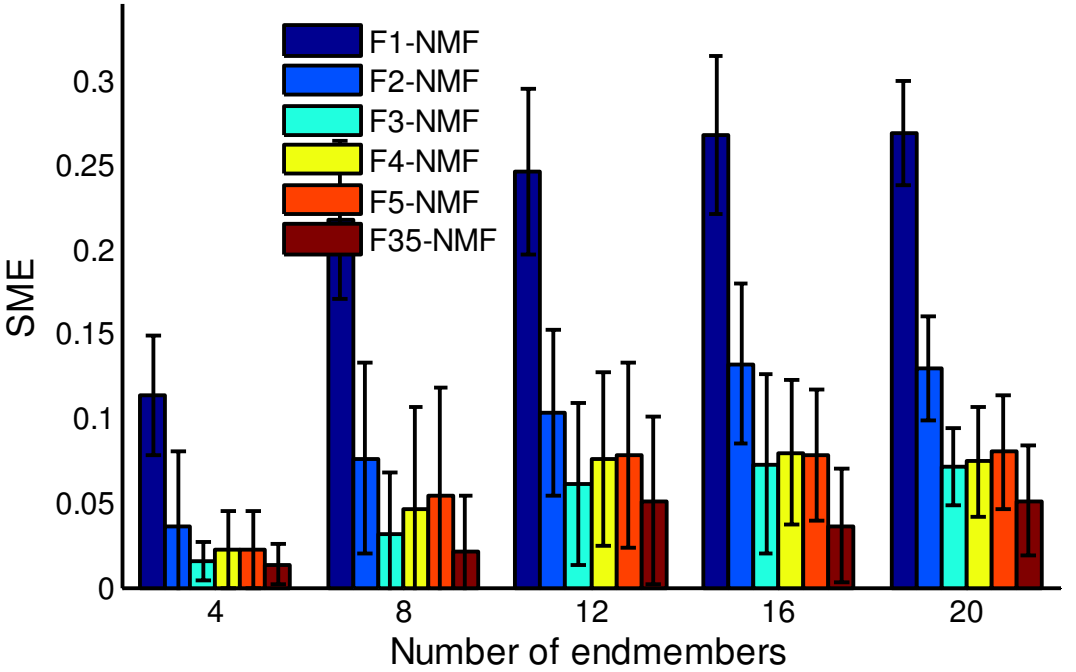

F1-NMF
F2-NMF
F3-NMF
F4-NMF
F5-NMF
F35-NMF
SME
0
0.05
0.1
0.15
0.2
0.25
0.3
4
8
12
16
20
Number of endmembers

Figure 4

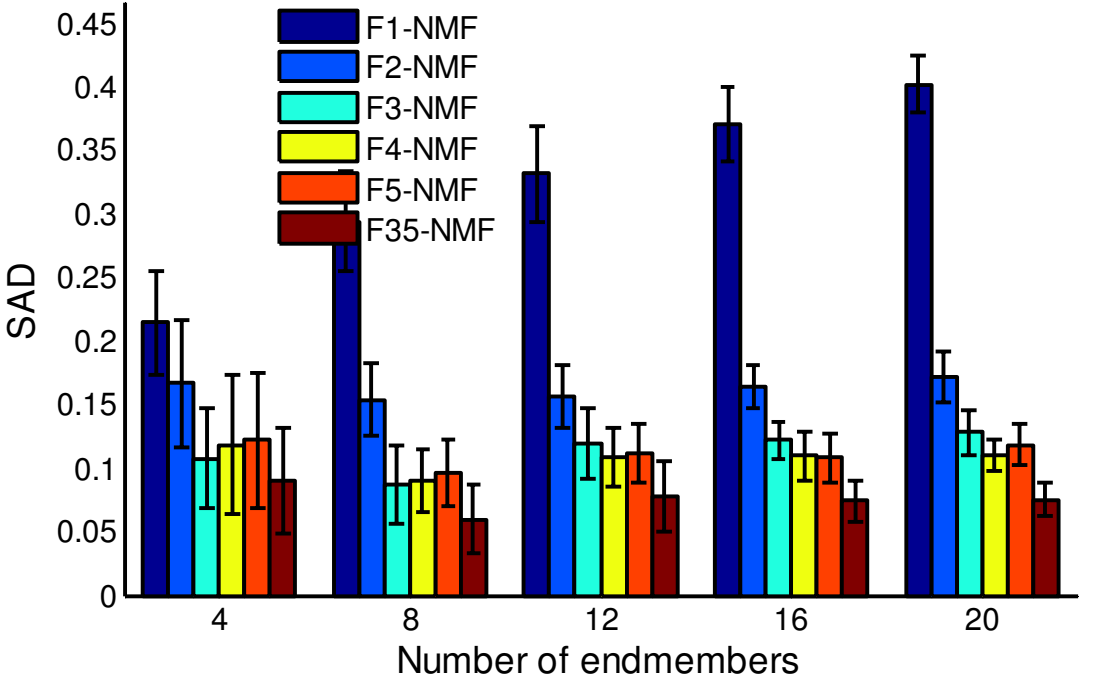

F1-NMF
F2-NMF
F3-NMF
F4-NMF
F5-NMF
F35-NMF
SAD
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
4
8
12
16
20
Number of endmembers

Figure 5

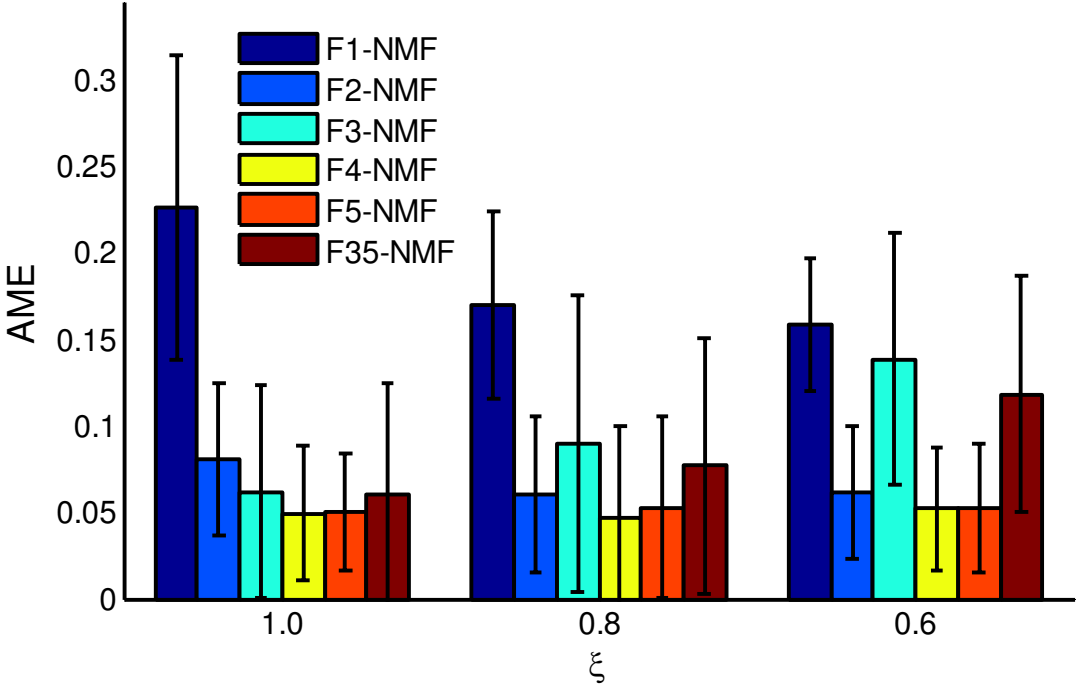

F1-NMF
F2-NMF
F3-NMF
F4-NMF
F5-NMF
F35-NMF
AME
0
0.05
0.1
0.15
0.2
0.25
0.3
1.0
0.8
0.6
ξ

Figure 6

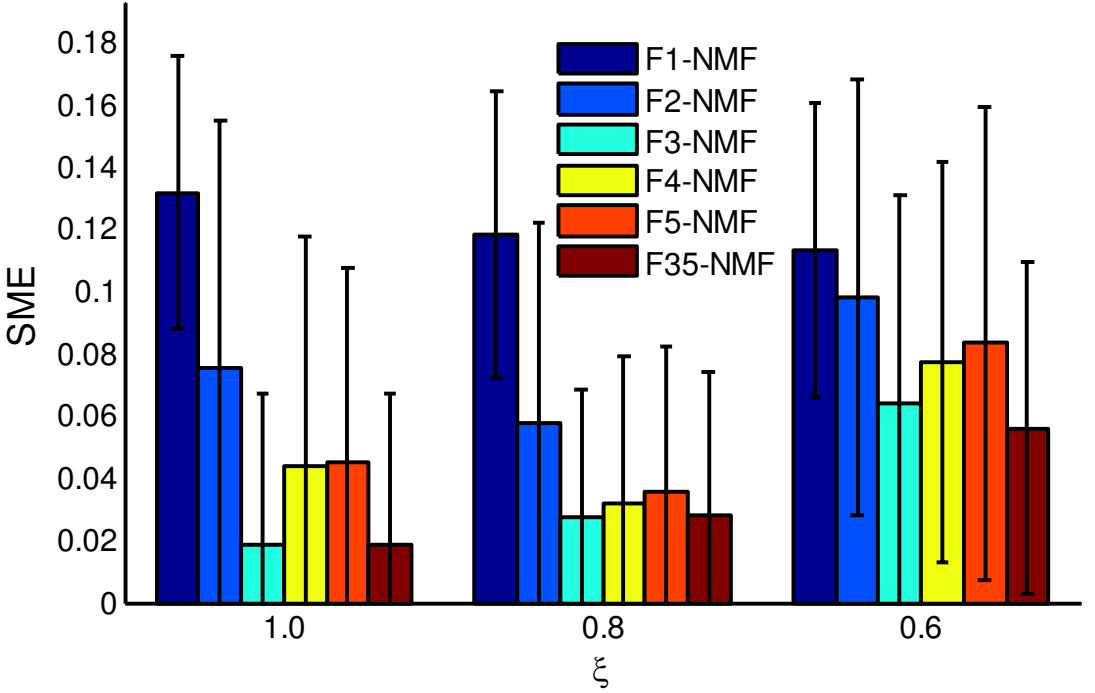

F1-NMF
F2-NMF
F3-NMF
F4-NMF
F5-NMF
F35-NMF
SME
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
1.0
0.8
0.6
ξ

Figure 7

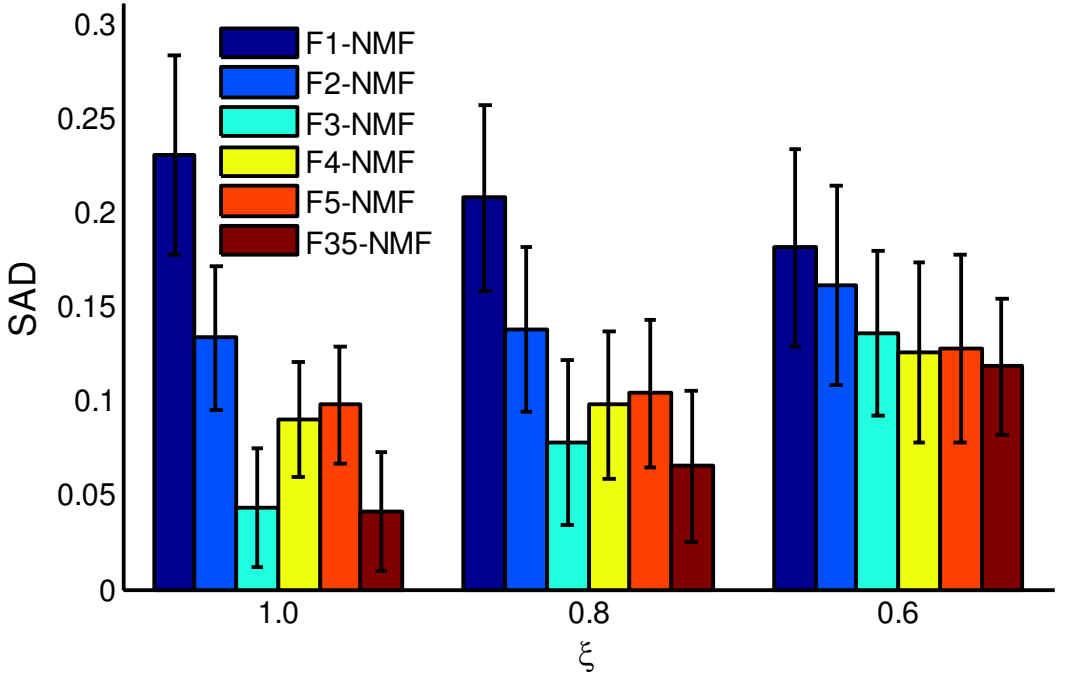

F1-NMF
F2-NMF
F3-NMF
F4-NMF
F5-NMF
F35-NMF
SAD
0.3
0.25
0.2
0.15
0.1
0.05
0
1.0
0.8
0.6
ξ

Figure 8

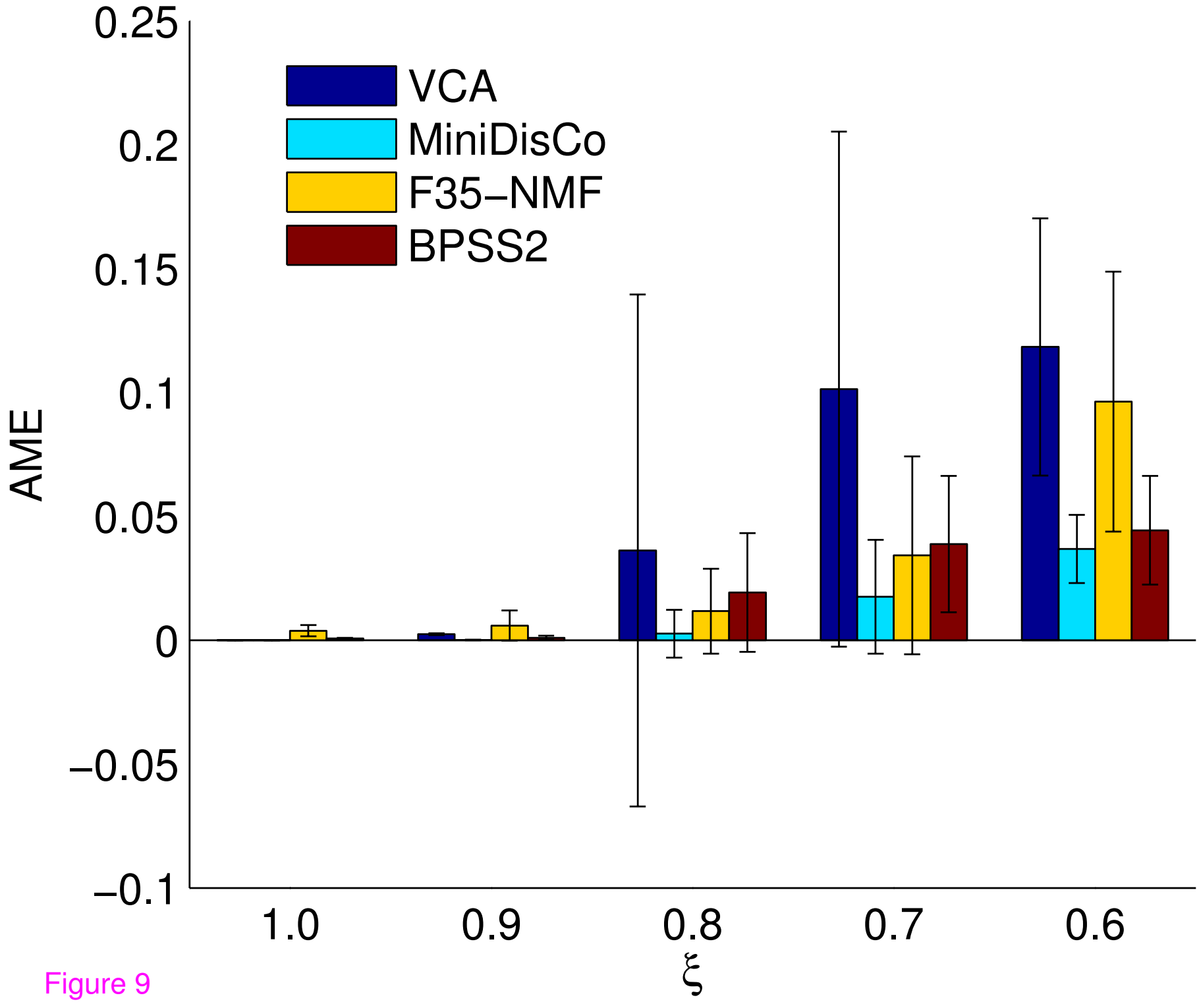

0.25
0.2
0.15
0.1
0.05
0
−0.05
−0.1
AME
VCA
MiniDisCo
F35−NMF
BPSS2
1.0
0.9
0.8
0.7
0.6
ξ

Figure 9

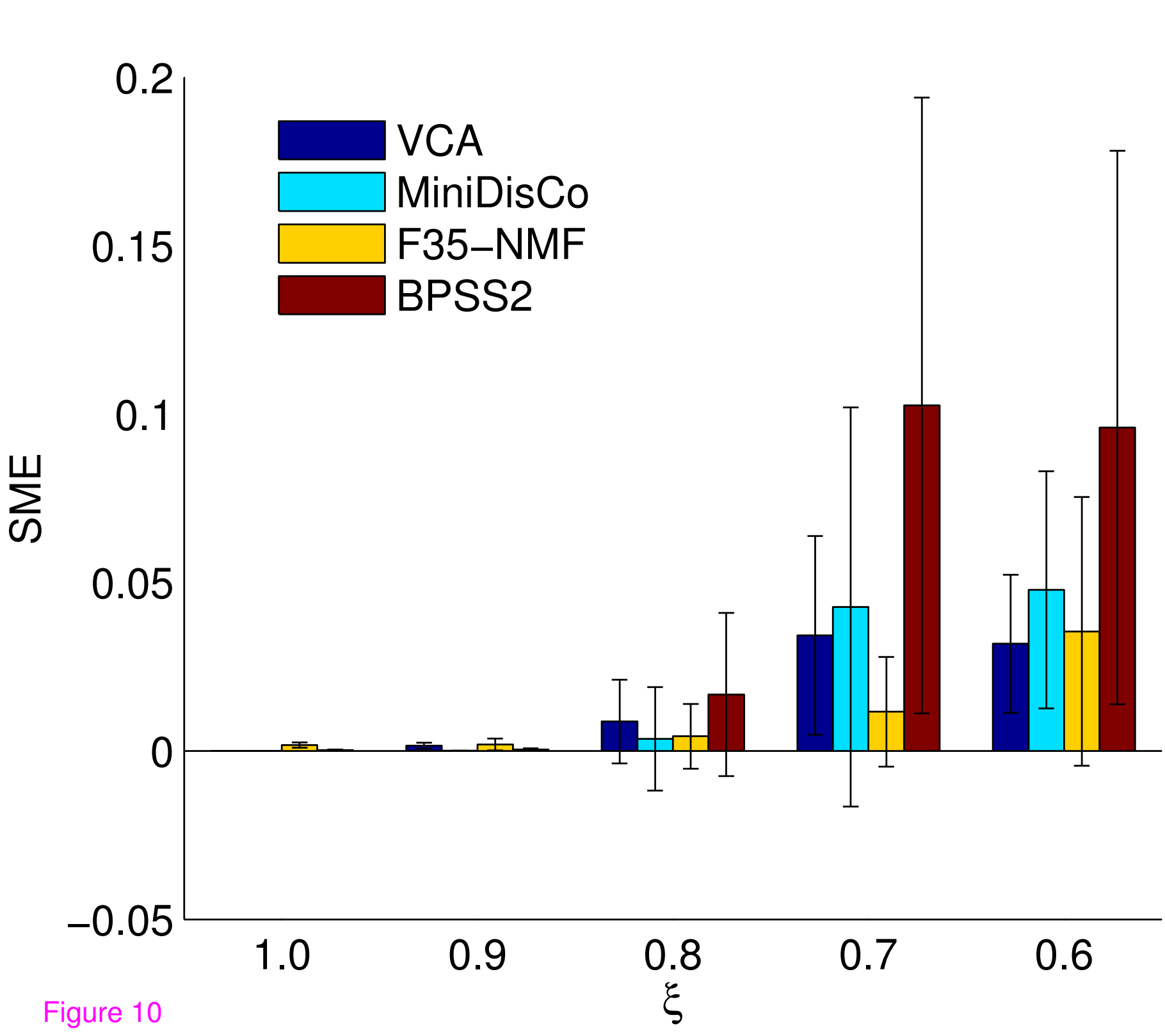

0.2
0.15
0.1
0.05
0
−0.05
SME
VCA
MiniDisCo
F35−NMF
BPSS2
1.0
0.9
0.8
0.7
0.6
ξ

Figure 10

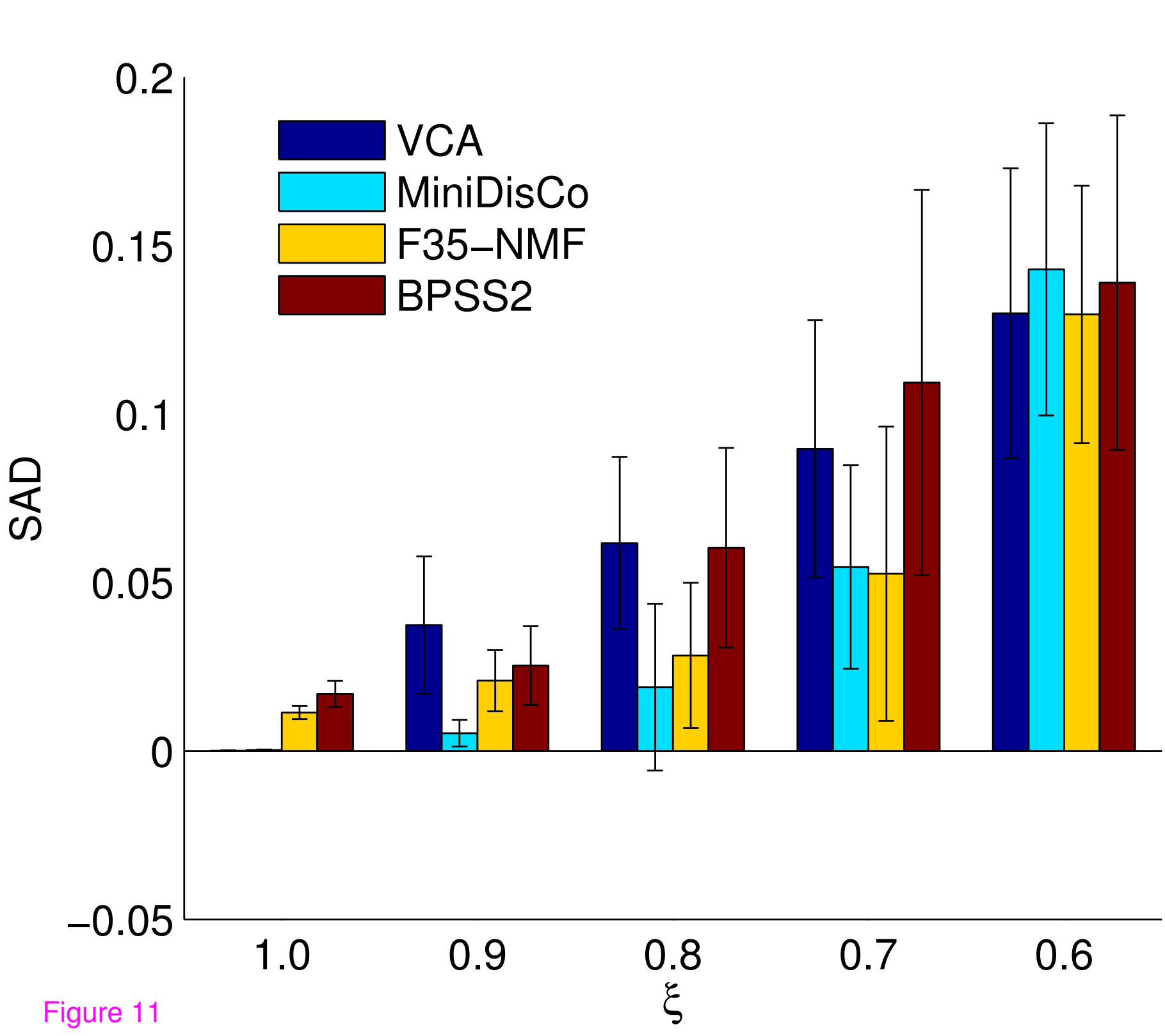

0.2
0.15
0.1
0.05
0
−0.05
SAD
VCA
MiniDisCo
F35−NMF
BPSS2
1.0
0.9
0.8
0.7
0.6
ξ

Figure 11

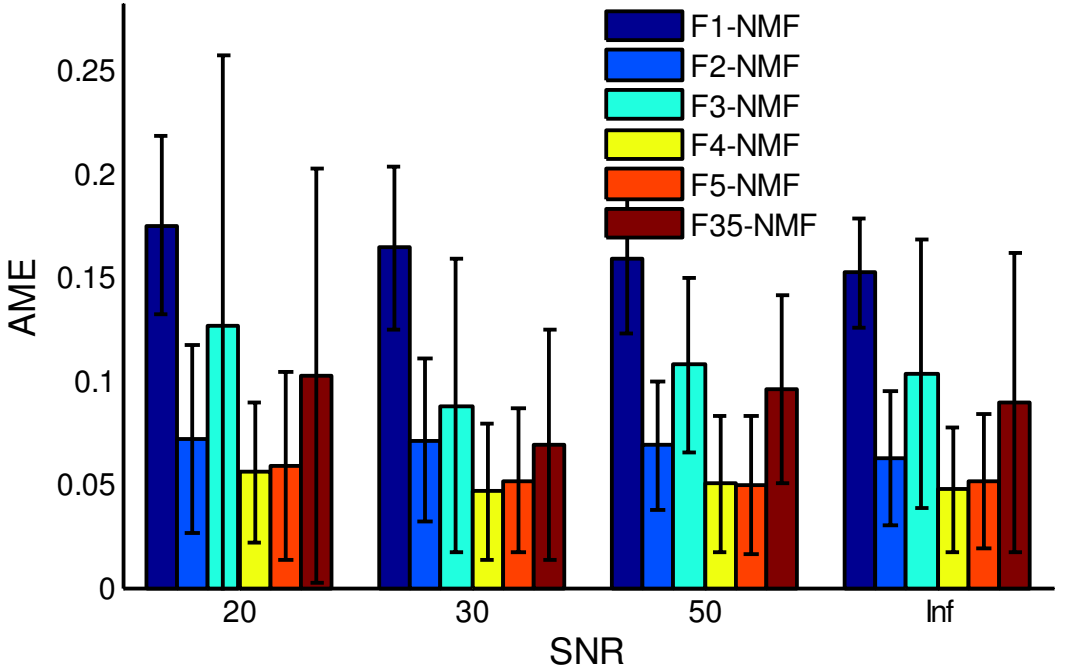

F1-NMF
F2-NMF
F3-NMF
F4-NMF
F5-NMF
F35-NMF
AME
0.25
0.2
0.15
0.1
0.05
0
20
30
50
Inf
SNR

Figure 12

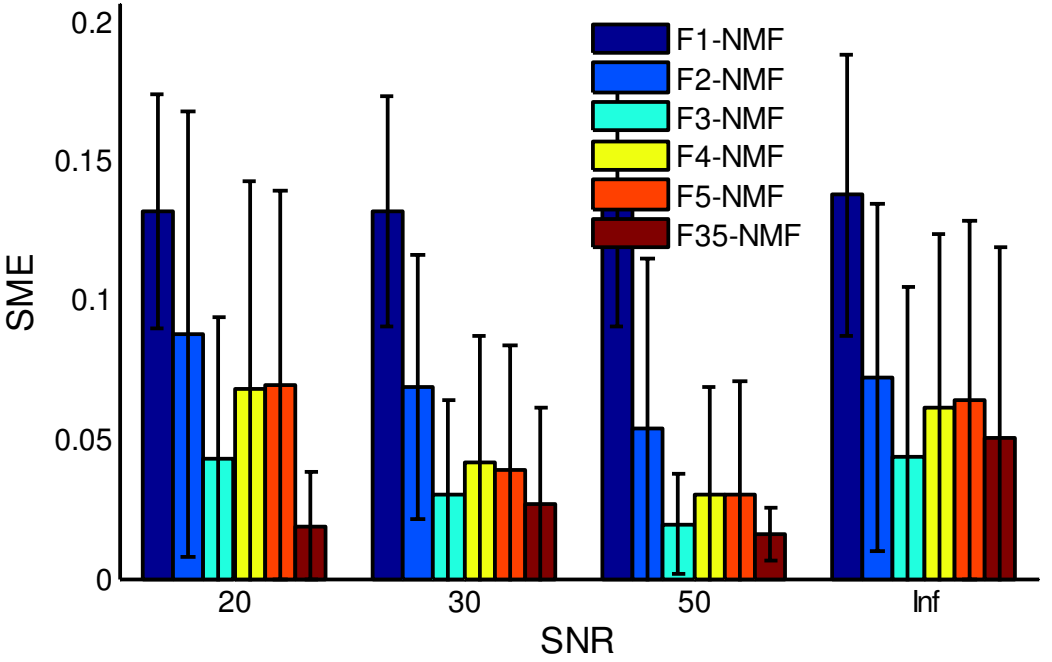

0.2
0.15
0.1
0.05
0
SME
F1-NMF
F2-NMF
F3-NMF
F4-NMF
F5-NMF
F35-NMF
20
30
50
Inf
SNR

Figure 13

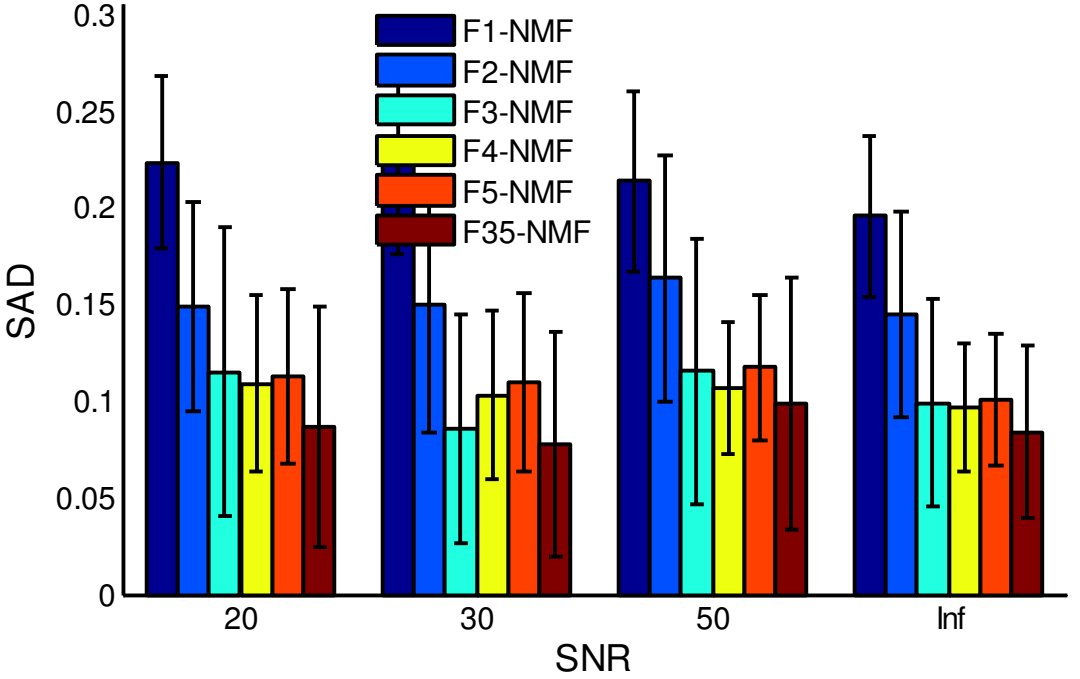

F1-NMF
F2-NMF
F3-NMF
F4-NMF
F5-NMF
F35-NMF
SAD
0.3
0.25
0.2
0.15
0.1
0.05
0
20
30
50
Inf
SNR

Figure 14

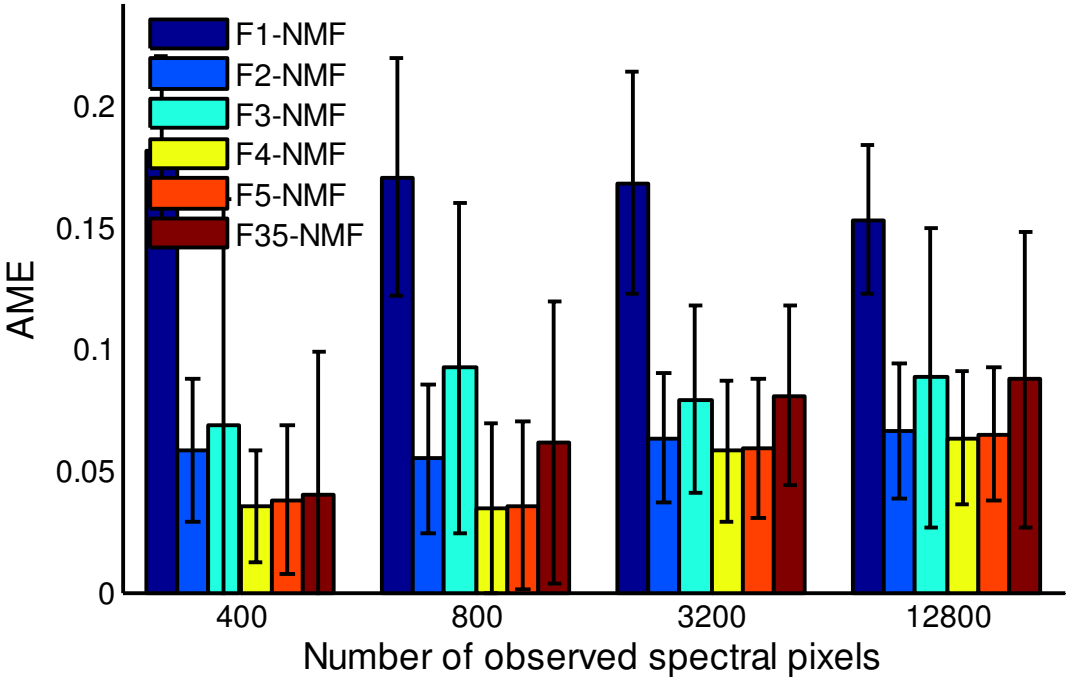

F1-NMF
F2-NMF
F3-NMF
F4-NMF
F5-NMF
F35-NMF
AME
0
0.05
0.1
0.15
0.2
400
800
3200
12800
Number of observed spectral pixels

Figure 15

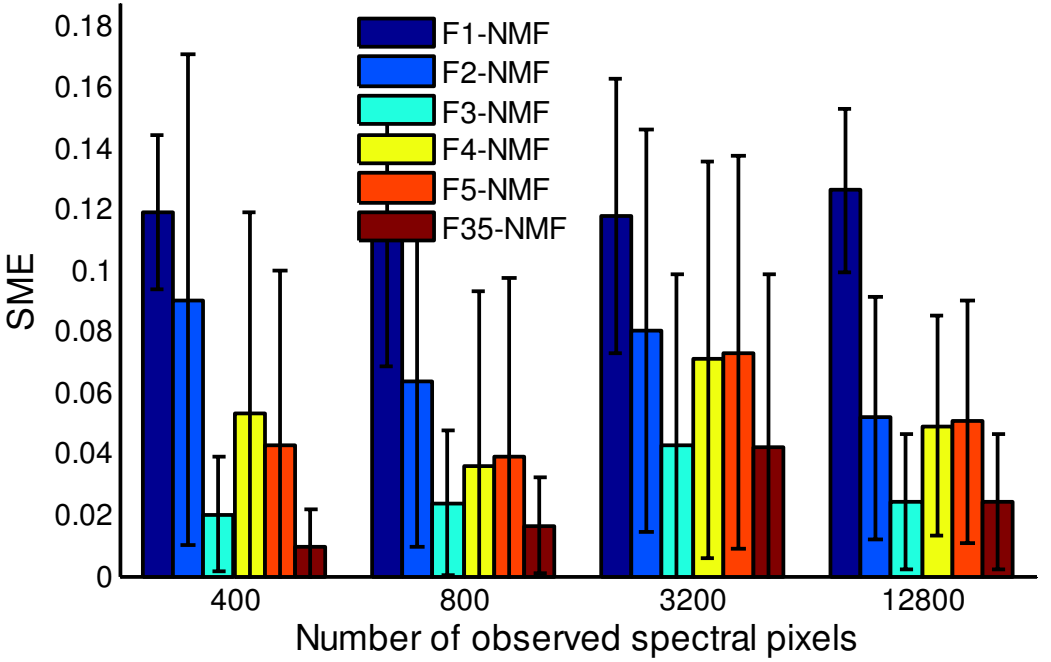

F1-NMF
F2-NMF
F3-NMF
F4-NMF
F5-NMF
F35-NMF
SME
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
400
800
3200
12800
Number of observed spectral pixels

Figure 16

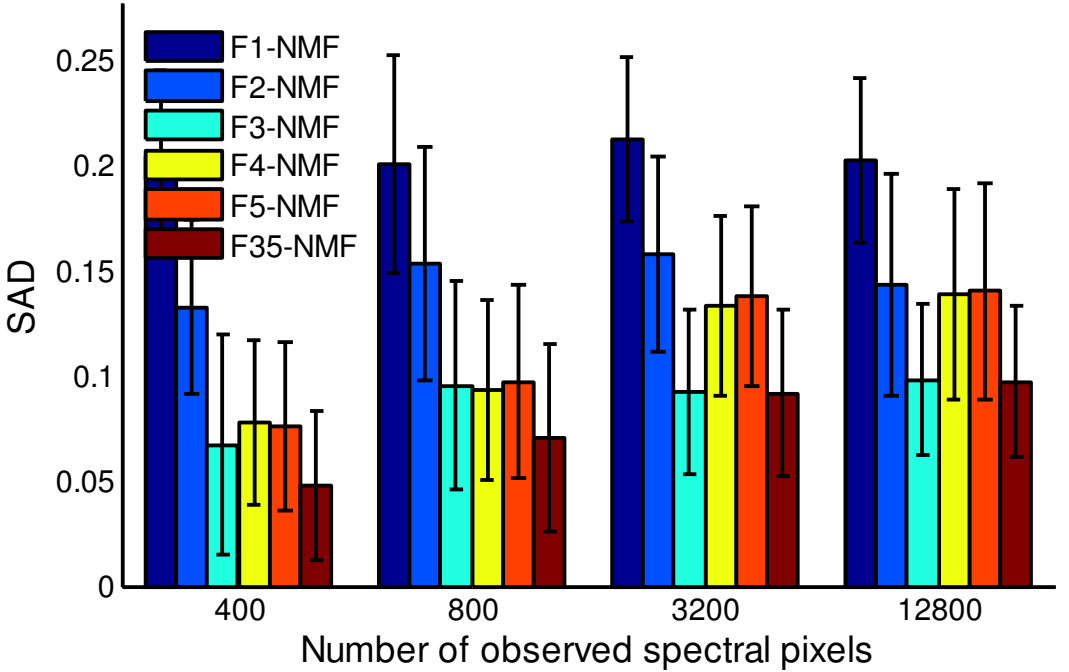

F1-NMF
F2-NMF
F3-NMF
F4-NMF
F5-NMF
F35-NMF
SAD
0.25
0.2
0.15
0.1
0.05
0
400
800
3200
12800
Number of observed spectral pixels

Figure 17

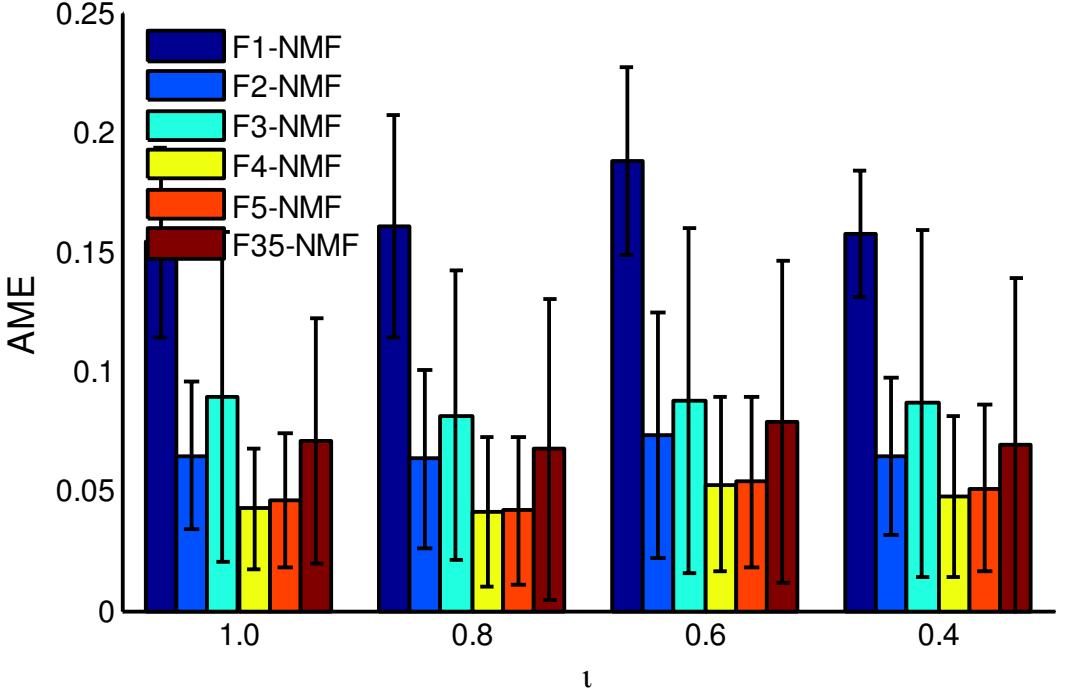

0.25
0.2
0.15
0.1
0.05
0
AME
F1-NMF
F2-NMF
F3-NMF
F4-NMF
F5-NMF
F35-NMF
1.0
0.8
0.6
0.4
ι

Figure 18

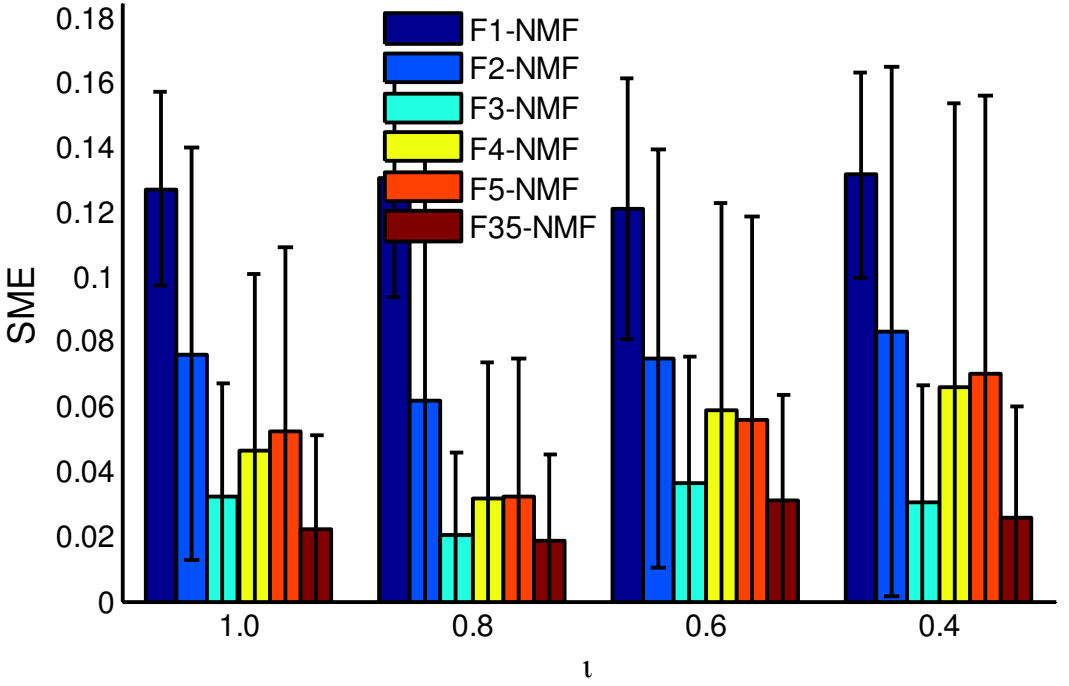

F1-NMF
F2-NMF
F3-NMF
F4-NMF
F5-NMF
F35-NMF
SME
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
1.0
0.8
0.6
0.4
ι

Figure 19

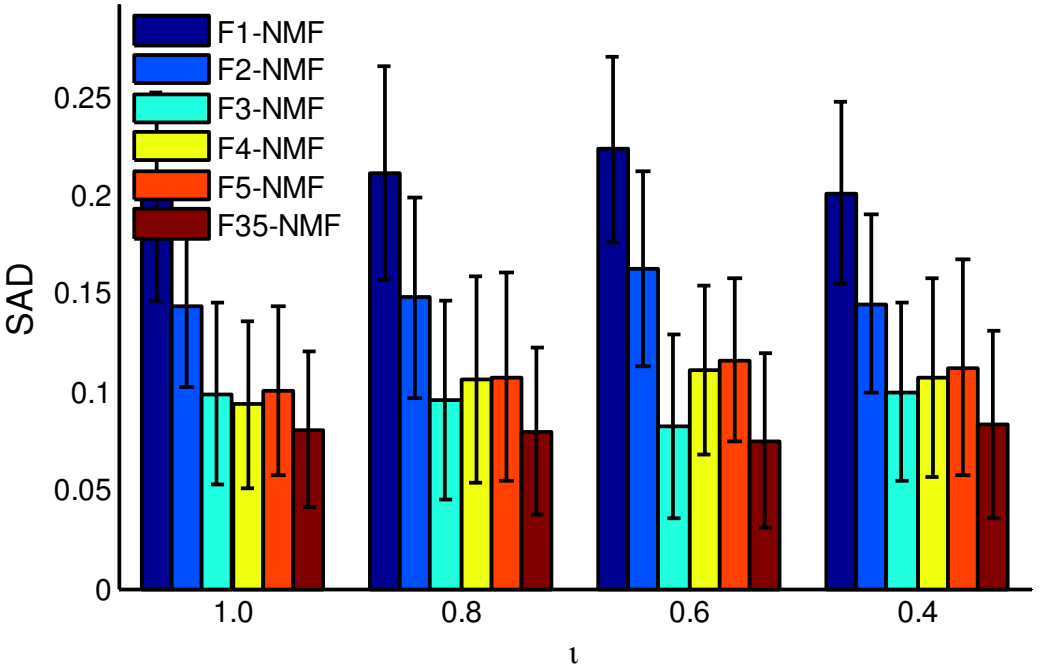

F1-NMF
F2-NMF
F3-NMF
F4-NMF
F5-NMF
F35-NMF
SAD
0
0.05
0.1
0.15
0.2
0.25
1.0
0.8
0.6
0.4
ι

Figure 20

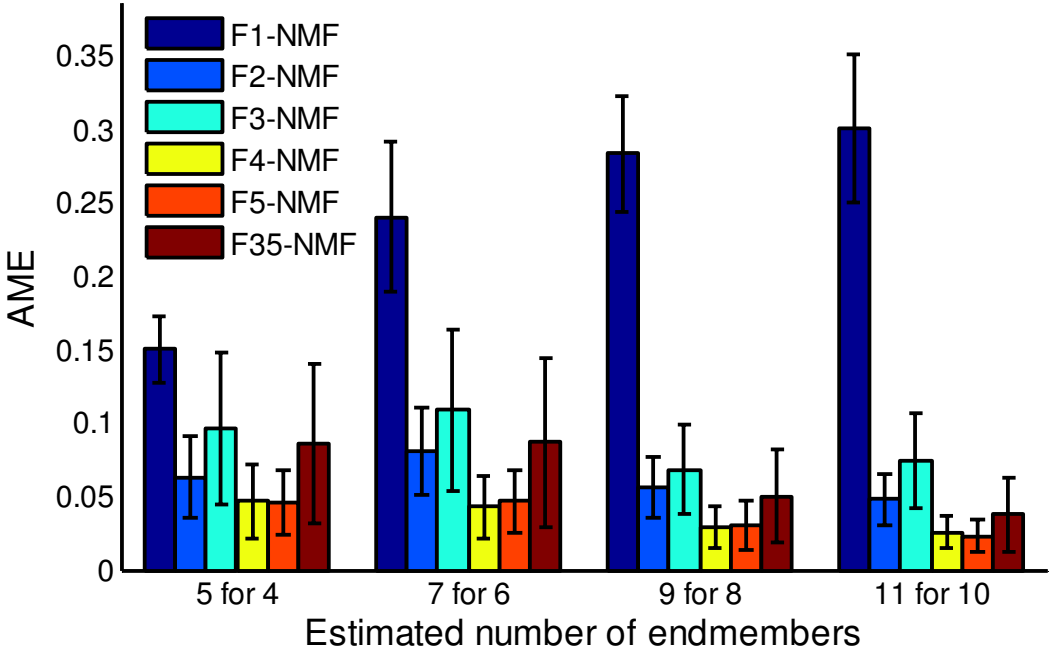

F1-NMF
F2-NMF
F3-NMF
F4-NMF
F5-NMF
F35-NMF
AME
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
5 for 4
7 for 6
9 for 8
11 for 10
Estimated number of endmembers

Figure 21

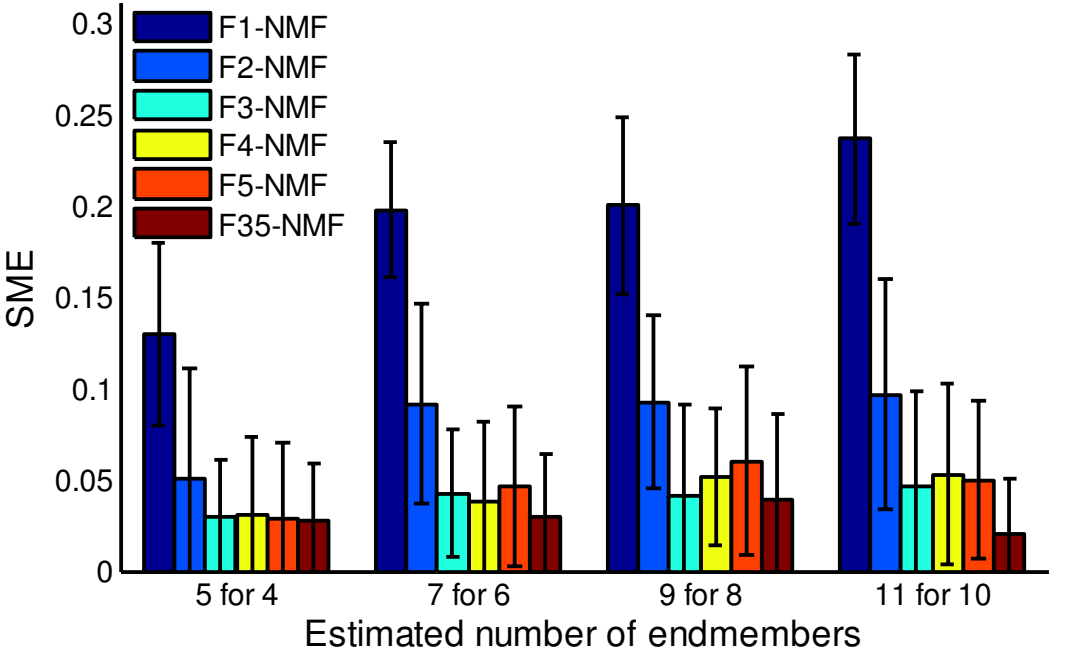

F1-NMF
F2-NMF
F3-NMF
F4-NMF
F5-NMF
F35-NMF
SME
0.3
0.25
0.2
0.15
0.1
0.05
0
5 for 4
7 for 6
9 for 8
11 for 10
Estimated number of endmembers

Figure 22

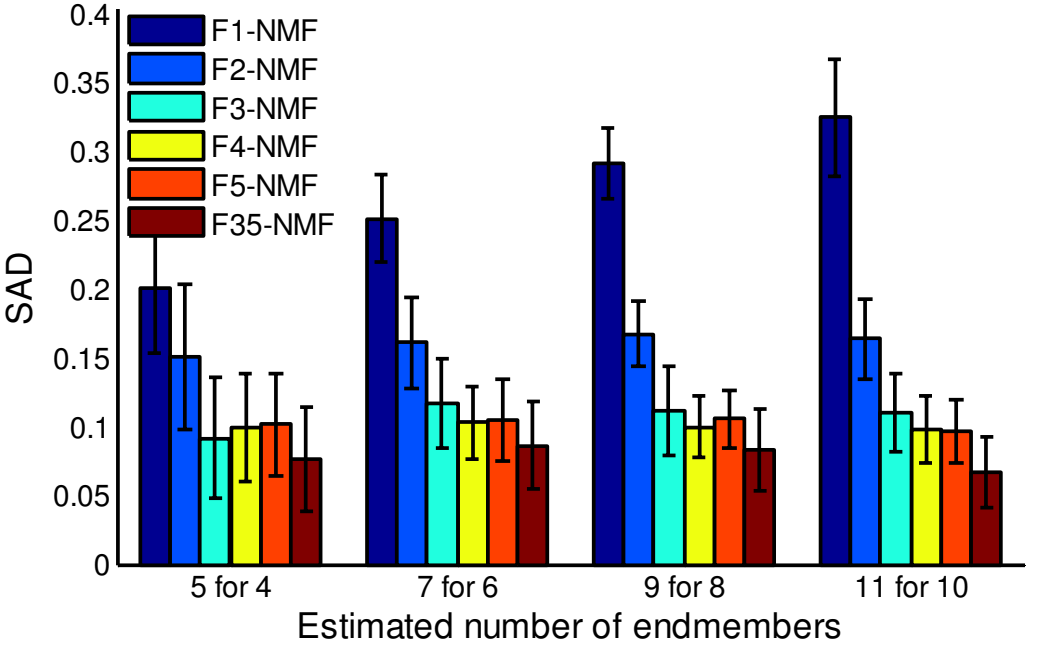

F1-NMF
F2-NMF
F3-NMF
F4-NMF
F5-NMF
F35-NMF
SAD
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
5 for 4
7 for 6
9 for 8
11 for 10
Estimated number of endmembers

Figure 23

Figure 2

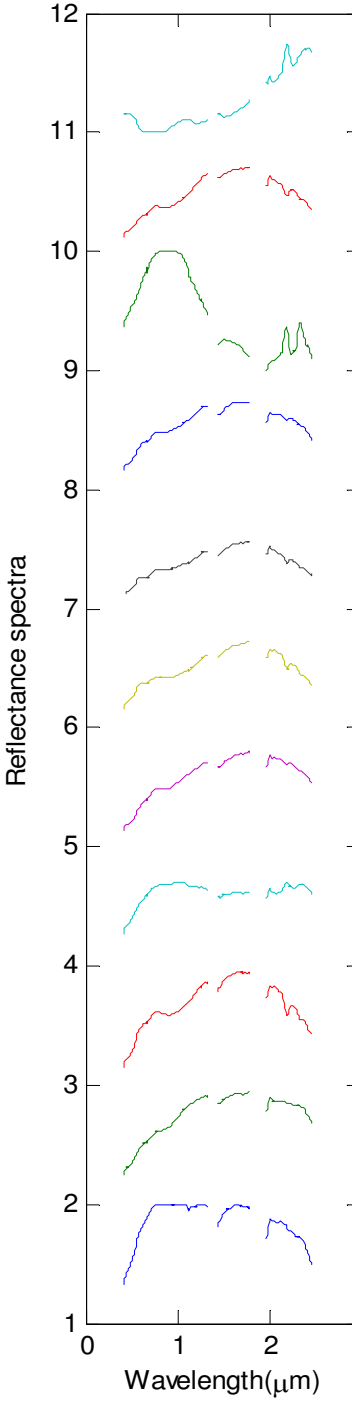


Figure 25

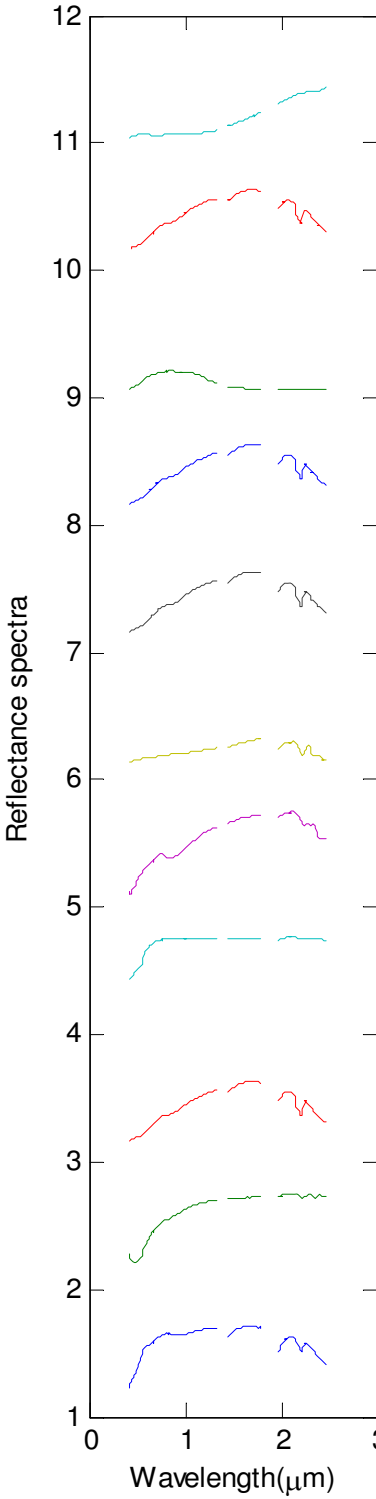


Figure 26

Figure 27

Figure 28

Figure 29

Figure 30


Figure 31


Figure 32

Figure 33


Figure 34

Figure 35

Figure 36

Figure 37