A robust test for nonlinear mixture detection in hyperspectral images

Yoann Altmann, Nicolas Dobigeon, Jean-Yves Tourneret, José Carlos M. Bermudez

A ROBUST TEST FOR NONLINEAR MIXTURE DETECTION IN HYPERSPECTRAL IMAGES

Y. Altmann, N. Dobigeon and J-Y. Tourneret*

University of Toulouse
IRIT-ENSEEIHT
Toulouse, France

J.C.M. Bermudez†

Federal University of Santa Catarina
Florianópolis, SC, Brazil

ABSTRACT

This paper studies a pixel by pixel nonlinearity detector for hyperspectral image analysis. The reflectances of linearly mixed pixels are assumed to be a linear combination of known pure spectral components (endmembers) contaminated by additive white Gaussian noise. Nonlinear mixing, however, is not restricted to any prescribed nonlinear mixing model. The mixing coefficients (abundances) satisfy the physically motivated sum-to-one and positivity constraints. The proposed detection strategy considers the distance between an observed pixel and the hyperplane spanned by the endmembers to decide whether that pixel satisfies the linear mixing model (null hypothesis) or results from a more general nonlinear mixture (alternative hypothesis). The distribution of this distance is derived under the two hypotheses. Closed-form expressions are then obtained for the probabilities of false alarm and detection as functions of the test threshold. The proposed detector is compared to another nonlinearity detector recently investigated in the literature through simulations using synthetic data. It is also applied to a real hyperspectral image.

***Index Terms*—** Nonlinearity detection, Hyperspectral images, Linear mixing model.

1. INTRODUCTION

Identifying the macroscopic materials present in a scene as well as their proportions in each pixel of this scene is of prime interest when analyzing hyperspectral images. Most spectral unmixing (SU) strategies assume that the pixel reflectances are linear combinations of the endmembers [1–5]. However, as explained in [6], the linear mixing model (LMM) can be inappropriate for some hyperspectral images, such as those containing sand, trees or vegetation areas. Nonlinear mixing models provide an interesting alternative for overcoming the inherent limitations of the LMM. Some nonlinear models have been proposed in the literature to handle specific kinds of nonlinearity [7, 8]. For instance, the bidirectional reflectance-based model studied in [9] has been introduced for intimate mixtures. The bilinear models recently studied in [10–15] mainly focus on scattering effects observed in images with relief such as vegetation or urban areas.

Even if the consideration of nonlinear effects in hyperspectral images can provide more accurate results in terms of endmember and abundance identification, nonlinear models generally require a higher computational complexity than approaches based on the LMM. Consequently, it is clearly of interest to identify the linearly mixed pixels in an image (which can be more easily analyzed) and to apply more specific unmixing methods only to nonlinearly mixed pixels. A pixel by pixel nonlinearity detector based on a polynomial post-nonlinear mixing model (PPNMM) was recently introduced in [8]. This detector assumed a specific nonlinear mixing model under the alternative hypothesis. A statistical test was then proposed based on the asymptotical properties of the estimators of the model parameters under the linear and nonlinear hypotheses. However, assuming a post-nonlinear mixing model under the alternative hypothesis can be too restrictive and can yield inaccurate detection results when the actual mixing does not obey the specified model. This paper addresses the problem of determining whether an observed pixel of an hyperspectral image is a linear function of endmembers or results from a generic nonlinear mixing. Differently from [8], the proposed nonlinearity detector is based only on the LMM associated with the null hypothesis. This makes the detector more robust to the nonlinearity that defines the alternative hypothesis. Under the null hypothesis, the linear mixture of the endmembers belongs to a low-dimensional hyperplane. Hence, we propose to design a statistical test based on the distance between the observed pixel and that hyperplane for deciding between the null and alternative hypotheses. As in [8], we assume that the endmembers contained in the image have been estimated by a geometric endmember extraction algorithm (EEA).

This paper is organized as follows. Section 2 introduces the linear and nonlinear models used for hyperspectral image unmixing under the null and alternative hypotheses. Section 3 derives the distribution of the squared distance between the observed pixel and the low-dimensional subspace containing linear mixtures of the endmembers. Section 4 derives two statistical tests for nonlinearity detection depending on the prior knowledge about the noise variance (that can be either known or unknown). Some simulation results conducted on synthetic and real data are shown in Section 5. Conclusions and future works are finally reported in Section 6.

2. MIXING MODELS

Let $\mathbf{y}$ be the $L \times 1$ pixel observed at L different spectral bands. The LMM assumes that $\mathbf{y}$ results from a mixture of R known endmembers $\mathbf{m}_1, \ldots, \mathbf{m}_R$ as follows

$$\mathbf{y} = \mathbf{M}\boldsymbol{a} + \mathbf{e} \tag{1}$$

where $\mathbf{M} = [\mathbf{m}_1, \ldots, \mathbf{m}_R]$ is the $L \times R$ endmember matrix, the elements a_r of $\boldsymbol{a} = [a_1, \ldots, a_R]^T$ are the proportions of each endmember in the mixture and $\mathbf{e}$ is an $L \times 1$ independent white Gaussian vector such that $\mathbf{e} \sim \mathcal{N}\left(\mathbf{0}_L, \sigma^2 \mathbf{I}_L\right)$. The elements of the abundance

*Part of this work has been supported by Direction Generale de l'Armement, French Ministry of Defence, and by the Hypanema ANR Project ANR Project n°ANR- 12-BS03-003.

†This work has been supported in part by Capes - Proc. 3367-11-8.

vector $\boldsymbol{a}$ satisfy the physical positivity and sum-to-one constraints

$$\sum_{r=1}^{R} a_r = 1, \;\; a_r \geq 0, \forall r \in \{1, \ldots, R\}. \tag{2}$$

Consider now the hyperplane $\mathcal{H}$ defined by

$$\mathcal{H} : \left\{ \boldsymbol{z} \middle| \boldsymbol{z} = \mathbf{M}\boldsymbol{a}, \sum_{r=1}^{R} a_r = 1 \right\}. \tag{3}$$

In the noise-free case, $\mathcal{H}$ lies in an $(R-1)$-dimensional subspace embedding all observations distributed according to the LMM.

We also consider a general nonlinear mixing model as follows

$$\mathbf{y} = \mathbf{M}\boldsymbol{a} + \boldsymbol{\mu} + \mathbf{e} \tag{4}$$

where $\boldsymbol{\mu}$ is an $L \times 1$ deterministic vector that does not belong to $\mathcal{H}$, i.e., $\boldsymbol{\mu} \notin \mathcal{H}$ and $\boldsymbol{a}$ satisfies the constraints (2). Note that $\boldsymbol{\mu}$ can be a nonlinear function of the endmember matrix $\mathbf{M}$ and/or the abundance vector $\boldsymbol{a}$ and should be denoted as $\boldsymbol{\mu}(\mathbf{M}, \boldsymbol{a})$. However, the arguments $\mathbf{M}$ and $\boldsymbol{a}$ are omitted in this paper for brevity.

Given an observation vector $\mathbf{y}$, we formulate the detection of nonlinear mixtures as the following binary hypothesis testing problem:

$$\begin{cases} H_0 : & \mathbf{y} \text{ is distributed according to (1)} \\ H_1 : & \mathbf{y} \text{ is distributed according to (4).} \end{cases} \tag{5}$$

Using the statistical properties of the noise $\mathbf{e}$, we obtain $\mathrm{E}[\mathbf{y}|H_0] = \mathbf{M}\boldsymbol{a} \in \mathcal{H}$ whereas $\mathrm{E}[\mathbf{y}|H_1] = \mathbf{M}\boldsymbol{a} + \boldsymbol{\mu} \notin \mathcal{H}$. As a consequence, it makes sense to consider the squared Euclidean distance

$$\delta^2(\mathbf{y}) = \min_{\boldsymbol{z} \in \mathcal{H}} \|\mathbf{y} - \boldsymbol{z}\|^2 \tag{6}$$

between the observed pixel $\mathbf{y}$ and the hyperplane $\mathcal{H}$ to decide which hypothesis (H_0 or H_1) is true. The next section studies the distribution of $\delta^2(\mathbf{y})$ under the two hypotheses H_0 and H_1.

3. DISTRIBUTIONS OF THE SQUARED DISTANCE UNDER HYPOTHESES H_0 AND H_1

We now design a statistical test for hypotheses H_0 and H_1 by studying the distribution of $\delta^2(\mathbf{y})$ under each hypothesis.

Under H_1 ($\boldsymbol{\mu} \neq 0$) or H_0 ($\boldsymbol{\mu} = 0$), given the sum-to-one constraint on the abundance vector, the mixing model (4) can be rewritten as

$$\tilde{\mathbf{y}} = \mathbf{y} - \mathbf{m}_R = \boldsymbol{K}\boldsymbol{c} + \boldsymbol{\mu} + \mathbf{e} \tag{7}$$

where $\boldsymbol{c} = [a_1, \ldots, a_{R-1}]^T$ is $(R-1) \times 1$ and $\boldsymbol{K} = [\mathbf{m}_1 - \mathbf{m}_R, \ldots, \mathbf{m}_{R-1} - \mathbf{m}_R]$ is $L \times (R-1)$. Hence, $\delta^2(\mathbf{y})$ can be computed by solving the unconstrained least squares (LS) problem

$$\delta^2(\mathbf{y}) = \min_{\boldsymbol{c} \in \mathbb{R}^{(R-1)}} \|\tilde{\mathbf{y}} - \boldsymbol{K}\boldsymbol{c}\|^2. \tag{8}$$

It is well known that the solution to this problem is given by

$$\hat{\boldsymbol{c}}_{\mathrm{LS}} = \left(\boldsymbol{K}^T\boldsymbol{K}\right)^{-1}\boldsymbol{K}^T\tilde{\mathbf{y}} \tag{9}$$

yielding the following $L \times 1$ residual vector $\hat{\mathbf{e}}$

$$\begin{aligned} \hat{\mathbf{e}} &= \tilde{\mathbf{y}} - \boldsymbol{K}\hat{\boldsymbol{c}}_{\mathrm{LS}} \\ &= \left[\mathbf{I}_L - \boldsymbol{K}\left(\boldsymbol{K}^T\boldsymbol{K}\right)^{-1}\boldsymbol{K}^T\right]\tilde{\mathbf{y}} \\ &= \boldsymbol{H}\left(\boldsymbol{\mu} + \mathbf{e}\right) \end{aligned} \tag{10}$$

where $\boldsymbol{H} = \mathbf{I}_L - \boldsymbol{K}\left(\boldsymbol{K}^T\boldsymbol{K}\right)^{-1}\boldsymbol{K}^T$ is an $L \times L$ projection matrix of rank $K = L - R + 1$. Using $\delta^2(\mathbf{y}) = \hat{\mathbf{e}}^T\hat{\mathbf{e}}$ and $\hat{\mathbf{e}} \sim \mathcal{N}\left(\boldsymbol{H}\boldsymbol{\mu}, \sigma^2\boldsymbol{H}\right)$, straightforward computations lead to [16]

$$\frac{1}{\sigma^2}\delta^2(\mathbf{y})|H_1 \sim \chi^2_K\left(\frac{\boldsymbol{\mu}^T\boldsymbol{H}\boldsymbol{\mu}}{\sigma^2}\right) \tag{11}$$

where $\chi^2_K(\lambda)$ denotes the noncentral χ^2 distribution with K degrees of freedom and noncentrality parameter λ. The distribution of $\delta^2(\mathbf{y})$ under H_0 can be obtained by setting $\boldsymbol{\mu} = \mathbf{0}$ in (11), yielding

$$\frac{1}{\sigma^2}\delta^2(\mathbf{y})|H_0 \sim \chi^2_K(0) = \chi^2_K \tag{12}$$

where χ^2_K is the χ^2 distribution with K degrees of freedom. Notice that the distributions of $\delta^2(\mathbf{y})$ under H_0 and H_1 depend on the known matrix $\mathbf{M}$ but also on the usually unknown noise variance σ^2 and nonlinearity vector $\boldsymbol{\mu}$. In the following we study nonlinearity detectors constructed for known and unknown noise variance σ^2.

4. NONLINEARITY DETECTION

As mentioned above, the distributions of $\delta^2(\mathbf{y})$ under H_0 and H_1 depend on σ^2 and on $\boldsymbol{\mu}$, which are usually unknown. We address first the scenario in which the noise variance is known but the nonlinearity under H_1 is unknown.

4.1. Known σ^2, unknown $\boldsymbol{\mu}$

For known σ^2, the distribution of $\delta^2(\mathbf{y})$ is perfectly known under H_0 and partially known under H_1. Thus, we use a statistical test that does not depend on the unknown nonlinearity $\boldsymbol{\mu}$ to decide between H_0 and H_1. Here we propose to use the following statistical test

$$T = \frac{\delta^2(\mathbf{y})}{\sigma^2} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{13}$$

where η is a threshold related to the probability of false alarm (PFA) of the test

$$P_{\mathrm{FA}} = \mathbb{P}\left[T > \eta \middle| H_0\right] \tag{14}$$

or equivalently,

$$\eta = \mathrm{F}^{-1}_{\chi^2_K}\left(1 - P_{\mathrm{FA}}\right) \tag{15}$$

where $\mathrm{F}^{-1}_{\chi^2_K}$ is the inverse cumulative distribution function of the χ^2_K-distribution. For a given $\boldsymbol{\mu}$, the power $P_{\mathrm{D}}(\boldsymbol{\mu})$ of the test is

$$P_{\mathrm{D}}(\boldsymbol{\mu}) = 1 - \mathrm{F}_{\chi^2_K(\lambda)}(\eta) \tag{16}$$

where $\mathrm{F}_{\chi^2_K(\lambda)}$ is the cumulative distribution function of the $\chi^2_K(\lambda)$-distribution and $\lambda = \sigma^{-2}\boldsymbol{\mu}^T\boldsymbol{H}\boldsymbol{\mu}$. Notice that the probability of detection (PD) $P_{\mathrm{D}}(\boldsymbol{\mu})$ is an increasing function of λ for a fixed threshold η. This makes sense, as the higher the power $\boldsymbol{\mu}^T\boldsymbol{H}\boldsymbol{\mu}$ of the nonlinearity orthogonal to $\mathcal{H}$, the better the detection performance. Moreover, the lower the noise variance, the better the nonlinearity detection. Unfortunately, the noise variance is unknown in most practical applications. In these cases, the test (13) cannot be used. We study a nonlinearity detector for unknown σ^2 in the next section.

4.2. Unknown σ^2, unknown $\boldsymbol{\mu}$

A simple solution if σ^2 is unknown is to replace the actual noise variance in (13) by its estimate $\hat{\sigma}^2$. This yields the following test

$$T^* = \frac{\delta^2(\mathbf{y})}{\hat{\sigma}^2} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{17}$$

where η is the threshold defined in (15). The PFA and PD of the test (17) are then given by

$$\begin{aligned} P_{\text{FA}}^* &= \mathbb{P}\left[T^* > \eta \middle| H_0\right] = \mathbb{P}\left[T > \frac{\hat{\sigma}^2}{\sigma^2}\eta \middle| H_0\right] \\ P_{\text{D}}^*(\boldsymbol{\mu}) &= \mathbb{P}\left[T^* > \eta \middle| H_1\right] = \mathbb{P}\left[T > \frac{\hat{\sigma}^2}{\sigma^2}\eta \middle| H_1\right]. \end{aligned} \tag{18}$$

The better the estimation of σ^2, the closer the distributions of T and T^* and thus the closer the performances of the tests (13) and (17). Here we propose to estimate σ^2 through an eigenanalysis of the sample covariance matrix of a set of N pixels assumed to share the same noise variance. The estimate $\hat{\sigma}^2$ is then determined as the average of the $p \leq L$ smallest eigenvalues of the sample covariance matrix. The accuracy of the estimator will depend on the choice of p, which will be discussed in the simulation section. For a given P_{FA} and associated threshold η given by (14), the distribution of T^* is shifted to the left if σ^2 is overestimated, i.e., $\hat{\sigma}^2 > \sigma^2$. This will lead to a P_{FA}^* of the test (17) that is lower than the P_{FA} of the test (13). Conversely, if $\hat{\sigma}^2 < \sigma^2$, the distribution of T^* is shifted to the right leading to $P_{\text{FA}}^* \geq P_{\text{FA}}$. Thus, it seems reasonable to overestimate σ^2 to ensure P_{FA}^* is upper bounded by a P_{FA} fixed by the user. This observation will be used in Section 5 to adjust the value of p.

5. SIMULATIONS

5.1. Synthetic data: known σ^2, unknown $\boldsymbol{\mu}$

We first investigate the performance of the test (13), which assumes σ^2 known. We consider a mixture of $R = 3$ materials (green grass, olive green paint and galvanized steel metal) whose spectral signatures $\mathbf{m}_r$ composed of $L = 826$ bands have been extracted from the spectral libraries provided with the ENVI software [17]. The abundance vector is fixed to $\boldsymbol{a} = [0.3, 0.6, 0.1]^T$ and the noise variance to $\sigma^2 = 10^{-3}$. The nonlinearity $\boldsymbol{\mu}$ is set as follows

$$\boldsymbol{\mu} = \nu \sum_{i=1}^{R-1} \sum_{j=i+1}^{R} a_i a_j \mathbf{m}_i \odot \mathbf{m}_j \tag{19}$$

where ν is a scaling factor selected from the set $\{0.4, 0.5, 0.6, 0.7\}$. This nonlinearity corresponds to the generalized bilinear model (GBM) studied in [13] with $\gamma_{i,j} = \nu$ for all (i, j). The specific form in (19) was chosen so that the impact of the nonlinearity is governed by a single parameter ν. Fig. 1 shows the theoretical and empirical ($N = 20000$ noise realizations) receiver operating characteristics (ROCs) [18, p. 74-75] for the test (13). Each value of ν corresponds to a different noncentrality parameter λ for the noncentral χ^2 distribution, ranging from $\lambda \approx 49$ to $\lambda \approx 150$. These results confirm that the performance of test (13) improves for larger values of λ (or ν).

5.2. Synthetic data: unknown σ^2, unknown $\boldsymbol{\mu}$

We now study the performance of the nonlinearity detector when σ^2 is replaced by its estimate $\hat{\sigma}^2$. Fig. 2 shows the ROCs of test (17) for $\lambda = 70$ and for three values of $\hat{\sigma}^2$: $0.95\sigma^2$, σ^2 and $1.05\sigma^2$. Note that

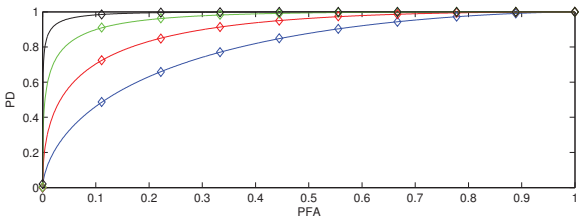

Fig. 1. Actual (solid lines) and empirical (diamonds) receiver operating characteristics (ROCs) of the first test (known noise variance) for $\nu = 0.4$ (blue), $\nu = 0.5$ (red), $\nu = 0.6$ (green) and $\nu = 0.7$ (black).

all ROCs coincide. However, different estimates $\hat{\sigma}^2$ correspond to different points on the curve for a fixed PFA. For instance, if the PFA is fixed to $P_{FA} = 0.1$ and the noise variance is correctly estimated, the corresponding P_D is around $P_D \approx 0.65$ (see Fig. 2 (middle)). If $\hat{\sigma}^2 = 0.95\sigma^2$, the PFA of test (17) rises to $P_{\text{FA}}^* \approx 0.41$, leading to $P_{\text{D}}^* \approx 0.92$. Conversely, if $\hat{\sigma}^2 = 1.05\sigma^2$, the PFA of test (17) falls to $P_{\text{FA}}^* \approx 0.01$, leading to $P_{\text{D}}^* \approx 0.27$.

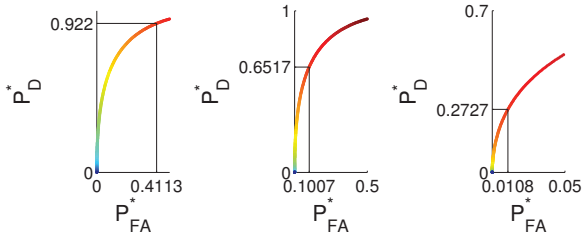

Fig. 2. Theoretical ROCs of the test (17) (unknown noise variance) for $\hat{\sigma}^2 = 0.95\sigma^2$ (left), $\hat{\sigma}^2 = \sigma^2$ (middle) and $\hat{\sigma}^2 = 1.05\sigma^2$ (right). The point corresponding with $P_{FA} = 0.1$ is the intersection of the black lines.

To investigate the ability of the proposed nonlinearity detector to detect different types of nonlinearities, we unmixed a synthetic image composed of $N = 10^4$ pixels generated according to four different mixing models. The $R = 3$ endmembers contained in this image are the same described in Section 5.1. We considered the following nonlinear mixing model

$$\mathbf{y}(n) = \mathbf{M}\boldsymbol{a}(n) + \cos(\theta)\boldsymbol{\mu}_1(n) + \sin(\theta)\boldsymbol{\mu}_2(n) + \mathbf{e}_n \tag{20}$$

for $n = 1, \ldots, N$, where $\mathbf{e}_n$ is a Gaussian noise vector such that $\mathbf{e}_n \sim \mathcal{N}\left(\mathbf{0}_L, \sigma^2\mathbf{I}_L\right)$ and

$$\boldsymbol{\mu}_1(n) = b\left[(\mathbf{M}\boldsymbol{a}(n)) \odot (\mathbf{M}\boldsymbol{a}(n))\right] \tag{21}$$

where $\odot$ denotes the Hadamard product and b is a fixed real parameter. The nonlinearity $\boldsymbol{\mu}_1$ corresponds to the nonlinear mixing model studied in [8] for nonlinearity detection. The $L \times 1$ nonlinearity vector $\boldsymbol{\mu}_2(n)$ is built to ensure that $\boldsymbol{\mu}_2(n)$ is orthogonal to the columns of $\mathbf{M}$ and to $\boldsymbol{\mu}_1(n)$ with $\|\boldsymbol{\mu}_1(n)\| = \|\boldsymbol{\mu}_2(n)\|$ ($\|\cdot\|$ denoting the ℓ_2 norm). The angle $\theta \in [0, \pi/2]$ is chosen to tune the contributions of $\boldsymbol{\mu}_1(n)$ and $\boldsymbol{\mu}_2(n)$ while ensuring $\|\cos(\theta)\boldsymbol{\mu}_1(n) + \sin(\theta)\boldsymbol{\mu}_2(n)\|^2 = \|\boldsymbol{\mu}_1(n)\|^2$ for any value of θ.

The considered set of pixels was divided into four 50×50 sub-images as follows. The first synthetic sub-image $\mathcal{S}_1$ was generated

using the standard linear mixing model (LMM) whereas the sub-images $\mathcal{S}_2$, $\mathcal{S}_3$ and $\mathcal{S}_4$ were generated according to the nonlinear mixing model (20) with $b = 0.1$ and $\theta \in \{\pi/4, 3\pi/8, \pi/2\}$. For each sub-image, the abundance vectors $\boldsymbol{a}_n, n = 1, \ldots, 2500$, were generated uniformly in the admissible set defined by the positivity and sum-to-one constraints. All sub-images were corrupted by an additive white Gaussian noise of variance $\sigma^2 = 10^{-3}$ corresponding to an average SNR ≈ 21dB. Table 1 shows the means and standard deviations of the noise variance estimates obtained for different values of p (for 50 Monte Carlo runs). This table shows that fixing $p = L - R + 1$ provides accurate estimates of σ^2 for these examples. Fig. 3 compares the empirical ROCs constructed from the number of pixels detected as linear and nonlinear for the different tests (known and unknown noise variance). The empirical ROCs for the test studied in [8] are also displayed in these figures. Fig. 3 (top left) shows that the three detectors are able to respect the PFA constraint. The three other subfigures of Fig. 3 display the ROCs for the three different values of θ. For small values of θ, the norm of the nonlinearity projection onto the vector $\boldsymbol{\mu}_1$ is large. Hence, the nonlinearity detector based on the PPNMM studied in [8] outperforms the tests studied in this paper (top right subfigure). However, the performance of test in [8] degrades as the portion of the nonlinearity that is orthogonal to $\boldsymbol{\mu}_1$ becomes predominant (bottom subfigures). The two proposed tests (known and unknown noise variance) perform similarly. Moreover, these two tests seem to be more robust to the type of nonlinearity. Finally, the proposed tests (13) and (17) only require one projection (10) of each pixel (and eventually the noise variance estimation procedure) while the test studied in [8] requires the minimization of a more complex cost function and the derivation of Cramèr-Rao bounds, leading to higher computation costs when compared to the proposed method.

	$\mathcal{S}_1$	$\mathcal{S}_2$	$\mathcal{S}_3$	$\mathcal{S}_4$
$L - p = 1$	19.94 ($\pm$0.27)	21.32 ($\pm$0.35)	20.69 ($\pm$0.30)	19.96 ($\pm$0.27)
$L - p = 2$	**9.99** ($\pm$0.01)	**9.98** ($\pm$0.01)	**9.99** ($\pm$0.01)	**9.99** ($\pm$0.01)
$L - p = 3$	9.97 ($\pm$0.01)	9.97 ($\pm$0.01)	9.98 ($\pm$0.01)	9.97 ($\pm$0.01)
$L - p = 4$	9.95 ($\pm$0.01)	9.95 ($\pm$0.01)	9.96 ($\pm$0.01)	9.96 ($\pm$0.01)
$L - p = 5$	9.94 ($\pm$0.01)	9.94 ($\pm$0.01)	9.94 ($\pm$0.01)	9.94 ($\pm$0.01)

Table 1. Means and standard deviations (in brackets) of the estimated noise variance ($\times 10^{-4}$) for different values of p ($\sigma^2 = 10^{-3}$).

5.3. Real data

The real image considered in this section is composed of spectral bands and was acquired in 1997 by the satellite AVIRIS over the Moffett Field, CA. A subimage of size 50×50 pixels (which we considered in previous work) was chosen to evaluate the proposed nonlinearity detector. The scene is mainly composed of water, vegetation, and soil. The corresponding endmembers were estimated by the vertex component analysis (VCA) [19] with $R = 3$. Fig. 4 shows an example of detection map (for $P_{FA} = 10^{-3}$) obtained with the proposed detector. It can be seen that nonlinear mixtures are mainly detected in the coastal area and in regions where mixtures of vegetation and soil occur (as also observed in [13]).

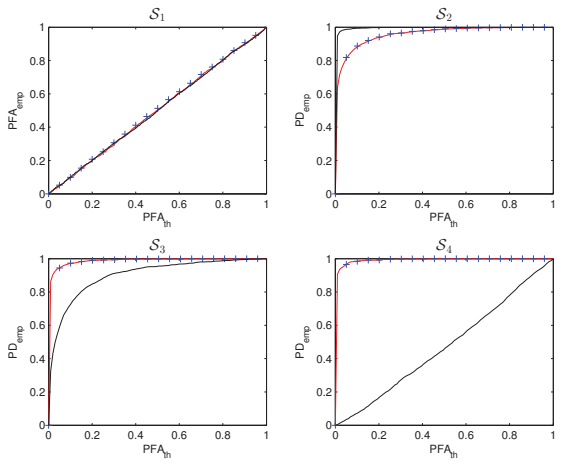

Fig. 3. Empirical ROCs of the tests (13) (red lines), (17) (blue crosses) and the test studied in [8] (black lines) for $\mathcal{S}_1$ to $\mathcal{S}_4$.

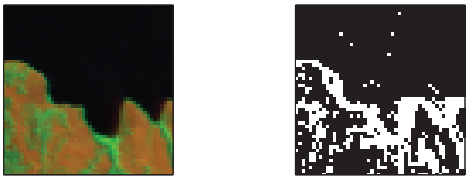

Fig. 4. Left: Moffett Field region of interest in synthetic colors. Right: Detection map obtained using the proposed detector for $P_{FA} = 10^{-3}$. Black (resp. white) pixels correspond to pixels detected as linearly (resp. nonlinearly) mixed.

6. CONCLUSIONS

This paper presented a new pixel by pixel nonlinearity detector for hyperspectral image analysis when the noise variance is known. The proposed detector is based on the distance between the observed pixel and the low dimensional subspace spanned by the endmembers. For unknown noise variance, a similar detector was proposed by replacing the actual noise variance by an accurate estimator based on eigenvalues resulting from the eigenanalysis of the sample covariance matrix of a set of image pixels. The main advantages of the proposed method are the absence of prior knowledge about the type of the nonlinearity and its low computational cost. Simulations on synthetic data illustrated the robustness of the method to detect various nonlinearities. The results obtained with a real AVIRIS hyperspectral image are also encouraging. Future works include the detection of nonlinearity in the case of unknown (or partially unknown) endmembers.

7. RELATION TO PRIOR WORK

As mentioned above, one of the main properties of the proposed nonlinearity detection method is that it does not rely on prior knowledge about the type of the nonlinearity used for spectral unmixing. Conversely, the pixel by pixel nonlinearity detector presented in [8] assumed a polynomial form for the nonlinearity with respect to the endmembers. We showed in this paper that the proposed detector has clear advantages when the nonlinear mixture of the endmembers cannot be approximated by a polynomial.

8. REFERENCES

[1] M. D. Craig, "Minimum volume transforms for remotely sensed data," *IEEE Trans. Geosci. and Remote Sensing*, pp. 542–552, 1994.

[2] D. C. Heinz and C.-I. Chang, "Fully constrained least-squares linear spectral mixture analysis method for material quantification in hyperspectral imagery," *IEEE Trans. Geosci. and Remote Sensing*, vol. 29, no. 3, pp. 529–545, March 2001.

[3] O. Eches, N. Dobigeon, and J.-Y. Tourneret, "Estimating the number of endmembers in hyperspectral images using the normal compositional model an a hierarchical bayesian algorithm," *IEEE J. Sel. Topics Signal Processing*, vol. 3, no. 3, pp. 582–591, June 2010.

[4] L. Miao and H. Qi, "Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization," *IEEE Trans. Geosci. and Remote Sensing*, vol. 45, no. 3, pp. 765777, March 2007.

[5] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observations Remote Sensing*, vol. 5, no. 2, pp. 354–379, April 2012.

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

[7] J. Chen, C. Richard, and P. Honeine, "Nonlinear unmixing of hyperspectral data based on a linear-mixture/nonlinear-fluctuation model," *IEEE Trans. Signal Process.*, vol. 61, no. 2, pp. 480–492, 2013.

[8] Y. Altmann, A. Halimi, N. Dobigeon, and J.-Y. Tourneret, "Supervised nonlinear spectral unmixing using a postnonlinear mixing model for hyperspectral imagery," *IEEE Trans. Image Processing*, vol. 21, no. 6, pp. 3017–3025, 2012.

[9] B. W. Hapke, "Bidirectional reflectance spectroscopy. I. Theory," *J. Geophys. Res.*, vol. 86, pp. 3039–3054, 1981.

[10] B. Somers, K. Cools, S. Delalieux, J. Stuckens, D. Van der Zande, W. W. Verstraeten, and P. Coppin, "Nonlinear hyperspectral mixture analysis for tree cover estimates in orchards," *Remote Sensing of Environment*, vol. 113, no. 6, pp. 1183–1193, 2009.

[11] J. M. P. Nascimento and J. M. Bioucas-Dias, "Nonlinear mixture model for hyperspectral unmixing," *Proc. of the SPIE*, vol. 7477, pp. 74770I–74770I–8, 2009.

[12] W. Fan, B. Hu, J. Miller, and M. Li, "Comparative study between a new nonlinear model and common linear model for analysing laboratory simulated-forest hyperspectral data," *Remote Sensing of Environment*, vol. 30, no. 11, pp. 2951–2962, June 2009.

[13] A. Halimi, Y. Altmann, N. Dobigeon, and J.-Y. Tourneret, "Nonlinear unmixing of hyperspectral images using a generalized bilinear model," *IEEE Trans. Geosci. and Remote Sensing*, vol. 49, no. 11, pp. 4153–4162, Nov. 2011.

[14] I. Meganem, P. Deliot, X. Briottet, Y. Deville, and S. Hosseini, "Physical modelling and non-linear unmixing method for urban hyperspectral images," in *Proc. IEEE GRSS Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, June 2011, pp. 1–4.

[15] P. Gader, D. Dranishnikov, A. Zare, and J. Chanussot, "A sparsity promoting bilinear unmixing model," in *Proc. IEEE GRSS Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, June 2012.

[16] A. Papoulis, *Probability, random variables, and stochastic processes*, Communications and signal processing. McGraw-Hill, 2 edition, 1991.

[17] RSI (Research Systems Inc.), *ENVI User's guide Version 4.0*, Boulder, CO 80301 USA, Sept. 2003.

[18] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection theory*, Prentice Hall, 1998.

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