DETECTING NONLINEAR MIXTURES IN HYPERSPECTRAL IMAGES

Yoann Altmann, Nicolas Dobigeon and Jean-Yves Tourneret

University of Toulouse, IRIT/INP-ENSEEIHT, France

{Yoann.Altmann, Nicolas.dobigeon, Jean-Yves.Tourneret}@enseeiht.fr

ABSTRACT

This paper presents a nonlinear mixing model for nonlinearity detection in hyperspectral images. The proposed model assumes that the pixel reflectances are nonlinear functions of pure spectral components contaminated by an additive white Gaussian noise. These nonlinear functions are approximated using polynomial functions leading to a polynomial post-nonlinear mixing model. The parameters involved in the resulting model are estimated using a least squares method. A generalized likelihood ratio test based on the asymptotic estimator distributions is then proposed to decide if the observed pixel results from the commonly used linear mixing model or from a more general nonlinear mixture. The derivation of a lower bound associated with the unmixing problem subject to the physical constraints of the abundance vectors allows the statistical test to be computed by assuming that the estimators achieve the considered bound. The performance of the detection strategy is evaluated thanks to simulations conducted on synthetic data.

***Index Terms*—** Hyperspectral images, nonlinearity detection, constrained estimation, post-nonlinear mixing model.

1. INTRODUCTION

Identification of the macroscopic materials present in a hyperspectral as well as their proportions in each pixel scene is of prime interest when analyzing hyperspectral images. Most of the related spectral unmixing (SU) strategies assume that the pixel reflectances are linear combinations of the endmembers [1–3]. However, as explained in [4], 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. For instance, the bidirectional reflectance-based model studied in [5] has been introduced for intimate mixtures. The bilinear models recently studied in [6–9] mainly focuses on scattering effects, e.g., observed in vegetation areas. Radial basis function networks [10, 11] and kernel-based models [12] have also been investigated for nonlinear SU. This paper considers a specific nonlinear model studied in [13] for nonlinear SU and referred to as *polynomial post-nonlinear mixing model* (PPNMM). The PPNMM belongs to the wide class of post-nonlinear mixing models introduced in [14] for source separation problems. It is a flexible generalization of the standard LMM that can accurately model many different nonlinearities [13]. The PPNMM has the nice property to be characterized by few parameters. In particular, its nonlinearity part is governed by a single real parameter referred to as nonlinearity parameter. The parameters of the PPNMM can be estimated using standard Bayesian or least squares (LS) methods (see [13] for details).

We assume that the endmembers contained in the image have been estimated by an endmember extraction algorithm (EEA), resulting in a supervised SU procedure. Consequently, the unknown PPNMM parameters are the abundances, the nonlinearity coefficient and the noise variance for each pixel of the image. This paper proposes to extract the endmembers contained in the hyperspectral image using a geometric EEA. Once the endmembers have been extracted from the image, the abundances and the nonlinearity parameter involved in the PNMM will be estimated using the subgradient method presented in [13].

This paper addresses the problem of determining whether an observed pixel of an hyperspectral image is a linear or nonlinear function of endmembers. One of the most interesting properties of the PPNMM is that it generalizes the LMM thanks to a unique nonlinearity parameter whose value characterizes the nonlinearity in the considered pixel. In particular, when the nonlinearity parameter equals zero, the resulting mixing model is linear. Consequently, it seems natural to use this parameter for deriving new nonlinearity detectors. It is precisely the objective of this paper which is organized as follows. Section 2 introduces the PPNMM for hyperspectral image unmixing. Section 3 derives the statistical test for nonlinearity detection based on the parameter estimates provided by the LS unmixing procedure. Some simulation results conducted on synthetic data are shown in Section 4.

2. POLYNOMIAL POST-NONLINEAR MIXING MODEL

This section introduces the nonlinear mixing model used for nonlinearity detection in hyperspectral images and the associated estimation algorithm.

2.1. PPNMM model

The L-spectrum $\mathbf{y} = [y_1, \ldots, y_L]^T$ of a mixed pixel is generally defined as a nonlinear transformation $\boldsymbol{g}(\cdot)$ of a linear mixture of R endmembers $\mathbf{m}_1, \ldots, \mathbf{m}_R$ contaminated by additive noise

$$\mathbf{y} = \boldsymbol{g}\left(\sum_{r=1}^{R} a_r \mathbf{m}_r\right) + \mathbf{n} = \boldsymbol{g}\left(\mathbf{M}\mathbf{a}\right) + \mathbf{n} \tag{1}$$

where $\mathbf{m}_r = [m_{r,1}, \ldots, m_{r,L}]^T$ is the spectrum of the rth material present in the scene, a_r is its corresponding proportion, R is the number of endmembers contained in the image and $\boldsymbol{g}(\cdot)$ is an appropriate nonlinear function. Moreover, L is the number of spectral bands and $\mathbf{n}$ is an additive independent and identically distributed (i.i.d) zero-mean Gaussian noise sequence with variance σ^2, denoted as $\mathbf{n} \sim \mathcal{N}\left(\mathbf{0}_L, \sigma^2 \mathbf{I}_L\right)$, where $\mathbf{I}_L$ is the $L \times L$ identity matrix. Note

Part of this work has been funded by a grant from DGA (French Ministry of Defence).

that the usual matrix and vector notations $\mathbf{M} = [\mathbf{m}_1, \ldots, \mathbf{m}_R]$ and $\mathbf{a} = [a_1, \ldots, a_R]^T$ have been used in the right hand side of (1).

Because of the lack of knowledge about the nonlinearity in (1), we proposed in [13] to approximate $\boldsymbol{g}(\cdot)$ using a second order polynomial $\boldsymbol{g}_b(\cdot)$ defined by

$$\mathbf{y} = \boldsymbol{g}_b(\mathbf{Ma}) + \mathbf{n} = \mathbf{Ma} + b(\mathbf{Ma}) \odot (\mathbf{Ma}) + \mathbf{n} \tag{2}$$

where $\odot$ denotes the Hadamard (term-by-term) product. Note that the resulting PPNMM includes bilinear terms such as those considered in [6–9]. However, the nonlinear terms are characterized by a single amplitude parameter b, which significantly simplifies the analysis.

Due to physical considerations, the abundance vector $\mathbf{a}$ satisfy the following positivity and sum-to-one constraints

$$\sum_{r=1}^{R} a_r = 1, \;\; a_r \geq 0, \forall r \in \{1, \ldots, R\}. \tag{3}$$

Moreover, it has been shown in [13] that the PPNMM is general enough to handle a wide class of nonlinear models.

2.2. Parameter estimation

The parameters of the PPNMM can be estimated by minimizing the following LS criterion

$$\begin{aligned} J(\mathbf{a}, b) &= \frac{1}{2} \|\mathbf{y} - \boldsymbol{g}_b(\mathbf{Ma})\|^2 \\ &= \frac{1}{2} \|\mathbf{y} - \mathbf{Ma} - b(\mathbf{Ma}) \odot (\mathbf{Ma})\|^2 \end{aligned} \tag{4}$$

subject to the constraints (3), where $\|\mathbf{x}\| = \sqrt{\mathbf{x}^T\mathbf{x}}$ is the standard ℓ^2 norm. After estimating $\mathbf{a}$ and b, the noise variance σ^2 can be determined as follows

$$\hat{\sigma}^2 = \frac{1}{L} \left\| \mathbf{y} - \mathbf{M}\hat{\mathbf{a}} - \hat{b}(\mathbf{M}\hat{\mathbf{a}}) \odot (\mathbf{M}\hat{\mathbf{a}}) \right\|^2. \tag{5}$$

Since the additive noise vector $\mathbf{n}$ is i.i.d zero-mean and Gaussian, the resulting estimator of $\boldsymbol{\theta} = [\mathbf{a}^T, b, \sigma^2]^T$ is the maximum likelihood estimator (MLE) of $\boldsymbol{\theta}$, denoted as $\hat{\boldsymbol{\theta}}$. Consequently, the estimator $\hat{\boldsymbol{\theta}} = [\hat{\mathbf{a}}^T, \hat{b}, \hat{\sigma}^2]^T$ is asymptotically efficient and asymptotically distributed according to a Gaussian distribution [15, Chap. 7] . Note that the asymptotic region corresponds to $L \to +\infty$. Since L is very large (some hundreds of spectral bands) for hyperspectral images, the asymptotic region will be achieved in most practical applications[1]. The two LS algorithms considered in [13] (i.e., linearization-based and subgradient-based algorithms) for minimizing (4) subject to the constraints (3) provide very similar performance. As a consequence, this paper will concentrate on one estimator only, namely the subgradient-based estimator. The next section derives a nonlinearity detector based on the MLEs of $\boldsymbol{\theta}$.

3. NONLINEARITY DETECTION

As shown in Section 2, the PPNMM allows the nonlinearity to be characterized by the parameter b for each pixel of the scene. An arbitrary threshold could be used to decide if the observed pixel is better modeled by the LMM or by a general nonlinear model defined by (2). However, it would be difficult to choose the appropriate threshold in order to guarantee a given probability of false alarm (PFA) or a given probability of detection (PD). In this section, we propose a statistical test for a pixel-by-pixel nonlinearity detection based on the distribution of $\hat{b}$. Based on the asymptotic properties of the MLE and on the large number of spectral bands available for a hyperspectral image, it makes sense to assume that $\hat{b}$ is distributed according to the following Gaussian distribution[2]

$$\hat{b} \sim \mathcal{N}\left(b, s^2\right) \tag{6}$$

where $s^2 \triangleq s^2(\mathbf{a}, b, \sigma^2)$ is the variance of the estimator $\hat{b}$. It is important to note that the variance of $\hat{b}$ is a function of the parameters $\mathbf{a}$, b and σ^2. Obviously, when the observation vector $\mathbf{y}$ results from the LMM (i.e., $b = 0$), then

$$\hat{b} \sim \mathcal{N}\left(0, s_0^2\right) \tag{7}$$

where $s_0^2 = s^2(\mathbf{a}, 0, \sigma^2)$. This interesting property can be used for testing the mixing model appropriate to the observation vector. The resulting nonlinearity detection problem can be considered as a two hypothesis testing problem, where the hypotheses are defined as

$$\begin{cases} H_0 & : \; \mathbf{y} \text{ is distributed according to the LMM} \\ H_1 & : \; \mathbf{y} \text{ is distributed according to the PPNMM} \end{cases} \tag{8}$$

Hypothesis H_0 is characterized by $b = 0$ whereas nonlinear models (H_1) lead to $b \neq 0$. As a consequence, the two hypotheses in (8) can be rewritten as

$$\begin{cases} H_0 & : \; \hat{b} \sim \mathcal{N}\left(0, s_0^2\right) \\ H_1 & : \; \hat{b} \sim \mathcal{N}\left(b, s_1^2\right) \end{cases} \tag{9}$$

where $s_1^2 = s^2(\mathbf{a}, b, \sigma^2)$ and $b \neq 0$.

3.1. Known parameters a and σ^2

For a given observation vector $\mathbf{y}$ and its corresponding estimated nonlinearity parameter $\hat{b}$, we propose to decide between hypotheses H_0 and H_1 using a classical generalized likelihood ratio test (GLRT). Using (6) and (7), the probability density functions of the test statistic $\hat{b}$ under the two hypothesis can be written

$$p(\hat{b}|H_0) = \left(\frac{1}{2\pi s_0^2}\right)^{\frac{1}{2}} \exp\left(-\frac{\hat{b}^2}{2s_0^2}\right) \tag{10}$$

$$p(\hat{b}|H_1) = \left(\frac{1}{2\pi s_1^2}\right)^{\frac{1}{2}} \exp\left(-\frac{(\hat{b}-b)^2}{2s_1^2}\right). \tag{11}$$

The corresponding GLRT consists of comparing the test statistic

$$\frac{\sup\limits_{b} \; p(\hat{b}|H_1)}{p(\hat{b}|H_0)} \tag{12}$$

to an appropriate threshold. Obviously, $p(\hat{b}|H_1)$ is maximized for $b = \hat{b}$. Straightforward computations lead to the following test strategy

$$T = \frac{\hat{b}^2}{s_0^2} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{13}$$

[1] The asymptotic behavior of the considered MLEs will be discussed in Section 4.

[2] This assumption will be validated in the simulation results.

where η is a threshold that is related to the test PFA as follows

$$P_{\mathrm{FA}} = \mathbb{P}\left[\frac{\hat{b}^2}{s_0^2} > \eta \middle| H_0\right] = 2\phi(-\sqrt{\eta}) \quad (14)$$

where $\phi(\cdot)$ is the cumulative distribution function of the normalized Gaussian distribution. For a given value of b, the power of the test $P_{\mathrm{D}}(b)$ can be computed as follows

$$\begin{aligned} P_{\mathrm{D}}(b) &= \mathbb{P}\left[\frac{\hat{b}^2}{s_0^2} > \eta \middle| H_1\right] = \mathbb{P}\left[\frac{\hat{b}^2}{s_0^2} > \eta \middle| b \neq 0\right] \\ &= 1 + \phi\left(\frac{-s_0\sqrt{\eta} - b}{s_1}\right) - \phi\left(\frac{s_0\sqrt{\eta} - b}{s_1}\right). \quad (15) \end{aligned}$$

It can be observed that for a given value of the threshold η, the probability of detection $P_{\mathrm{D}}(b)$ is an increasing function of $|b|$, which is an intuitive result. In order to apply the detection strategy (13) and to compute the corresponding P_{FA} and $P_{\mathrm{D}}(b)$, we need to know the parameters s_0 and s_1 whose determination is the objective of the next section.

3.2. Unknown parameters a and σ^2

The test (13) assumes known parameters $\mathbf{a}$ and σ^2 to compute $s_0^2 = s^2(\mathbf{a}, 0, \sigma^2)$. However, these parameters are unknown in most practical applications. To alleviate this problem, we propose to approximate the variance of $\hat{b}$ under H_0 by an appropriate estimator $\hat{s}_0^2$ leading to

$$\hat{T} = \frac{\hat{b}^2}{\hat{s}_0^2} \underset{H_0}{\overset{H_1}{\gtrless}} \eta^*. \quad (16)$$

More precisely, in order to build $\hat{s}_0^2$, we propose to compute the constrained CRLB of $\boldsymbol{\theta} = [\mathbf{a}^T, b, \sigma^2]^T$ under hypothesis H_0 (i.e., $b = 0$). The CCRLB of b is then given by the $(R+1)$th diagonal element of $\mathrm{CCRLB}(\boldsymbol{\theta})$ denoted as $\mathrm{CCRLB}(b; \mathbf{a}, \sigma^2)$. The estimator of the variance of $\hat{b}$ under hypothesis H_0 required to compute the test statistic in (16) is then defined as $\hat{s}_0^2 = \mathrm{CCRLB}(0; \hat{\mathbf{a}}, \hat{\sigma}^2)$ where $\hat{\mathbf{a}}$ and $\hat{\sigma}^2$ are the MLEs of $\mathbf{a}$ and σ^2. The next section study the performance of the nonlinearity detector defined by (16) for synthetic hyperspectral data.

4. SYNTHETIC DATA

The statistical test proposed in (16) assumes the efficiency and normality of the estimator $\hat{b}$ resulting from the unmixing procedure. Note that is has been shown that the asymptotic region in term of MLE efficiency is usually achieved in the hyperspectral imagery context (i.e., for large L and high signal to-noise ratio (SNR)) (see [16] for details).

The performance of the proposed nonlinearity detector is investigated by testing independently each pixel of a 100×100 synthetic image generated according to the PPNMM. The abundance vectors $\mathbf{a}_n, n = 1, \ldots, 10000$, have been randomly drawn from a uniform distribution over the simplex defined by the positivity and sum-to-one constraints. All pixels have been corrupted by an additive Gaussian noise of variance $\sigma^2 = 3 \times 10^{-3}$, corresponding to $\mathrm{SNR} \simeq 15$dB. The nonlinearity parameters have been chosen in the set $\{5\sigma^2, 10\sigma^2, 20\sigma^2, 30\sigma^2\}$, defining four different nonlinearity levels. Fig. 1 presents the actual nonlinearity parameters and the detection maps using the subgradient-based estimation procedure for $P_{\mathrm{FA}} = 0.01$ and $P_{\mathrm{FA}} = 0.05$. The white (resp. black) pixels are detected as nonlinearly (resp. linearly) distributed pixels. Note that similar results would be obtained when using the Taylor-based estimation procedure (see [16] for details).

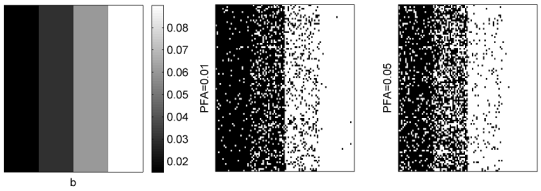

Fig. 1. Actual values of b (left) and detection maps for $\mathrm{P_{FA}} = 0.01$ (middle) and $\mathrm{P_{FA}} = 0.05$ (right) using the subgradient-based algorithm. Black (resp. white) pixels correspond to pixels detected as linearly (resp. nonlinearly) mixed.

The capacity of the PPNMM to detect various nonlinearities is investigated by unmixing a 100×100 synthetic image generated according to four different mixing models. The $R = 3$ endmembers contained in this image have been extracted from the spectral libraries provided with the ENVI software (i.e., green grass, olive green paint and galvanized steel metal). The considered image is divided into four 50×50 sub-images as follows. The first synthetic sub-image $\mathcal{S}_1$ has been generated using the standard linear mixing model (LMM). A second sub-image $\mathcal{S}_2$ has been generated according to the bilinear mixing model introduced in [8], referred to as "Fan model" (FM). A third sub-image $\mathcal{S}_3$ has been generated according to the generalized bilinear mixing model (GBM) recently introduced in [9], whereas a fourth sub-image $\mathcal{S}_4$ has been generated according to the proposed PPNMM. For each sub-image, the abundance vectors $\mathbf{a}_n, n = 1, \ldots, 2500$, have been randomly generated according to a uniform distribution over the admissible set defined by the positivity and sum-to-one constraints. All sub-images have been corrupted by an additive white Gaussian noise corresponding to $\mathrm{SNR} = 15$dB. The nonlinearity coefficients are uniformly drawn in the set $(0, 1)$ for the GBM and the parameter b has been generated uniformly in the set $(-0.3, 0.3)$ for the PPNMM. Fig. 2 shows the detection maps obtained with the GLRT for $P_{\mathrm{FA}} = 0.05$. From this figure, it can be seen that the location of the nonlinear mixtures on the detection maps is straightforward. Note that the GBM and the PPNMM generalize the LMM. As a consequence, mixed pixels can be close to the simplex corresponding to the noise-free case LMM and can be detected as linearly distributed pixels. For the FM, only almost pure pixels are close to that simplex, leading to a larger number of pixels detected as nonlinear. This remark is illustrated in Fig. 3 which shows the location of the pixels detected as nonlinear in the 3-dimensional space spanned by the three dominant axes resulting from a principal component analysis.

5. CONCLUSIONS AND FUTURE WORKS

A nonlinearity detector was presented for hyperspectral image analysis. This detector assumed that the hyperspectral image pixels are related to the endmembers by a polynomial post-nonlinear mixing model generalizing the widely used linear mixing model. A subgradient-based algorithm was used to estimate the model parameters. To determine the estimator variance a constrained Cramer-Rao lower bound was derived to handle the constraints for the abundance

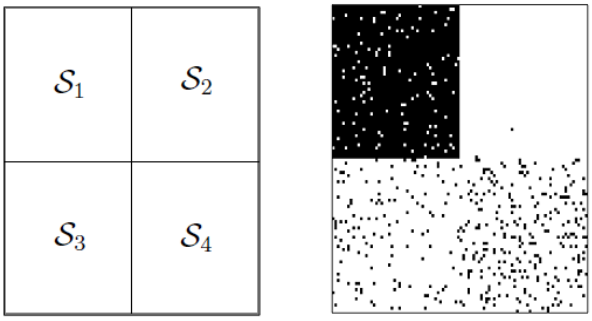

Fig. 2. Left: Actual location of the four sub-images S_1 (LMM), S_2 (FM), S_3 (GBM) and S_4 (PPNMM). Right: Associated detection map using the subgradient-based algorithm. Black (resp. white) pixels correspond to pixels detected as linearly (resp. nonlinearly) mixed.

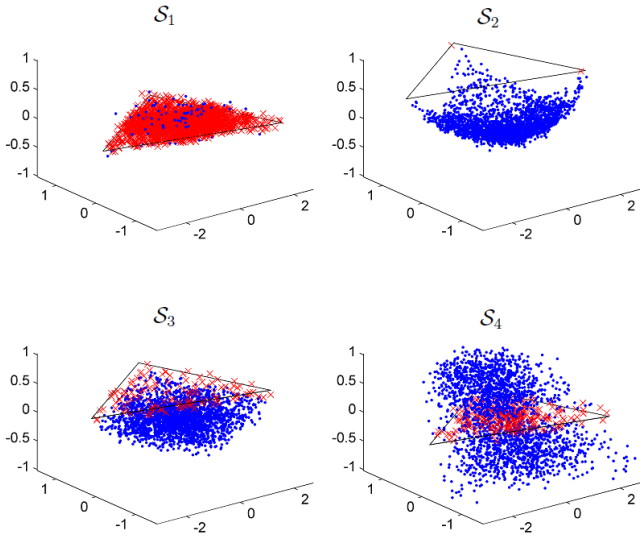

Fig. 3. Pixels detected as linear (red crosses) and nonlinear (blue dots) for the four sub-images S_1 (LMM), S_2 (FM), S_3 (GBM) and S_4 (PPNMM). The simplex corresponding to the noise-free case LMM is depicted in black lines.

vector. Based on the distribution of the resulting estimator, a generalized likelihood ratio test was derived to decide if a pixel of a hyperspectral image is a linear combination of endmembers or results from a general nonlinear mixture of the endmembers. Results obtained on synthetic and real images illustrated the accuracy of the polynomial post-nonlinear model for detecting nonlinearities in hyperspectral images. Future works include the consideration of spatial correlation between pixels of the hyperspectral image to improve unmixing and detection results.

6. REFERENCES

[1] M. Craig, "Minimum volume transforms for remotely sensed data," *IEEE Trans. Geosci. and Remote Sensing*, vol. 32, no. 3, 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, C. Mailhes, and J.-Y. Tourneret, "Bayesian estimation of linear mixtures using the normal compositional model," *IEEE Trans. Image Processing*, vol. 19, no. 6, pp. 1403–1413, june 2010.

[4] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Processing Magazine*, pp. 44–57, Jan. 2002.

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

[6] B. Somers, K. Cools, S. Delalieux, J. Stuckens, D. V. 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.

[7] J. M. P. Nascimento and J. M. Bioucas-Dias, "Nonlinear mixture model for hyperspectral unmixing," *Proceedings of the SPIE*, vol. 7477, pp. 74 770I–74 770I–8, 2009.

[8] 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.

[9] 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.

[10] K. J. Guilfoyle, M. L. Althouse, and C.-I. Chang, "A quantitative and comparative analysis of linear and nonlinear spectral mixture models using radial basis function neural networks," *IEEE Geosci. and Remote Sensing Lett.*, vol. 39, no. 8, pp. 2314–2318, Aug. 2001.

[11] Y. Altmann, N. Dobigeon, S. McLaughlin, and J.-Y. Tourneret, "Nonlinear unmixing of hyperspectral images using radial basis functions and orthogonal least squares," in *Proc. IEEE Int. Conf. Geosci. and Remote Sensing (IGARSS)*, july 2011, pp. 1151 –1154.

[12] J. Broadwater, R. Chellappa, A. Banerjee, and P. Burlina, "Kernel fully constrained least squares abundance estimates," in *Proc. IEEE Int. Conf. Geosci. and Remote Sensing (IGARSS)*, july 2007, pp. 4041 –4044.

[13] Y. Altmann, A. Halimi, N. Dobigeon, and J.-Y. Tourneret, "Supervised nonlinear spectral unmixing using a post-nonlinear mixing model for hyperspectral imagery," *IEEE Trans. Image Processing*, 2011, accepted under minor revisions.

[14] M. Babaie-Zadeh, C. Jutten, and K. Nayebi, "Separating convolutive post non-linear mixtures," in *Proc. of the 3rd Workshop on Independent Component Analysis and Signal Separation (ICA2001)*, San Diego, 2001, pp. 138–143.

[15] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation theory*. Englewood Cliffs NJ: Prentice Hall, 1993.

[16] Y. Altmann, A. Halimi, N. Dobigeon, and J.-Y. Tourneret, "Nonlinearity detection in hyperspectral images using a polynomial post-nonlinear mixing model," University of Toulouse, France, Tech. Rep., Nov. 2011. [Online]. Available: http://altmann.perso.enseeiht.fr/