



**HAL**  
open science

# Estimating abundance fractions of materials in hyperspectral images by fitting a post-nonlinear mixing model

Jie Chen, Cédric Richard, Paul Honeine

► **To cite this version:**

Jie Chen, Cédric Richard, Paul Honeine. Estimating abundance fractions of materials in hyperspectral images by fitting a post-nonlinear mixing model. Proc. IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 2013, Gainesville, Florida, USA, United States. 10.1109/WHISPERS.2013.8080639 . hal-01965999

**HAL Id: hal-01965999**

**<https://hal.science/hal-01965999>**

Submitted on 27 Dec 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# ESTIMATING ABUNDANCE FRACTIONS OF MATERIALS IN HYPERSPECTRAL IMAGES BY FITTING A POST-NONLINEAR MIXING MODEL

Jie Chen<sup>(1,2)</sup>, Cédric Richard<sup>(1)</sup>, Paul Honeine<sup>(2)</sup>

<sup>(1)</sup> Université de Nice Sophia-Antipolis, CNRS, Observatoire de la Côte d’Azur, France

<sup>(2)</sup> Université de Technologie de Troyes, CNRS, Institut Charles Delaunay, France  
{jie.chen, cedric.richard}@unice.fr, paul.honeine@utt.fr

## ABSTRACT

Within the area of hyperspectral data processing, nonlinear unmixing techniques have emerged as promising alternatives for overcoming the limitations of linear methods. In this paper, we consider the class of post-nonlinear mixing models of the partially linear form. More precisely, these composite models consist of a linear mixing part and a nonlinear fluctuation term defined in a reproducing kernel Hilbert space, both terms being parameterized by the endmember spectral signatures and their respective abundances. These models consider that the reproducing kernel may also depend advantageously on the fractional abundances. An iterative algorithm is then derived to jointly estimate the fractional abundances and to infer the nonlinear functional term.

**Index Terms**— Nonlinear unmixing, post-nonlinear mixing model, kernel methods, hyperspectral data processing

## 1. INTRODUCTION

There has been tremendous effort in the last decade to solve the linear unmixing problem, in the supervised, semi-supervised and unsupervised learning setting, motivated by its simplicity and acceptable approximation error in many scenarios. However, in the very recent years, researchers have started developing nonlinear unmixing techniques, as nonlinear models may provide more accurate estimations of mixtures, in particular when there are significant physical interactions of photons scattered by multiple materials [1].

Several algorithms have been proposed to address the nonlinear unmixing problem. For instance, bilinear models were considered to handle complex scenarios such as multilayered scenes [2, 3], by introducing additional interaction terms in the linear model. An unmixing algorithm based on a manifold learning process was investigated in [4], under the assumption that hyperspectral data may be embedded into a low-dimensional manifold. Kernel methods aim to avoid high computational complexity by using more simple physics-inspired models. Some kernels were designed to be sufficiently flexible for matching various nonlinearity de-

grees [5]. Finally, artificial neural networks were considered to conduct unmixing with training data [6].

More recently, it has been argued that both linear and nonlinear mixtures can be present in a single pixel [5, 7, 8], meaning that partially linear models may be advantageously used. In [7], we proposed a kernel-based nonlinear unmixing technique based on the partially linear assumption. In [8], the authors combined the linear mixture model with intimate mixtures. In [5], we investigated this problem under the framework of functional learning theory, where the tradeoff between the linear and the nonlinear components is automatically adjusted by solving a convex multi-kernel learning problem. However, in [5, 7], in order to make the problem convex and tractable, the nonlinear component of the model was assumed to be only a function of the material spectral signatures, and thus to be independent of their abundances.

In this paper, we consider a more general problem where the nonlinear mixing term also depends on the fractional abundances. This means that the reproducing kernel depends on the abundances, and is directly learned from data. A specific kernel is proposed, and an efficient iterative algorithm based on the augmented Lagrangian principle is derived to jointly estimate the abundances and to infer the nonlinear functional term.

## 2. NOTATIONS AND THE MIXTURE MODEL

### 2.1. Notations

Suppose that in a scene, there exist  $R$  significant endmembers with spectral signature  $\mathbf{m}_i \in \mathbb{R}^L$ , for  $i = 1, \dots, R$ , where  $L$  denotes the number of spectral bands. Let  $\mathbf{r} \in \mathbb{R}^L$  be an observed pixel, and  $\boldsymbol{\alpha} \in \mathbb{R}^R$  be the associated vector of abundances. Let us denote by  $\mathbf{M} = [\mathbf{m}_1, \dots, \mathbf{m}_R] \in \mathbb{R}^{L \times R}$  the matrix of the endmember spectra. For the sake of convenience, the  $\ell$ -th row of  $\mathbf{M}$  will be denoted by  $\mathbf{m}_{\lambda_\ell}^\top \in \mathbb{R}^L$ , that is, the column vector  $\mathbf{m}_{\lambda_\ell}$  contains the  $R$  endmember signatures at the  $\ell$ -th wavelength band.

## 2.2. Post-nonlinear mixture model

In order to estimate the mixing ratios of the endmembers, we modeled in [5] the nonlinear mixing mechanism  $\psi$  by a linear trend parameterized by the abundance vector  $\alpha$ , combined with a nonlinear fluctuation term represented by  $\psi_{\text{nlm}}$ , namely

$$\begin{aligned} \psi(\mathbf{m}_{\lambda_\ell}, \alpha) &= \psi_{\text{lin}}(\mathbf{m}_{\lambda_\ell}, \alpha) + \psi_{\text{nlm}}(\mathbf{m}_{\lambda_\ell}) \\ \text{with } \psi_{\text{lin}}(\mathbf{m}_{\lambda_\ell}, \alpha) &= \alpha^\top \mathbf{m}_{\lambda_\ell} \end{aligned}$$

where the nonlinear fluctuation  $\psi_{\text{nlm}}(\mathbf{m}_{\lambda_\ell})$  is independent of the abundance vector  $\alpha$ . It can be expected that incorporating the abundance into the nonlinear function should be advantageous for a more accurate modeling. In this paper, we consider the following post-nonlinear model for  $\psi_{\text{nlm}}$

$$\begin{aligned} \psi(\mathbf{m}_{\lambda_\ell}, \alpha) &= \psi_{\text{lin}}(\mathbf{m}_{\lambda_\ell}, \alpha) + \psi_{\text{nlm}}(\alpha^\top \mathbf{m}_{\lambda_\ell}) \\ \text{with } \psi_{\text{lin}}(\mathbf{m}_{\lambda_\ell}, \alpha) &= \alpha^\top \mathbf{m}_{\lambda_\ell}. \end{aligned} \quad (1)$$

This model can characterize many different nonlinearities. For example, with the parabolic transform  $\psi_{\text{nlm}}(x) = x^2$ , model (1) is a polynomial post-nonlinear mixture model that mimics second-order interactions between material signatures, including cross-terms of the bilinear mixture model discussed in [3].

## 3. KERNEL-BASED NONLINEAR UNMIXING WITH THE POST-NONLINEAR MODEL

### 3.1. Problem formulation

Considering the problem (1), we intend to estimate the abundance vector  $\alpha$  and to infer the nonlinear function  $\psi_{\text{nlm}}$  by solving the following functional optimization problem<sup>1</sup>

$$\begin{aligned} \psi^* &= \arg \min_{\psi} \frac{1}{2} (\|\psi_{\text{lin}}\|_{\mathcal{H}_{\text{lin}}}^2 + \|\psi_{\text{nlm}}\|_{\mathcal{H}_{\text{nlm}}}^2) + \frac{1}{2\mu} \sum_{\ell=1}^L e_\ell^2 \\ \text{where } \psi &= \psi_{\text{lin}} + \psi_{\text{nlm}} \quad \text{with } \psi_{\text{lin}}(\mathbf{m}_{\lambda_\ell}) = \alpha^\top \mathbf{m}_{\lambda_\ell} \\ \text{subject to } e_\ell &= r_\ell - (\alpha^\top \mathbf{m}_{\lambda_\ell} + \psi_{\text{nlm}}(\alpha^\top \mathbf{m}_{\lambda_\ell})) \\ \alpha &\succeq \mathbf{0} \quad \text{and} \quad \mathbf{1}^\top \alpha = 1. \end{aligned} \quad (2)$$

The functional  $\psi_{\text{nlm}}$  is restricted to be an element of the reproducing kernel Hilbert space  $\mathcal{H}_{\text{nlm}}$ , with the reproducing property

$$\psi_{\text{nlm}}(x) = \langle \psi_{\text{nlm}}, \kappa_{\text{nlm}}(\cdot, x) \rangle,$$

where  $\kappa_{\text{nlm}}$  is the reproducing kernel associated to the Hilbert space  $\mathcal{H}_{\text{nlm}}$ . It should be noticed that solving problem (2) is challenging because it is no longer convex, and the dual variables cannot be expressed in closed-form as in [5]. Fortunately, we can take advantage of its dual form for seeking a local optimum in an efficient way, via a variable splitting technique.

<sup>1</sup>The sum-to-one constraint, which may not be necessary in some contexts, can be removed without effecting the structure of the algorithm.

Let us introduce a new variable  $\mathbf{u}$  that replaces  $\alpha$  in the nonlinear fluctuation  $\psi_{\text{nlm}}$ . An equality constraint that connects  $\mathbf{u}$  and  $\alpha$  must also be introduced to ensure the equivalence between the transformed problem and the original one. The problem (2) is now formulated as

$$\begin{aligned} \alpha^* &= \underset{\psi, \alpha, \mathbf{u}, e}{\operatorname{argmin}} J(\psi, \alpha, \mathbf{u}, e) \\ &\text{subject to } \mathbf{u} = \alpha \end{aligned} \quad (3)$$

with

$$\begin{aligned} J(\psi, \alpha, \mathbf{u}, e) &= \frac{1}{2} (\|\psi_{\text{lin}}\|_{\mathcal{H}_{\text{lin}}}^2 + \|\psi_{\text{nlm}}\|_{\mathcal{H}_{\text{nlm}}}^2) + \frac{1}{2\mu} \sum_{\ell=1}^L e_\ell^2 \\ \text{where } \psi &= \psi_{\text{lin}} + \psi_{\text{nlm}} \quad \text{with } \psi_{\text{lin}}(\mathbf{m}_{\lambda_\ell}) = \alpha^\top \mathbf{m}_{\lambda_\ell} \\ \text{subject to } e_\ell &= r_\ell - (\alpha^\top \mathbf{m}_{\lambda_\ell} + \psi_{\text{nlm}}(\mathbf{u}^\top \mathbf{m}_{\lambda_\ell})) \\ \alpha &\succeq \mathbf{0} \quad \text{and} \quad \mathbf{1}^\top \alpha = 1. \end{aligned}$$

Thanks to this strategy, the derivatives with respect to each of the two variables  $\alpha$  and  $\mathbf{u}$  can now be evaluated separately. On the one hand, the variable  $\alpha$  is estimated using an algorithm similar to the one derived in [5]. On the other hand, the variable  $\mathbf{u}$  is updated by using a gradient descent approach.

### 3.2. Problem solving with an ADMM-like algorithm

The augmented Lagrange function of the problem (3) is

$$\begin{aligned} J_a^o(\psi, \alpha, \mathbf{u}, e, \nu) &= J(\psi, \alpha, \mathbf{u}, e) \\ &\quad + \nu^\top (\mathbf{u} - \alpha) + \frac{1}{2\zeta} \|\mathbf{u} - \alpha\|^2 \end{aligned}$$

where  $\zeta > 0$  is a positive parameter and  $\nu \in \mathbb{R}^R$  is the dual variable associated to the constraint  $\mathbf{u} = \alpha$ . For simplicity, the function  $J_a^o$  can also be written in an equivalent scaled form  $J_a$  with  $\mathbf{d}$  the scaled dual variable [9]

$$J_a(\psi, \alpha, \mathbf{u}, e, \mathbf{d}) = J(\psi, \alpha, \mathbf{u}, e) + \frac{1}{2\zeta} \|\mathbf{u} - \alpha - \mathbf{d}\|^2$$

Minimizing the above cost function can be performed by iterating the following three steps:

- **Step 1:** Update the variables  $\alpha$ ,  $\psi$  and  $e$  by

$$\alpha^{(k+1)}, \psi^{(k+1)}, e^{(k+1)} = \underset{\alpha, \psi, e}{\operatorname{argmin}} J_a(\psi, \alpha, \mathbf{u}^{(k)}, e, \mathbf{d}^{(k)})$$

- **Step 2:** Update the variable  $\mathbf{u}$  via a gradient descent

$$\mathbf{u}^{(k+1)} = \mathbf{u}^{(k)} - \eta \nabla_{\mathbf{u}} J_a(\psi^{(k+1)}, \alpha^{(k+1)}, \mathbf{u}, e^{(k+1)}, \mathbf{d}^{(k)})$$

with  $\eta$  the step size.

- **Step 3:** Update the scaled dual variable  $\mathbf{d}$  by

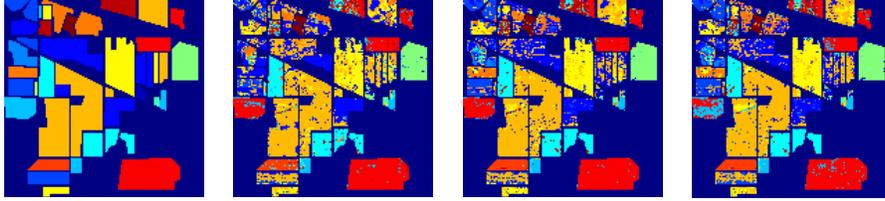
$$\mathbf{d}^{(k+1)} = \mathbf{d}^{(k)} + (\mathbf{u}^{(k+1)} - \alpha^{(k+1)}).$$



$$\max_{\beta, \gamma, \lambda} \bar{G}_a(\beta, \gamma, \lambda) = -\frac{\rho}{2} \begin{pmatrix} \frac{\beta}{\gamma} \\ \frac{\gamma}{\lambda} \end{pmatrix}^T \begin{pmatrix} \frac{1}{\rho}(\mathbf{K} + \mu\mathbf{I}) & \mathbf{M} & -\mathbf{M}\mathbf{1} \\ \mathbf{M}^T & \mathbf{I} & -\mathbf{1} \\ -\mathbf{1}^T \mathbf{M}^T & -\mathbf{1} & R \end{pmatrix} \begin{pmatrix} \beta \\ \gamma \\ \lambda \end{pmatrix} + \begin{pmatrix} \mathbf{r} - \frac{\rho}{\zeta} \mathbf{M}(\mathbf{u}^{(k)} - \mathbf{d}^{(k)}) \\ -\frac{\rho}{\zeta}(\mathbf{u}^{(k)} - \mathbf{d}^{(k)}) \\ \frac{\rho}{\zeta} \mathbf{1}^T (\mathbf{u}^{(k)} - \mathbf{d}^{(k)}) - 1 \end{pmatrix}^T \begin{pmatrix} \beta \\ \gamma \\ \lambda \end{pmatrix}$$

subject to  $\gamma \succeq \mathbf{0}$

with  $\mathbf{K} = \mathbf{K}_{\text{nlm}} + \rho \mathbf{M} \mathbf{M}^T$  and  $\rho = \frac{\zeta}{\zeta+1}$



**Fig. 2.** Classification map with 10% training data. From left to right: ground-truth, FCLS (61.36%), K-Hype (71.39%), Proposed (73.38%)

**Table 1.** RMSE comparison

	$R = 3$			$R = 5$		
	linear	bilinear	PNMM	linear	bilinear	PNMM
FCLS	<b>0.0037±0.00002</b>	0.0758±0.0019	0.0604±0.0017	<b>0.0134±0.0002</b>	0.1137±0.0032	0.1428±0.0039
ExtM	0.0079±0.0001	0.0312±0.0013	0.0601±0.0016	0.0157±0.0003	0.0575±0.0024	0.1427±0.0040
K-Hype	0.0208±0.0004	0.0349±0.0013	0.0446±0.0020	0.0231±0.0004	0.0307±0.0008	0.0398±0.0012
Proposed	0.0123±0.0002	<b>0.0288±0.0009</b>	<b>0.0159±0.0003</b>	0.0225±0.0005	<b>0.0245±0.0006</b>	<b>0.0299±0.0008</b>

**Table 2.** Classification accuracies with abundance features

Training percentage	5%	10%	15%
FCLS	56.41	61.36	62.32
K-Hype	67.67	71.39	74.68
Proposed	68.64	73.38	74.18

allowed us to evaluate the performance of unmixing algorithms via an alternative way that benefits from the ground truth reference of classification tasks. The scene used in our experiment is the celebrated Indian Pines region captured by AVIRIS. In this experiment, we used FCLS, K-Hype and the proposed algorithm for unmixing-based feature extraction. Training set was constructed by randomly selecting 5%, 10% and 15% of pixels. The comparative results are illustrated in Table 2 and Figure 2. These classification results clearly highlight the advantage of nonlinear unmixing algorithms, and in particular the proposed algorithm.

## 5. REFERENCES

- [1] 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 Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 2, pp. 354–379, 2012.
- [2] N. Raksuntorn and Q. Du, "Nonlinear spectral mixture analysis for hyperspectral imagery in an unknown environment," *IEEE Geoscience and Remote Sensing Letters*, vol. 7, no. 4, pp. 836–840, 2010.
- [3] A. Halimi, Y. Altman, N. Dobigeon, and J.-Y. Tourneret, "Nonlinear unmixing of hyperspectral images using a generalized bilinear model," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 11, pp. 4153–4162, 2011.
- [4] C. M. Bachmann, T. L. Ainsworth, and R. A. Fusina, "Exploiting manifold geometry in hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 3, pp. 441–454, 2005.
- [5] J. Chen, C. Richard, and P. Honeine, "Nonlinear unmixing of hyperspectral data based on a linear-mixture/nonlinear-fluctuation model," *IEEE Transactions on Signal Processing*, vol. 61, no. 2, pp. 480–492, 2013.
- [6] J. Plaza, P. Martínez, R. Pérez, and A. Plaza, "Nonlinear neural network mixture models for fractional abundance estimation in AVIRIS hyperspectral images," in *Proc. of XIII JPL Airborne Earth Science Workshop*, 2004.
- [7] J. Chen, C. Richard, and P. Honeine, "A novel kernel-based nonlinear unmixing scheme of hyperspectral images," in *Proc. of 45th Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, Pacific Grove (CA), USA, Nov. 2011.
- [8] R. Close and P. Gader, "Estimating the percentage of linear and nonlinear mixing at the subpixel level in hyperspectral imaging," in *Proc. IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Munich, Germany, Jul. 2012.
- [9] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.
- [10] D. C. Heinz and C.-I. Chang, "Fully constrained least squares linear mixture analysis for material quantification in hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 3, pp. 529–545, 2001.
- [11] F. Abdallah, C. Richard, and R. Lengelle, "An improved training algorithm for nonlinear kernel discriminants," *IEEE Transactions on Signal Processing*, vol. 52, no. 10, pp. 2798–2806, 2004.