Robust unmixing algorithms for hyperspectral imagery

ROBUST UNMIXING ALGORITHMS FOR HYPERSPECTRAL IMAGERY

Abderrahim Halimi[(1)], *Yoann Altmann*[(1)], *Gerald S. Buller*[(1)], *Steve McLaughlin*[(1)],
William Oxford[(2)], Damien Clarke[(2)], and Jonathan Piper[(2)]

[(1)] School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh U.K.
[(2)] Defence Science and Technology Laboratory, Porton Down, Salisbury, Wiltshire U.K.

ABSTRACT

The linear mixture model (LMM) assumes a hyperspectral pixel spectrum to be a linear combination of endmember spectra corrupted by additive noise. This model is widely used for spectral unmixing mainly because of its simplicity. However, the LMM can be inappropriate in presence of nonlinear effects, endmember variability or outliers. This paper presents a comparison between recent robust hyperspectral unmixing algorithms. The mixture models are first introduced followed by the description of their associated unmixing algorithms. The algorithms are then analyzed when considering a real image acquired over the region of Porton Down in England. The results discuss the behavior of each algorithm to unmix these data and compare their ability to detect the natural or man-made outliers in the scene. The obtained results highlight the potential of the studied mixture models to overcome the current limitations of the LMM.

***Index Terms*—** Hyperspectral imagery, robust unmixing, Bayesian estimation, optimization, MCMC, anomaly detection and analysis

I. INTRODUCTION

Spectral unmixing (SU) of hyperspectral images (HSI) has been the subject of intensive interest over the last two decades. It consists of distinguishing the materials (endmembers) and quantifying their proportions (abundances) in each pixel of an observed image. The linear mixture model (LMM) is the widely used model for SU mainly because of its simplicity. However, this model can be inappropriate for some hyperspectral scenarios, namely in the presence of a nonlinearity (NL) such as multiple scattering [1], [2], endmember variability (EV) [3]–[5], or outliers (rarely represented materials, faulty detector,...). These effects emphasize the need for robust HS unmixing strategies able to deal with these mismodelling effects [6]–[8].

This paper compares three recent mixture models accounting for the presence of residual components (RC) or outliers. The three models assume that the observed pixels result from a convex combination of the endmembers of the scene, corrupted by an additive term modeling deviations from the classical LMM. This residual term has different characteristics depending on the studied phenomenon. For instance, [6] considers a spatially-sparse positive RC suitable for the capture of nonlinearity effects, [7] proposes a spatial-spectral correlated support (i.e., the set of spatial/spectral locations where anomalies occur) for the RC and [8] presents a spatial-spectral smooth RC to better approximate endmember variability and/or nonlinearity. The properties of these models will be further discussed in Section II.

Estimating the abundances and the RC associated with these mixture models is an ill-posed problem that requires the introduction of prior knowledge regarding these parameters. From a Bayesian perspective, this goal can be achieved by assigning them appropriate prior distributions. From an optimization perspective, the prior knowledge can be seen as additional regularization terms. The two approaches (i.e., Bayesian and optimization formulations) will be further discussed in the paper when describing the three models. Indeed, the authors of [6] consider a coordinate descent algorithm (CDA) to minimize a regularized similarity measure. Two distance measures have been considered corresponding to the squared Euclidean distance (SED), and the Kullback-Leibler divergence (KLD). The regularization terms include the positivity of the RC and the spatial sparsity of the energies of the RC by considering an ℓ_{21} mixed norm. This norm is known as a collaborative regularization since it uses the information of the residuals in all the spectral bands to promote spatial group-sparsity. The authors of [7] adopt a Bayesian approach and define an anomaly prior model to capture the spatial/spectral structure of the potential RC. They then use a Markov chain Monte-Carlo (MCMC) method to approximate appropriate Bayesian estimators (i.e., minimum mean squares estimator (MMSE) and maximum a posteriori (MAP) estimator) [9]. The Bayesian model considers a Gaussian likelihood (related to the noise statistics) while assuming band-dependent noise variance. The model also enforces a spatial-spectral correlation on the RC support by considering a 3D Markov random fields model [5], [10]. In [8], the authors adopt a Bayesian approach while they approximate the MAP estimator by maximizing the resulting posterior distribution using a CDA. As in [7], [11], [12], the Bayesian model considers centered Gaussian noise with band-dependent variance. This model enforces a spectral smoothness on the RC by considering Gaussian process [13] and spatial correlation between the RC energies by using a gamma Markov random fields prior [14], [15]. The studied models and estimation algorithms are analyzed using a hyperspectral image acquired over Porton Down in England. The results obtained are very promising and show the potential of the studied mixture models to overcome the limitations of the LMM.

The paper is structured as follows. Section II introduces the considered mixture models to deal with the RC. The optimization formulation, the hierarchical Bayesian models and the estimation algorithms associated with each model are described in Section III. Section IV shows results obtained using a real hyperspectral image. Conclusions and future work are finally reported in Section V.

II. MIXTURE MODELS

The robust formulations considered are based on a residual component analysis model [16] that is expressed as the sum of a linear model and a residual term. The general observation model for the $(L \times 1)$ pixel spectrum $\boldsymbol{y}_n$ is given by

$$\begin{aligned} \boldsymbol{y}_n &= \sum_{r=1}^{R} a_{r,n}\boldsymbol{m}_r + \boldsymbol{\phi}_n + \boldsymbol{e}_n \\ &= \boldsymbol{M}\boldsymbol{a}_n + \boldsymbol{\phi}_n + \boldsymbol{e}_n, \end{aligned} \quad (1)$$

This work was supported by the EPSRC Grants EP/J015180/1, EP/N003446/1, and EP/K015338/1

where $\boldsymbol{a}_n = (a_{1,n}, \cdots, a_{R,n})^T$ is the $(R \times 1)$ vector of abundances associated with the nth pixel, R is the number of endmembers, $\boldsymbol{e}_n$ is an additive noise, L is the number of spectral bands, $\boldsymbol{M} = (\boldsymbol{m}_1, \cdots, \boldsymbol{m}_R)$ is the endmember matrix that is assumed known (either picked out from a spectral library if available, or extracted using an endmember extraction algorithm) and $\boldsymbol{\phi}_n$ is a residual term that might have different characteristics depending on the studied effect. Due to physical constraints, the abundance vector $\boldsymbol{a}_n$ satisfies the following positivity and sum-to-one (PSTO) constraints

$$a_{r,n} \geq 0, \forall r \in \{1, \ldots, R\} \quad \text{and} \quad \sum_{r=1}^{R} a_{r,n} = 1. \tag{2}$$

Eq. (1) shows a general model that can be adapted to account for different robust mixture models. In this paper, we consider three variants: (i) the robust nonnegative matrix factorization (RNMF) proposed in [6], (ii) the robust Bayesian linear unmixing (RBLU) proposed in [7], and (iii) the residual component analysis with mismodelling effects (RCA-ME) proposed in [8]. Table I presents the main outlier and noise prior statistical properties assumed by each method and the next sections describe in more details their properties.

Table I. Characteristics of the robust models. "Spar." stands for sparsity, "pos." for positivity, "val." for values and "correl." for correlation.

	Residuals			Statistics
	Pos.	Spatial	Spectral	
RNMF	✓	Spar. energies	None	-SED -KLD
RBLU	X	Correl. and spar. support	Correl. support	$\mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$
ME	X	Correl. energies	Correl. val.	$\mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$

II-A. RNMF

The RNMF model is given by [6]

$$\boldsymbol{y}_n^{\text{RNMF}} = \sum_{r=1}^{R} a_{r,n} \boldsymbol{m}_r + \boldsymbol{r}_n \tag{3}$$

where $\boldsymbol{r}_n, \forall n$ denote positive residuals that are due to the presence of nonlinear effects such as the bilinear models of [17]–[19] and the polynomial model of [20]. Therefore, this positivity assumption is physically well-motivated for multi-layered models (e.g., scenes where significant multiple reflections can occur). In addition, the RNMF model assumes that the nonlinear spectra are spatially sparse, i.e., the majority of the pixels follows the LMM while only some pixels include an additional residual term. Note, however, that the RNMF does not assume any spectral structure for the residuals. The RNMF estimates the parameter of interest using two measures of similarity between the observed spectrum $\boldsymbol{y}_{\ell,n}$ and the spectrum $\boldsymbol{y}_{\ell,n}^{\text{RNMF}}$. The first one is the SED, given by $\sum_{\ell,n} \left(\boldsymbol{y}_{\ell,n} - \boldsymbol{y}_{\ell,n}^{\text{RNMF}}\right)^2$, and the second one is the KLD, given by $\sum_{\ell,n} \left(\boldsymbol{y}_{\ell,n}^{\text{RNMF}} \log \frac{\boldsymbol{y}_{\ell,n}^{\text{RNMF}}}{\boldsymbol{y}_{\ell,n}} \right. \left. - \boldsymbol{y}_{\ell,n}^{\text{RNMF}} + \boldsymbol{y}_{\ell,n} \right)$. Finally, RNMF is an unsupervised algorithm that estimates both the endmembers, the abundances and the residuals. However, for comparison purposes, this paper considers the supervised case (known endmember matrix $\boldsymbol{M}$) and only estimates the remaining parameters. This can be easily implemented by skipping the endmember update step of the iterative estimation process (see [6] for details on the algorithm implementation).

II-B. RBLU

The RBLU model follows (1) while considering additive centered Gaussian noise with a diagonal covariance matrix $\boldsymbol{\Sigma} = \text{diag}\left(\boldsymbol{\sigma}^2\right)$ as follows $\boldsymbol{e}_n \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$, where $\boldsymbol{\sigma}^2 = \left(\sigma_1^2, \cdots, \sigma_L^2\right)^T$ is an $(L \times 1)$ vector containing the noise variances. In addition, the outliers $\boldsymbol{\phi}_n = \boldsymbol{z}_n \odot \boldsymbol{x}_n, \forall n$ are represented by a term-wise product $\odot$ between the support variable $\boldsymbol{z}_n$ (which is an $L \times 1$ vector of 0s and 1s) and the vector of values $\boldsymbol{x}_n$. Note that the outliers are assumed to share a correlated support $\boldsymbol{Z} = [\boldsymbol{z}_1, \cdots, \boldsymbol{z}_N]$, i.e., the spatial and spectral neighbors of a corrupted pixel (with outliers in a given spectral band) are more likely to contain corruption. The outlier structure is modeled via a 3D Ising model and generalizes the model studied in [21] which assumed the support sets of outliers to have a fixed structure. However, the values of these corruptions contained in $\boldsymbol{X} = [\boldsymbol{x}_1, \cdots, \boldsymbol{x}_N]$ are independent which enables flexibility in this model to capture different physical phenomena such as local spectral and/or spatial variations (e.g., intrinsic material spectral variability, illumination changes, scarcely represented endmembers,...). Note also that RBLU enforces spectral and spatial sparsity among the corrupted pixels. Finally, and as previously described, this paper considers the supervised case and RBLU will be used to estimate the parameters of interest while considering a known endmember matrix $\boldsymbol{M}$.

II-C. RCA-ME

In a similar fashion to RBLU, the RCA-ME model considers additive centered Gaussian noise with a diagonal covariance matrix $\boldsymbol{\Sigma} = \text{diag}\left(\boldsymbol{\sigma}^2\right)$. To represent different physical phenomena such as nonlinearity and/or endmember variability, the outliers $\boldsymbol{\phi}_n = \boldsymbol{d}_n, \forall n$ are assumed to present some spectral smoothness. Indeed, for bilinear and polynomial based nonlinearity models, the outliers result from a (nonlinear) combination of endmembers that are generally smooth (when the endmembers themselves are smooth). In addition, it has been shown in [8] that endmember variability can also be approximated by assuming spectrally smooth residuals. Note that RCA-ME also assumes the energies of the outliers to be spatially correlated. Note finally that RCA-ME accounts for spatial illumination variability by relaxing the abundance sum-to-one constraint, however, for a comparison purpose, this paper assumes the absence of this variability and maintains the abundance sum-to-one constraint.

III. UNMIXING ALGORITHMS

This section presents the optimization formulation, the hierarchical Bayesian models and the estimation algorithms associated with each model.

III-A. Optimization problem for RNMF

Considering a known endmember matrix $\boldsymbol{M}$, the RNMF estimates the unknown parameters by solving the following problem

$$\min_{\boldsymbol{A},\boldsymbol{R}} C_{\text{RNMF}}(\boldsymbol{A}, \boldsymbol{R}) = D(\boldsymbol{Y}|\boldsymbol{M}\boldsymbol{A} + \boldsymbol{R}) + \lambda ||\boldsymbol{R}||_{2,1} \tag{4}$$

s.t. the positivity of $\boldsymbol{A}$ and $\boldsymbol{R}$ and to the sum-to-one constraint on $\boldsymbol{A}$, where $\boldsymbol{A} = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_N]$, $\boldsymbol{R} = [\boldsymbol{r}_1, \cdots, \boldsymbol{r}_N]$, $D(x|y)$ is the distance measure between x and y and $||\boldsymbol{R}||_{2,1} = \sum_{n=1}^{N} \sqrt{\boldsymbol{r}_n^\top \boldsymbol{r}_n}$ is a mixed norm that imposes spatial sparsity on the energies of $\boldsymbol{R}$. This problem is solved using a coordinate descent algorithm (CDA) that sequentially updates the two matrices $\boldsymbol{R}$ and $\boldsymbol{A}$. More precisely, while fixing $\boldsymbol{A}$, RNMF updates $\boldsymbol{R}$ using a majorization-minimization algorithm. Then, fixing $\boldsymbol{R}$, RNMF updates $\boldsymbol{A}$ using a non-negative gradient descent algorithm. The resulting iterative algorithm was found experimentally to decrease the value of the objective function at each iteration [6].

III-B. Bayesian model for RBLU

Using Bayes rule, the noise statistics (likelihood) and the prior knowledge about the parameter of interest $\boldsymbol{\Theta}_1 = \left(\boldsymbol{A}, \boldsymbol{\Sigma}, \boldsymbol{X}, \boldsymbol{Z}, s^2, \boldsymbol{\beta}\right)$ lead to the following posterior distribution

$$f\left(\boldsymbol{\Theta}_1|\boldsymbol{Y}\right) \propto f(\boldsymbol{Y}|\boldsymbol{\Theta}_1) f\left(\boldsymbol{A}\right) f\left(\boldsymbol{\Sigma}\right) f\left(\boldsymbol{X}|s^2\right) f\left(\boldsymbol{Z}|\boldsymbol{\beta}\right) f\left(s^2\right) f\left(\boldsymbol{\beta}\right), \tag{5}$$

where s^2 and $\boldsymbol{\beta}$ are two hyperparameters related to the value of the outliers, and to the correlation of the outlier supports, respectively. RBLU uses a stochastic optimization mechanism to adjust the parameters $\boldsymbol{\beta}$ of the Markov random field associated with the support $\boldsymbol{Z}$. The remaining parameters are sampled according to the target posterior distribution (5) using an MCMC algorithm. The generated samples are then used to approximate the minimum mean square estimator (MMSE) or the maximum a-posteriori (MAP) estimator of the parameters. The obtained algorithm benefits from the good convergence properties of the MCMC approach [9] while it suffers from a high computational cost.

III-C. Bayesian model for RCA-ME

Similarly to RBLU, the CDA-ME adopts a Bayesian approach leading to the following posterior

$$f\left(\boldsymbol{\Theta}_2|\boldsymbol{Y}\right) \propto f(\boldsymbol{Y}|\boldsymbol{\Theta}_2) f\left(\boldsymbol{A}\right) f\left(\boldsymbol{\Sigma}\right) f\left(\boldsymbol{D}|\boldsymbol{\epsilon}\right) f\left(\boldsymbol{\epsilon}, \boldsymbol{w}\right), \tag{6}$$

where $\boldsymbol{\Theta}_2 = (\boldsymbol{A}, \boldsymbol{\Sigma}, \boldsymbol{D}, \boldsymbol{\epsilon}, \boldsymbol{w})$, $\boldsymbol{\epsilon}$ and $\boldsymbol{w}$ are two hyperparameters introducing the spatial correlation of the outlier energies (see [8] for more details). CDA-ME approximates the MAP estimator of $\boldsymbol{\Theta}$ by sequentially updating the different parameters. In each step, the posterior distribution is maximized w.r.t. one parameter, the other being fixed. Thus, the algorithm iteratively updates each parameter by maximizing its conditional distribution either by an analytical solution (as for $\boldsymbol{\epsilon}$, $\boldsymbol{w}$, $\boldsymbol{D}$, $\boldsymbol{\Sigma}$) or by using the fast SUNSAL algorithm [22] (as for $\boldsymbol{A}$). According to proposition 2.7.1 in [23], the limit points of the sequence generated by CDA are stationary points of (6), since the maximum of that function w.r.t. $\boldsymbol{\Theta}$ along each coordinate is unique.

IV. RESULTS

This section evaluates the performance of the robust algorithms described above when applied to a real hyperspectral data set. All simulations have been implemented using MATLAB R2015a on a computer with Intel(R) Core(TM) i7-4790 CPU@3.60GHz and 32GB RAM.

IV-A. Data description

The real image used in this section was acquired in July 2014 by Defence Science and Technology Laboratory (DSTL) over Porton Down (U.K.). The dataset contains $L = 140$ spectral bands recorded from the visible to near infrared (415 to 990 nm) with a spatial resolution of about 0.2 metres. The unmixing algorithms have been applied to a scene of size 400×200 pixels which is composed of $R = 5$ components: grass, tree, soil 1, road, and soil 2, and other man-made outliers as shown in Fig. 1 (see green and white spots). This image is interesting since it potentially includes nonlinearity effect between the tree and soil, endmember variability for the grass area and man-made sparse outliers, which makes it suitable for the assessment of the algorithms presented in this paper. The endmember associated with the $R = 5$ components have been manually selected by choosing the purest pixels in each component's region.

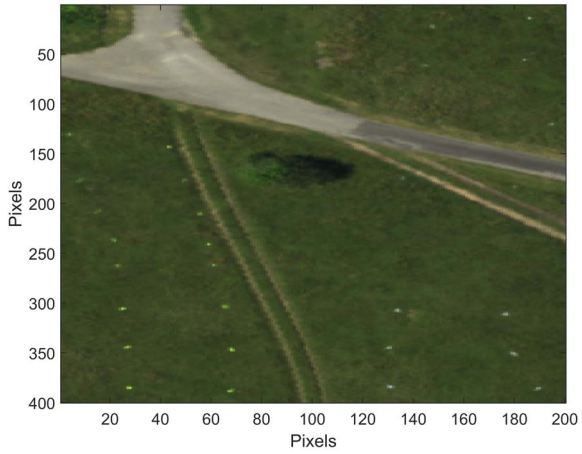

Fig. 1. An example of the hyperspectral image data used in the analysis

IV-B. Comparison between the algorithms

The abundances of the real image have been estimated by the robust algorithms and the LMM based SUNSAL algorithm [22]. After using SED and KLD similarity measures for the RNMF and $\lambda \in [0.01\lambda_0, \lambda_0, 2\lambda_0]$, where λ_0 is the suggested value in [6], we found that the best reconstruction error, given by $\text{RE} = \sqrt{\frac{1}{NL}\sum_{n=1}^{N} \|\hat{\boldsymbol{y}}_n - \boldsymbol{y}_n\|^2}$), is obtained when considering the KLD measure and $\lambda = 0.01\lambda_0$. The following RNMF results are related to this parameter configuration. Fig. 2 shows the obtained abundance maps where a white (black) pixel indicates a large (small) proportion of the corresponding materials. Overall, the obtained abundances are in good agreement except RMNF that show slightly different tree and grass abundances (which can be explained by the high similarity between the tree and grass spectra). Note that some of the man-made outliers are detected as a soil abundances (see far right-hand side plots of Fig. 2).

Table II shows the obtained RE and spectral angle mapper, $\text{SAM} = \frac{1}{N}\sum_{n=1}^{N} \arccos\left(\frac{\hat{\boldsymbol{y}}_n^T \boldsymbol{y}_n}{\|\boldsymbol{y}_n\| \, \|\hat{\boldsymbol{y}}_n\|}\right)$, between the observed and reconstructed spectra with the considered algorithms. This table shows that the robust algorithms provide smaller RE and SAM than the LMM based SUNSAL algorithm. However, this is achieved at a price of a higher computational cost especially for the MCMC-based RBLU algorithm. An interesting property of the robust algorithms is their ability to show outlier energy maps as presented in Fig. 3 (left). The nonlinearity effect present between the tree-soil is detected by all the algorithms. In contrast to RBLU and CDA-ME, RNMF does not detect the shadow as an outlier since it generally reduces the reflectance and can be seen as a negative outlier, which is ignored in RNMF because of the outlier positivity constraint. The three algorithms succeeded in capturing the endmember variability effect affecting the grass with CDA-ME detecting the larger region. It is also of interest to note that the algorithms detect the man-made sparse outliers. Considering the RE maps in Fig. 3 (right), it is clear that the RE of RNMF is higher in shadowed regions while RBLU and CDA-ME present low RE.

Fig. 4 shows some randomly selected outlier spectra obtained with the three robust algorithms. These spectra show a similar global shape while they highlight the properties of each algorithm. Indeed, it can be seen that the RNMF algorithm focuses on the nonlinearity effect (i.e., positive outliers) while it avoids negative spectra that are mainly related to shadowing and endmember

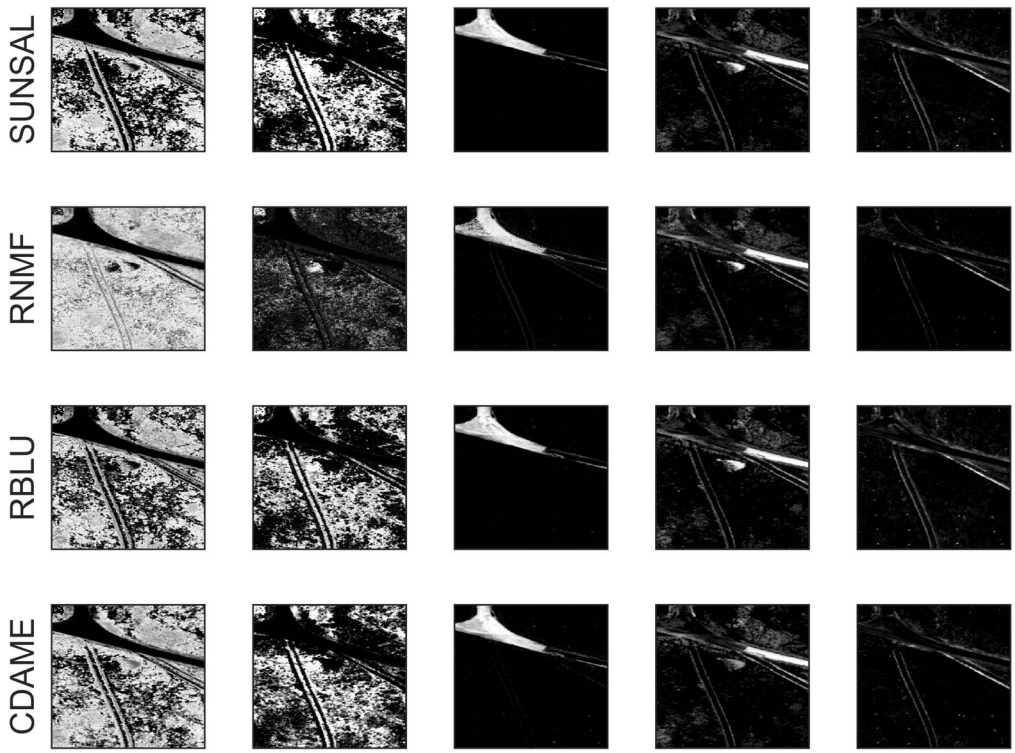

Fig. 2. Estimated abundance maps with different algorithms (all images have the same dynamic, i.e., between 0 and 1). From left to right: grass, tree, soil 1, road, and soil 2.

variability effects. Moreover, it is interesting to note that RBLU allows the absence of outliers (null spectra) which is an advantage of imposing sparsity on the support of the outliers instead of their values. Note finally that Fig. 4 also highlights the smooth spectral property of CDA-ME.

Finally, Table III summarizes the pros and cons of each model. A few comments should be made regarding the noise robustness and the endmember columns. Both RBLU and CDA-ME consider Gaussian noise with band dependent variances, thus, they are robust to the spectral variation of the noise. SUNSAL and RNMF (with SED) consider independent and identically distributed Gaussian noise which is less robust to the spectral variation of the noise. However, RNMF considers a general formulation that allows the use of the KLD measure, which makes it suitable to process images corrupted by non Gaussian noise. Note finally that both RNMF and RBLU can estimate the endmembers, while SUNSAL and CDA-ME assume a known endmember matrix $\boldsymbol{M}$. This property has been reported in table III for completeness, but the reader should be aware that the endmembers have been assumed known in this paper.

Table II. Unmixing performance on a real images (400×200 pixels).

	RE ($\times 10^{-3}$)	SAM ($\times 10^{-2}$)	Time (min)
SUNSAL	9.32	3.34	0.03
RNMF	7.17	2.05	23.92
RBLU	4.28	2.32	1440
CDA-ME	3.87	2.21	6.98

Table III. Comparison of the robust algorithms. (+++) best results, (-) fair results.

	Effects	Time	RE	Noise	Endm.
SUNSAL	LMM	+++	-	-	-
RNMF	LMM + NL	+	+	++	+
RBLU	LMM + NL EV + Shadow	-	++	++	+
CDAME	LMM + NL EV + Shadow	++	+++	++	-

V. CONCLUSIONS

This paper presented a comparison between three robust mixture models and their associated unmixing algorithms. The three mixture models are expressed as the sum of a linear model and a residual term. The latter depended on the model considered and showed different characteristics as reported in Table I. In contrast to RBLU and CDA-ME that considered Gaussian statistics, RNMF adopted a general formulation that allowed the use of two similarity measures.

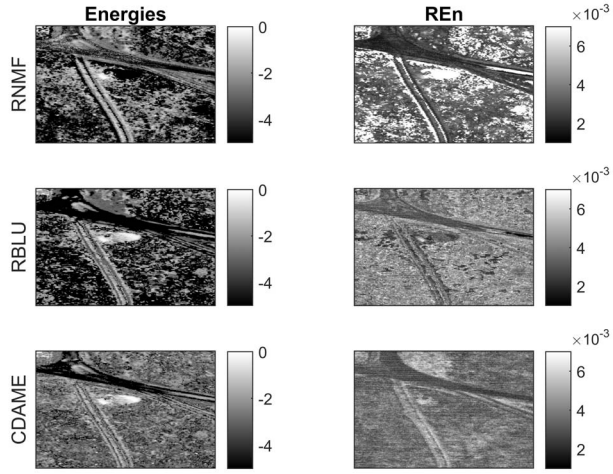

Fig. 3. (Left) square root of the energies of the residuals obtained with $||\hat{\boldsymbol{\phi}}_n||$. (Right) reconstruction error between the reconstructed pixel $\hat{\boldsymbol{y}}_n$ and the observed spectra $\boldsymbol{y}_n$ obtained with $\text{RE}_n = \frac{1}{\sqrt{L}}||\hat{\boldsymbol{y}}_n - \boldsymbol{y}_n||$.

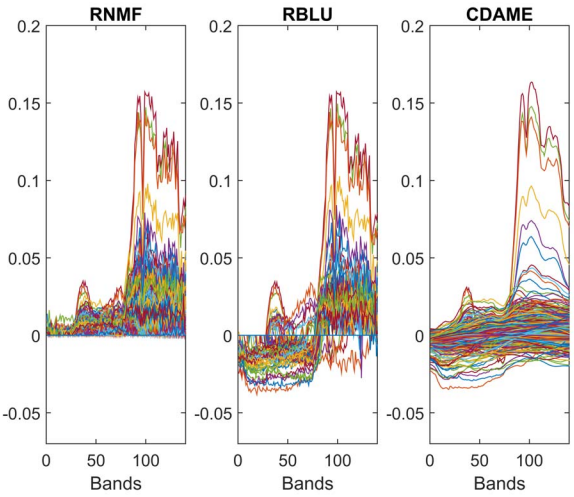

Fig. 4. Example of outlier spectra for (left) RNMF, (middle) RBLU and (right) CDAME.

Both RNMF and CDA-ME used a CDA to minimize their cost functions, while RBLU considered a Markov chain Monte-Carlo to approximate the Bayesian estimators. A comparison between these strategies was achieved by analyzing of a real image containing different effects (such as nonlinearity, endmember variability, outliers and shadow). The results obtained highlighted the benefit of the robust models with respect to the classical linear model. The pros and cons of each model were summarized in Table III, to provide the reader with a useful tool to choose the suitable model/algorithm for his case. Future work includes the use of these models to detect the presence of outliers using hypothesis tests [24]. Considering an efficient optimization algorithm to maximize the RBLU posterior distribution is also an interesting issue which deserves to be investigated.

VI. REFERENCES

[1] R. Heylen, M. Parente, and P. Gader, "A review of nonlinear hyperspectral unmixing methods," *IEEE J. Sel. Topics Appl. Earth Observat. Remote Sens.*, vol. 7, no. 6, pp. 1844–1868, June 2014.

[2] N. Dobigeon, J.-Y. Tourneret, C. Richard, J. Bermudez, S. McLaughlin, and A. Hero, "Nonlinear unmixing of hyperspectral images: Models and algorithms," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 82–94, Jan 2014.

[3] B. Somers, G. P. Asner, L. Tits, and P. Coppin, "Endmember variability in spectral mixture analysis: A review," *Remote Sensing of Environment*, vol. 115, no. 7, pp. 1603 – 1616, 2011.

[4] A. Zare and K. Ho, "Endmember variability in hyperspectral analysis: Addressing spectral variability during spectral unmixing," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 95–104, Jan 2014.

[5] A. Halimi, N. Dobigeon, and J.-Y. Tourneret, "Unsupervised unmixing of hyperspectral images accounting for endmember variability," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 4904–4917, 2015.

[6] C. Févotte and N. Dobigeon, "Nonlinear hyperspectral unmixing with robust nonnegative matrix factorization," *IEEE Trans. Image Process.*, pp. 4810–4819, Dec. 2015.

[7] Y. Altmann, S. McLaughlin, and A. Hero, "Robust linear spectral unmixing using anomaly detection," *IEEE Trans. Computational Imaging*, pp. 74–85, June 2015.

[8] A. Halimi, P. Honeine, and J. M. Bioucas-Dias, "Hyperspectral unmixing in presence of endmember variability, nonlinearity or mismodelling effects," in *ArXiv e-prints*, Nov. 2015.

[9] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*. New York: Springer-Verlag, 1999.

[10] O. Eches, N. Dobigeon, and J.-Y. Tourneret, "Enhancing hyperspectral image unmixing with spatial correlations," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, Nov 2011.

[11] Y. Altmann, N. Dobigeon, S. McLaughlin, and J.-Y. Tourneret, "Unsupervised post-nonlinear unmixing of hyperspectral images using a Hamiltonian Monte Carlo algorithm," *IEEE Trans. Image Process.*, vol. 23, no. 6, pp. 2663–2675, June 2014.

[12] Y. Altmann, M. Pereyra, and S. McLaughlin, "Bayesian nonlinear hyperspectral unmixing with spatial residual component analysis," *IEEE Trans. Image Process.*, vol. 1, no. 3, pp. 174–185, Sept. 2015.

[13] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. London, England: The MIT Press, 2006.

[14] O. Dikmen and A. Cemgil, "Gamma markov random fields for audio source modeling," *IEEE Trans. Audio, Speech, Language Process.*, vol. 18, no. 3, pp. 589–601, March 2010.

[15] A. Halimi, G. S. Buller, S. McLaughlin, and P. Honeine, "Bayesian filtering of smooth signals: Application to altimetry," in *ArXiv e-prints*, Feb. 2016.

[16] A. A. Kalaitzis and N. D. Lawrence, "Residual components analysis," in *Proc. ICML*, 2012, pp. 1–3.

[17] 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," *International Journal of Remote Sensing*, vol. 30, no. 11, pp. 2951–2962, June 2009.

[18] J. M. Bioucas-Dias and J. M. P. Nascimento, "Nonlinear mixture model for hyperspectral unmixing," in *Proc. SPIE Image and Signal Processing for Remote Sensing XV*, L. Bruzzone, C. Notarnicola, and F. Posa, Eds., vol. 7477, no. 1. SPIE, 2009, p. 74770I.

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

[20] 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 Process.*, vol. 21, no. 6, pp. 3017–3025, June 2012.

[21] G. E. Newstadt, A. O. Hero, and J. Simmons, "Robust spectral unmixing for anomaly detection," in *Proc. IEEE-SP Workshop Stat. and Signal Processing*, June 2014, pp. 109–112.

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

[23] D. P. Bertsekas, *Nonlinear programming*. Belmont, Massachusetts: Athena Scientific, 1995.

[24] Y. Altmann, N. Dobigeon, and J. Y. Tourneret, "Nonlinearity detection in hyperspectral images using a polynomial post-nonlinear mixing model," *IEEE Trans. Image Process.*, vol. 22, no. 4, pp. 1267–1276, April 2013.