



Heriot-Watt University
Research Gateway

Nonlinear hyperspectral unmixing accounting for spatial illumination variability

Citation for published version:

Halimi, A, Honeine, P, Bioucas-Dias, JM, Buller, GS & McLaughlin, S 2016, 'Nonlinear hyperspectral unmixing accounting for spatial illumination variability', Paper presented at 8th Workshop on Hyperspectral Image and Signal Processing 2016, Los Angeles, United States, 21/08/16 - 24/08/16.

Link:

[Link to publication record in Heriot-Watt Research Portal](#)

Document Version:

Peer reviewed version

General rights

Copyright for the publications made accessible via Heriot-Watt Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

Heriot-Watt University has made every reasonable effort to ensure that the content in Heriot-Watt Research Portal complies with UK legislation. If you believe that the public display of this file breaches copyright please contact open.access@hw.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

NONLINEAR HYPERSPECTRAL UNMIXING ACCOUNTING FOR SPATIAL ILLUMINATION VARIABILITY

Abderrahim Halimi¹, Paul Honeine², Jose Bioucas-Dias³, Gerald S. Buller¹, Steve McLaughlin¹

¹ Heriot-Watt University, School of Engineering and Physical Sciences, Edinburgh, U.K

² LITIS lab, Université de Rouen, France

³ Instituto de Telecomunicações and Instituto Superior Técnico, Universidade de Lisboa, Portugal.

ABSTRACT

This paper presents a new supervised algorithm for nonlinear hyperspectral unmixing. Based on the residual component analysis model, the proposed model assumes the linear model to be corrupted by an additive term that accounts for bilinear interactions between the endmembers. The proposed formulation considers also the effect of the spatial illumination variability. The parameters of the proposed model are estimated using a Bayesian strategy. This approach introduces prior distributions on the parameters of interest to take into account their known constraints. The resulting posterior distribution is optimized using a coordinate descent algorithm which allows us to approximate the maximum a posteriori estimator of the unknown model parameters. The proposed model and estimation algorithm are validated on both synthetic and real images showing competitive results regarding the quality of the inferences and the computational complexity when compared to the state-of-the-art algorithms.

Index Terms— Hyperspectral, nonlinear unmixing, Bayesian estimation, coordinate descent, gamma Markov random field

1. INTRODUCTION

Spectral unmixing (SU) consists of identifying the macroscopic materials (endmembers) present in a hyperspectral image and quantifying the proportions (abundances) of these materials in all pixels of the 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 if there are volumetric scattering, or terrain relief, or intimate mixtures of materials [1]. Nonlinear (NL) mixture models (NLMM) appear then as an alternative to better account for those effects [2, 3]. There exist two main families for NLMMs: the first family is signal processing based and seeks to construct flexible models that can represent a wide range of nonlinearities. The second family is physical based which includes models accounting for multiple scattering such as polynomial [4] or bilinear models [5–9].

This paper introduces a residual component (RC) mixture model for nonlinear hyperspectral unmixing. The model generalizes the well-known LMM by accounting for the spatial illumination variation and the presence of a nonlinear residual term. This term is a modification of the polynomial term proposed in [10], to account for the bilinear interactions between the endmembers. Estimating the abundances associated with this mixture model is a challenging

problem. We propose here a hierarchical Bayesian model to estimate the parameters and hyperparameters associated with the RC model. This hierarchical model introduces prior distributions that enforce known physical constraints on the estimated parameters such as the sum-to-one and positivity of the abundances, and the positivity of the nonlinear coefficients. Moreover, the spatial correlation of the residual term has been introduced by considering Markov random fields (MRF) [11]. Using the likelihood and the considered prior distributions, the joint posterior distribution of the unknown parameter vector is then derived. The minimum mean square error (MMSE) and maximum a posteriori (MAP) estimators of these parameters cannot be easily computed from the obtained joint posteriors. In this paper, the MAP estimator is evaluated by considering a coordinate descent algorithm (CDA) [12, 13] that sequentially updates the abundances, the noise variances and the residual term. The proposed Bayesian model and estimation algorithm are validated using synthetic and real hyperspectral images. The obtained results are very promising and show the potential of the proposed mixture and Bayesian models and their associated inference algorithm.

The paper is structured as follows. Section 2 introduces the proposed mixture model to deal with the NL. The proposed hierarchical Bayesian model and its estimation algorithm are introduced in Sections 3 and 4. Section 5 is devoted to testing and validating the proposed technique using synthetic images with known ground truth. Section 6 shows results obtained using a real hyperspectral image. Conclusions and future work are finally reported in Section 7.

2. MIXTURE MODEL

The LMM is widely used because of its simplicity. However, there are a lot of situations where the linear model is not valid because of the presence of nonlinearity or other mismodelling effects. These effects require the elaboration of a more sophisticated model to improve the abundance estimates. This paper proposes a new formulation based on a residual component model [14] 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 $\mathbf{y}_{i,j}$ is given by

$$\mathbf{y}_{i,j} = c_{i,j} \mathbf{M} \mathbf{a}_{i,j} + \phi_{i,j}^{NL}(\mathbf{M}) + \mathbf{e}_{i,j} \quad (1)$$

where $\mathbf{a}_{i,j} = (a_{1,i,j}, \dots, a_{R,i,j})^T$ is an $(R \times 1)$ vector of abundances associated with the pixel (i, j) and satisfying the positivity and sum-to-one constraints ($a_{r,i,j} \geq 0, \forall r \in \{1, \dots, R\}$ and $\sum_{r=1}^R a_{r,i,j} = 1$), R is the number of endmembers, $\mathbf{e}_{i,j} \sim \mathcal{N}(\mathbf{0}_L, \Sigma)$ is an additive centered Gaussian noise with a diagonal covariance matrix $\Sigma = \text{diag}(\sigma^2)$, $\sigma^2 = (\sigma_1^2, \dots, \sigma_L^2)^T$ is an $(L \times 1)$ vector containing the noise variances, L is the number of

This work was supported in part by the EPSRC Grants EP/JO15180/1, EP/N003446/1, and EP/K015338/1, and in part by the HYPANEMA ANR Project under Grant ANR-12-BS03-003, and by Portuguese Fundação para a Ciência e Tecnologia (FCT), grant UID/EEA/5008/2013.

spectral bands, \mathbf{M} is a fixed endmember matrix that is assumed known (extracted using an endmember extraction algorithm) and $\phi_{i,j}^{NL}$ is a residual term that accounts for NL interactions. Similarly to [15, 16], model (1) includes the effect of illumination variation via the pixel dependent illumination coefficient $c_{i,j}$ (we consider the same parameter for all the endmembers). Moreover, the effect of bilinear interactions between the endmembers is introduced by considering the following residual term

$$\phi_{i,j}^{NL}(\mathbf{M}) = c_{i,j}^2 \mathbf{Q}(\mathbf{M}) \boldsymbol{\gamma}_{i,j} \quad (2)$$

where $\boldsymbol{\gamma}_{i,j} = \left(\gamma_{i,j}^{(1)}, \dots, \gamma_{i,j}^{(R)}, \gamma_{i,j}^{(1,2)}, \dots, \gamma_{i,j}^{(R-1,R)} \right)^T$, $\forall i, j$ is the $(D \times 1)$ vector of positive nonlinear coefficients, $D = \frac{R(R+1)}{2}$, $\mathbf{Q}(\mathbf{M}) = (\mathbf{m}_1 \odot \mathbf{m}_1, \dots, \mathbf{m}_R \odot \mathbf{m}_R, \sqrt{2}\mathbf{m}_1 \odot \mathbf{m}_2, \dots, \sqrt{2}\mathbf{m}_{R-1} \odot \mathbf{m}_R)$, is the $(L \times D)$ matrix gathering the interaction spectra, and \odot denotes the Hadamard (termwise) product. Note that model (1) reduces to the model [10] for $c_{i,j} = 1, \forall i, j$ and to the LMM for $\boldsymbol{\gamma}_{i,j} = 0$, and $c_{i,j} = 1, \forall i, j$. In addition, model (1) has a polynomial-like form as for the bilinear models (GBM [5], PP-NMM [4], Nascimento [7], Fan [8] and Meganem [9] models). Note finally that model (1) (with no illumination variation) has been studied in [10] when considering a Markov chain Monte-Carlo (MCMC) approach and have shown good performance for processing hyperspectral images. However, the MCMC estimation algorithm was computationally expensive, and we consider in this paper a faster algorithm based on a coordinate descent algorithm.

3. HIERARCHICAL BAYESIAN MODEL

The known parameters associated with model (1) are $\Theta = (\mathbf{A}, \mathbf{c}, \boldsymbol{\Gamma}, \boldsymbol{\sigma}^2, \boldsymbol{\epsilon})$, where \mathbf{A} (resp. $\boldsymbol{\Gamma}$) gathers all the abundance vectors (resp. nonlinear coefficients), and $\boldsymbol{\epsilon}$ is a hyperparameter described later in the text. A hierarchical Bayesian model is considered for the estimation of Θ . This approach is well suited for ill-posed inverse problems, since it accounts for the known parameter constraints using a priori distributions. The next sections present the likelihood and the proposed parameter prior distributions.

3.1. Likelihood

Using the observation model (1), the Gaussian properties of the noise sequence $e_{i,j}$, and exploiting independence between the observations in different spectral bands, yield the following Gaussian distribution for the likelihood

$$\mathbf{y}_{i,j} | \mathbf{a}_{i,j}, c_{i,j}, \boldsymbol{\gamma}_{i,j}, \boldsymbol{\Sigma} \sim \mathcal{N}(\boldsymbol{\mu}_{i,j}, \boldsymbol{\Sigma}). \quad (3)$$

where $\boldsymbol{\mu}_{i,j} = c_{i,j} \mathbf{M} \mathbf{a}_{i,j} + \phi_{i,j}^{NL}(\mathbf{M})$ and \sim means ‘‘is distributed according to’’. The independence between the observed pixels leads to $f(\mathbf{Y} | \Theta) = \prod_i \prod_j f(\mathbf{y}_{i,j} | \mathbf{a}_{i,j}, c_{i,j}, \boldsymbol{\gamma}_{i,j}, \boldsymbol{\Sigma})$, where \mathbf{Y} gather all the pixels.

3.2. Parameter priors

This section introduces the prior distributions that we have chosen for the parameters of interest \mathbf{A} , \mathbf{c} , $\boldsymbol{\Gamma}$ and $\boldsymbol{\Sigma}$.

3.2.1. Abundance matrix \mathbf{A}

The abundances should satisfy the positivity and sum-to-one (PSTO) constraints. Akin to [5, 17], and since there is no additional information about \mathbf{A} , we assign the abundances a uniform prior in the simplex $\mathcal{S} = \left\{ \mathbf{a}_{i,j} | a_{r,i,j} \geq 0, \forall r \text{ and } \sum_{r=1}^R a_{r,i,j} = 1 \right\}$.

3.2.2. Prior for \mathbf{c}

In absence of illumination variability, the parameter $c_{i,j}$ has a fixed value $\#1$ for all pixels [5, 18, 19]. This paper accounts for the illumination variability by allowing this parameter to fluctuate around $\#1$. This is achieved by considering the conjugate Gaussian prior $c_{i,j} \sim \mathcal{N}(1, \eta^2)$, where η^2 is a small fixed variance ($\eta^2 = 0.01$ in the rest of the paper). For simplicity, we denote ‘‘ $x | \theta \sim \dots$ ’’, by ‘‘ $x \sim \dots$ ’’ when the parameter θ is a user fixed parameter. Note finally that the joint prior of \mathbf{c} is obtained by assuming a priori independence between the coefficients $c_{i,j}$, as follows $f(\mathbf{c}) = \prod_{i,j} f(c_{i,j})$.

3.2.3. Nonlinear coefficients $\boldsymbol{\gamma}_{i,j}$

Due to physical constraints, the nonlinear coefficients should satisfy the positivity constraint. Similarly to [10], $\boldsymbol{\gamma}_{i,j}$ are assigned the following truncated Gaussian prior:

$$\boldsymbol{\gamma}_{i,j} | \epsilon_{i,j}^2 \sim \mathcal{N}_{(R+)^D}(\mathbf{0}_D, \epsilon_{i,j}^2 \mathbb{I}_D), \quad (4)$$

where \mathbb{I}_D denotes the $D \times D$ identity matrix and $\epsilon_{i,j}^2$ is a variance parameter that is pixel dependent. From (4), it is clear that this variance is related to the strength of the nonlinearities at the pixel (i, j) (via the norm $\|\boldsymbol{\gamma}_{i,j}\|^2$). Moreover, as in [10], we assume the nonlinear energies to vary smoothly from one pixel to another which will be introduced by considering a specific prior for $\epsilon_{i,j}^2$, as explained in Section 3.3. Note finally that the joint prior of $\boldsymbol{\Gamma}$ is obtained by assuming a priori independence between the nonlinear coefficients, as follows $f(\boldsymbol{\Gamma} | \boldsymbol{\epsilon}) = \prod_{i,j} f(\boldsymbol{\gamma}_{i,j} | \epsilon_{i,j}^2)$.

3.2.4. Noise variances

The noise variances are assigned a conjugate inverse gamma distribution as follows: $\sigma_\ell^2 \sim \mathcal{IG}(\varphi_\ell, \psi_\ell)$, where σ_ℓ^2 are assumed a priori independent leading to $f(\boldsymbol{\sigma}^2) = \prod_{\ell=1}^L f(\sigma_\ell^2)$. The hyperparameters φ_ℓ and ψ_ℓ are fixed to approximate the HySime estimated variances [20]. Note finally that the hyperparameters can also be set to $\varphi_\ell = \psi_\ell = 0$ in absence of prior knowledge about σ_ℓ^2 , leading to a noninformative Jeffreys’ prior.

3.3. Hyperparameter priors

Due to the spatial organization of hyperspectral images, we expect the energies of the nonlinear coefficients $\boldsymbol{\gamma}_{i,j}$ to vary smoothly from one pixel to another. This behavior is obtained by introducing an auxiliary variable \mathbf{w} (of size $N_{\text{row}} \times N_{\text{col}}$) and assigning a gamma Markov random field (GMRF) prior for $(\boldsymbol{\epsilon}, \mathbf{w})$ (see [10, 11] for more details regarding this prior). An interesting property of this joint prior is that the conditional prior distributions of $\boldsymbol{\epsilon}$ and \mathbf{w} reduce to conjugate inverse gamma (\mathcal{IG}) and gamma (\mathcal{G}) distributions as follows

$$\begin{aligned} \epsilon_{i,j}^2 | \mathbf{w}, \zeta &\sim \mathcal{IG}(4\zeta, 4\zeta \rho_{1,i,j}(\mathbf{w})), \\ \mathbf{w}_{i,j}^2 | \boldsymbol{\epsilon}, \zeta &\sim \mathcal{G}(4\zeta, 1/(4\zeta \rho_{2,i,j}(\boldsymbol{\epsilon}))), \end{aligned} \quad (5)$$

where

$$\begin{aligned} \rho_{1,i,j}(\mathbf{w}) &= (w_{i,j}^2 + w_{i+1,j}^2 + w_{i,j+1}^2 + w_{i+1,j+1}^2)/4, \\ \rho_{2,i,j}(\boldsymbol{\epsilon}) &= (\epsilon_{i,j}^{-2} + \epsilon_{i-1,j}^{-2} + \epsilon_{i,j-1}^{-2} + \epsilon_{i-1,j-1}^{-2})/4, \end{aligned} \quad (6)$$

and ζ is a fixed coupling parameter that controls the amount of spatial smoothness enforced by the GMRF. Note that (5) ensures that each $\epsilon_{i,j}^2$ is connected to four neighbor elements of \mathbf{w} and vice-versa

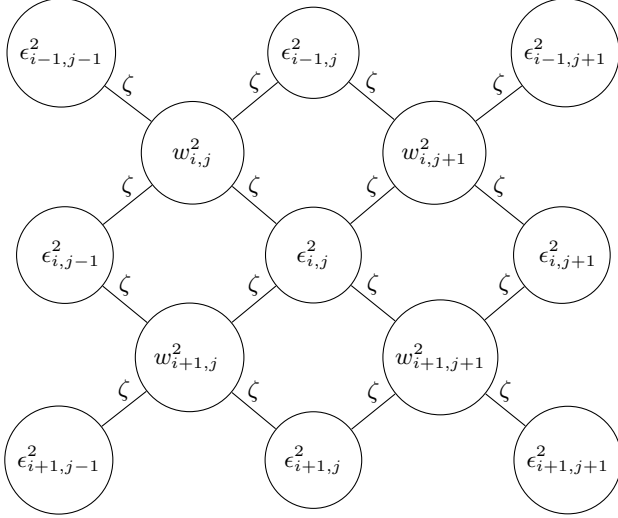


Fig. 1. Considered 1st order GMRF neighborhood structure.

(see Fig. 1). Note also that the energies $\epsilon_{i,j}^2$ are conditionally independent and the 1st order neighbors (i.e., the spatial correlation) is introduced via the interaction between ϵ and the auxiliary variables w .

3.4. Posterior distributions

The joint posterior distribution of the hierarchical Bayesian model can be computed from the following Bayes' rule

$$f(\Theta|\mathbf{Y}) \propto f(\mathbf{Y}|\Theta)f(\mathbf{A})f(\sigma^2)f(c)f(\gamma|\epsilon)f(\epsilon, \mathbf{w}), \quad (7)$$

where \propto means “proportional to” and we have assumed a priori independence between the parameters of each model. The MMSE and MAP estimators associated with the posterior (7) are not easy to determine. In this paper, and akin to [13], we propose to evaluate the MAP estimator by using an optimization technique maximizing the posterior (7) w.r.t. the parameters of interest.

4. COORDINATE DESCENT ALGORITHM

As a result of the large number of parameters to estimate, we propose a coordinate descent algorithm [12, 13] that sequentially updates the different parameters. In each step of the proposed CDA-NL algorithm, 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 as follows (see [21] for more details regarding the conditional distributions):

- Conditional of \mathbf{A} : truncated Gaussian distribution (whose one maximum is obtained with SUNSAL-FCLS¹ [19])
- Conditional of $\mathbf{\Gamma}$: positive truncated Gaussian distribution (whose one maximum is obtained with SUNSAL-CLS [19])
- Conditional of ϵ : inverse gamma distribution (analytical expression of the maximum)
- Conditional of \mathbf{w} : gamma distribution (analytical expression of the maximum)

¹SUNSAL-FCLS satisfies the PSTO constraints while SUNSAL-CLS only ensures the positivity constraint.

- Conditional of σ^2 : inverse gamma distribution (analytical expression of the maximum)
- Conditional of c : Gaussian distribution (analytical expression of the mean).

Regarding the sequence generated by the coordinate descent algorithm, the proposition 2.7.1 in [12] asserts that its limit points are stationary points of (7) provided that the maximum of that function w.r.t. Θ along each coordinate is unique. This is easily checked for all the parameters except for c^{NL} . Indeed, the cost function writes as a 4-order polynomial w.r.t. c^{NL} (leading to 3 possible maxima) and we have chosen the one that maximizes it in the interval [0.2, 3]. Note also that the cost function is not convex, thus, the solution obtained might depend on the initial values that need to be chosen carefully. Therefore, the abundances \mathbf{A} are initialized with SUNSAL-FCLS [19], the nonlinear coefficients are initialized to 0, the noise variance is initialized with the HySime estimate [20], the illumination coefficient c is initialized by considering the sum of the abundances that are estimated using only the positivity constraint with SUNSAL-CLS [19]. With these initializations, the proposed algorithm reached systematically minima of good quality in the considered simulations (see Sections 5 and 6).

5. SIMULATION RESULTS ON SYNTHETIC DATA

This section evaluates the performance of the proposed CDA-NL algorithm on a synthetic image denoted by I . This image contains 100×100 pixels, $L = 207$ spectral bands and has been generated using $R = 3$ endmembers extracted from the ENVI software library. The image has also been corrupted by i.i.d. Gaussian noise for a fair comparison with SU algorithms using this assumption. To investigate the behavior of the proposed approach w.r.t. different mixture models, image I has been generated according to 4 linear/nonlinear models. For this, an image partition into 4 classes has been generated by considering a Potts-Markov random field. The four spatial classes are associated with the LMM, NL model (1) (with $\epsilon^2 = 0.1$), GBM (with random nonlinear coefficients in [0.8; 1]) and PPNMM (with $b = 0.5$), respectively. Note that the abundances have been clustered inside the simplex using a Dirichlet distribution and that the illumination coefficient c increases linearly from the left to the right of the image in [0.9, 1.15]. Table 1 compares the proposed CDA-NL algorithm with the FCLS [18], SUNSAL-CLS [19], SKhype [22], and RCA-MCMC [23] algorithms. The criteria used are the quadratic distance between the actual and estimated abundances (RMSE), the reconstruction error (RE) and the spectral angle mapper (SAM) between the observed and reconstructed pixel spectra. Table 1 highlights the competitive performance of CDA-NL for the different mixture models and shows that CDA-NL has the best overall results. It also shows that CDA-NL has competitive computational times with respect to NL unmixing algorithms SKhype and RCA-MCMC.

6. RESULTS ON REAL DATA

This section illustrates the performance of the proposed algorithm when applied to a real hyperspectral image. This image was acquired over Moffett Field, CA, in 1997 by the AVIRIS. The considered dataset contains 100×100 pixels, $L = 152$ spectral bands (after removing water absorption bands) acquired in the interval $0.4 - 2.5\mu\text{m}$, has a spatial resolution of 100m and is mainly composed of three components: water, soil, and vegetation (see Fig. 2 (Top-left)). The VCA algorithm [24] was used to estimate $R = 3$ endmembers. The estimated CDA-NL abundance maps are in good agreement with

Table 1. Results on synthetic data.

	RMSE (classes) ($\times 10^{-2}$)				RMSE ($\times 10^{-2}$)	RE ($\times 10^{-2}$)	SAM ($\times 10^{-2}$)	Time (s)
	\mathcal{C}_1	\mathcal{C}_2	\mathcal{C}_3	\mathcal{C}_4				
	LMM	RCA-NL	GBM	PPNMM				
FCLS	10.24	44.72	15.48	23.98	24.76	15.74	10.64	1.7
SUNSAT-CLS	3.84	33.81	5.68	8.45	16.55	4.17	7.57	0.07
SKhype	1.67	11.92	2.21	2.81	5.87	—	—	547
RCA-MCMC	5.87	6.29	5.44	3.93	5.66	—	—	9009
CDA-NL	1.62	7.27	2.16	2.89	3.86	2.86	6.16	430

the state-of-the-art algorithms and are not shown here for brevity (see [21] for more details). Fig. 2 (top-right) shows the energies of the difference between the reconstructed signal and the linear model. It is clear that the residual components are mainly located in the coastal region and in water. Figs. 2 (bottom-left) and (bottom-right) show the effects of the illumination variation and the nonlinear term, respectively. It can be seen that the illumination variation mainly occurs in water while the nonlinear effects are located in the coastal zone and in presence of vegetation.

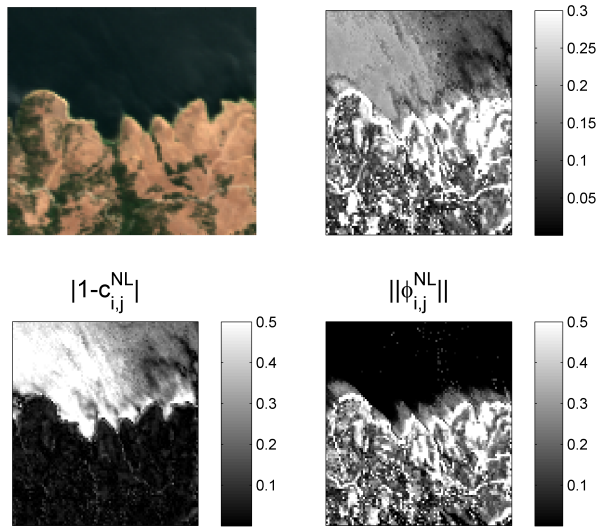


Fig. 2. (Top-left) Real Moffett image. (Top-right) Square root of the energies of the difference between the reconstructed signal and the linear model ($\|\hat{\mathbf{y}}_{i,j} - \mathbf{M}\hat{\mathbf{a}}_{i,j}\|$). (Bottom-left) estimated illumination variation $|1 - c_{i,j}|$. (Bottom-right) square root of the energies of the residual terms $\|\phi_{i,j}^{NL}\|$.

7. CONCLUSIONS

This paper introduced a hyperspectral mixture model and its associated Bayesian algorithm for nonlinear hyperspectral unmixing. The proposed model considered the spatial illumination variation and the presence of nonlinearity effects. A hierarchical Bayesian model was proposed to introduce the known constraints on the parameters of interest. Those parameters were estimated using a coordinate descent

algorithm that showed a reduced computational cost when compared to state-of-the-art algorithms. The proposed algorithm showed good performance when processing synthetic data generated with the linear and nonlinear models. Results on real data confirmed the good performance of the proposed algorithm and showed its ability to extract different features in the observed scene. Future work includes the estimation of the hyperparameters associated with the proposed Bayesian model. Generalizing the mixture model to include other forms of endmember variability is also an interesting issue which would deserve to be investigated.

8. REFERENCES

- [1] B. W. Hapke, “Bidirectional reflectance spectroscopy. I. Theory,” *J. Geophys. Res.*, vol. 86, pp. 3039–3054, 1981.
- [2] 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.
- [3] 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.
- [4] 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.
- [5] 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.
- [6] —, “Unmixing hyperspectral images using the generalized bilinear model,” in *Proc. IEEE Int. Conf. Geosci. Remote Sens. (IGARSS)*, July 2011, pp. 1886–1889.
- [7] 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.
- [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,” *International Journal of Remote Sensing*, vol. 30, no. 11, pp. 2951–2962, June 2009.

- [9] I. Meganem, P. Deliot, X. Briottet, Y. Deville, and S. Hosseini, "Linear-quadratic mixing model for reflectances in urban environments," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 544–558, Jan 2014.
- [10] Y. Altmann, M. Pereyra, and S. McLaughlin, "Bayesian nonlinear hyperspectral unmixing with spatial residual component analysis," *IEEE Trans. Comput. Imaging*, vol. 1, no. 3, pp. 174–185, Sept. 2015.
- [11] 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.
- [12] D. P. Bertsekas, *Nonlinear programming*. Belmont, Massachusetts: Athena Scientific, 1995.
- [13] A. Halimi, C. Mailhes, J.-Y. Tourneret, and H. Snoussi, "Bayesian estimation of smooth altimetric parameters: Application to conventional and delay/Doppler altimetry," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 4, Dec. 2015.
- [14] A. A. Kalaitzis and N. D. Lawrence, "Residual components analysis," in *Proc. ICML*, 2012, pp. 1–3.
- [15] M. A. Veganzones, L. Drumetz, G. Tochon, M. Dalla Mura, A. Plaza, J. M. Bioucas-Dias, and J. C., "A New Extended Linear Mixing Model to Address Spectral Variability," in *IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS 2014)*, Lausanne, Switzerland, June 2014.
- [16] G. Shaw and H. Burke, "Spectral imaging for remote sensing," *Lincoln Lab. J.*, vol. 14, no. 1, pp. 3–28, 2003.
- [17] N. Dobigeon, J.-Y. Tourneret, and C.-I Chang, "Semi-supervised linear spectral unmixing using a hierarchical Bayesian model for hyperspectral imagery," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 2684–2695, July 2008.
- [18] 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. Remote Sens.*, vol. 29, no. 3, pp. 529–545, March 2001.
- [19] 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.
- [20] J. M. Bioucas-Dias and J. M. P. Nascimento, "Hyperspectral subspace identification," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 8, pp. 2435–2445, Aug. 2008.
- [21] 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.
- [22] 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, Jan 2013.
- [23] Y. Altmann, N. Dobigeon, S. McLaughlin, and J.-Y. Tourneret, "Residual component analysis of hyperspectral images: Application to joint nonlinear unmixing and nonlinearity detection," *IEEE Trans. Image Process.*, vol. 23, no. 5, pp. 2148–2158, May 2014.
- [24] J. M. P. Nascimento and J. M. Bioucas-Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 898–910, April 2005.