A bilinear-bilinear Non-Negative Matrix Factorization method for Hyperspectral Unmixing

Olivier Eches and Mireille Guillaume, Ecole Centrale Marseille, Aix Marseille Université, CNRS, Institut Fresnel UMR 7122, 13397 Marseille, France.

Abstract—**Spectral unmixing of hyperspectral images consists of estimating pure material spectra with their corresponding proportions (or *abundances*). Non-linear mixing models for spectral unmixing are of very recent interest within the signal and image processing community. This letter proposes a new non-linear unmixing approach using the Fan bilinear-bilinear model and non-negative matrix factorization method that takes into account physical constraints on spectra (positivity) and abundances (positivity and sum-to-one). The proposed method is tested using a projected Gradient algorithm on synthetic and real data. The performances of this method are compared to linear approach and to recent non-linear approach.**

I. INTRODUCTION

Over the last decade, hyperspectral image unmixing has become one of the greatest topic of interest in the geoscience and remote sensing community. Classical unmixing algorithms assume that the image pixels are linear combinations of a given number of pure materials spectra or *endmembers* with corresponding fractions referred to as *abundances*. This linear mixing model has been widely employed by the community and provided interesting results (see [1] and references therein). However, this naive model proved to be inappropriate for complex environments such as forests [2] areas. In such situations, non-linear mixing models have been recently viewed as an alternative to the linear mixing model. In [3], Hapke first introduced a bidirectional reflectance-based model for intimate mixtures representation. Other types of models have been proposed for example in [4] where kernel functions are employed or in [5] with neural networks modeling non-linearities. In this letter, we are focusing on a quadratic-quadratic model introduced by Fan in [2] which casts the scattering effects of light in vegetation areas. Assuming that $\boldsymbol{R} = [\boldsymbol{r}_1, \ldots, \boldsymbol{r}_I]$ is the matrix representing the observed spectra on L bands of the considered image of I pixels, the model can be written as

$$\boldsymbol{R} = \boldsymbol{S}\boldsymbol{A} + \boldsymbol{S}_b\boldsymbol{A}_b + \boldsymbol{N}, \qquad (1)$$

with $\boldsymbol{S}$ representing the endmembers matrix of size $L \times J$ (J being the number of endmembers) and $\boldsymbol{A}$ the abundances matrix. $\boldsymbol{N}$ is the L-dimensional noise that can be distributed according to a normal distribution with zero mean and same variance for each band $\mathcal{N}(\boldsymbol{0}_L, \sigma_n^2 \boldsymbol{I}_L)$. $\boldsymbol{S}_b = [\boldsymbol{s}_1 \odot \boldsymbol{s}_2, \ldots, \boldsymbol{s}_{J-1} \odot \boldsymbol{s}_J]$ ($\odot$ is the Hadamard product) and $\boldsymbol{A}_b = [\boldsymbol{a}_1 \odot \boldsymbol{a}_2, \ldots, \boldsymbol{a}_{J-1} \odot \boldsymbol{a}_J]^T$ are respectively $L \times J(J-1)/2$ and $J(J-1)/2 \times I$ matrices that represent the non-linearities. More precisely, this model stems from a polynomial representation of light interactions with the different surfaces

inside a pixel. Then, the non-linear terms were obtained by retaining the first-order terms of the Taylor series development of this polynomial function, that represents only the second scattering between a couple of endmembers. Since abundances are proportions, they are submitted to positivity and sum-to-one constraints [2], i. e.

$$\boldsymbol{1}_{1J}\boldsymbol{A} = \boldsymbol{1}_{1I}, \quad \forall i, j \ \boldsymbol{A}_{ij} > 0, \qquad (2)$$

where $\boldsymbol{1}_{1J}$ and $\boldsymbol{1}_{1I}$ are the matrices of "ones" of size respectively $1 \times J$ and $1 \times I$. In Fig. 1, a scatterplot representation for 3 endmembers is given, thus showing the non-linearities influence on the linear model geometry.

Several contributions focused on linear-quadratic mixtures for blind source separation (BSS) problems [6] in the past years, as in [7]. Recently, the authors of [8] proposed a Bayesian source separation method combined with a MCMC (Markov Chain Monte Carlo) simulation method for estimating the parameters, applied for separation of scanned images. However, the proposed prior distributions were truncated and the method is computationally intensive. Alternatively, non-negative matrix factorization (NMF) methods [9] can be used to solve BSS problems under non-negativity constraints as in [10] where it has been successfully applied to identify constituent in chemical shift imaging or in [11] and [12] for hyperspectral linear unmixing. Indeed, jointly estimating the spectra and the corresponding abundances in a single step can be considered as a BSS. Recently, in [13], Meganem *et al.* proposed a NMF method for non-linear spectral unmixing based on a linear-quadratic model, different from (1), which is quadratic-quadratic. Indeed, this method, which is based on an urban scenario, do not integrate any sum-to-one constraint nor the bilinear dependance on the abundances. Other contributions focused on non-linear unmixing using a so-called *generalized bilinear model* [14], [15], very similar to (1). The authors present optimization methods [15] to solve the unmixing problem where the sum-to-one constraint is taken into account. Yet, it assumes the endmembers are known, i.e. estimated by endmember extraction algorithms like Vertex Component Analysis (VCA) [16] or extracted from a spectral library. Moreover, the algorithms are applied pixelwise.

In this letter, we propose a new NMF method for bilinear-bilinear spectra and abundances unmixing, using Fan model [2]. Inspired from [13] for the matrices construction and from [12] for the optimization method, we compute the gradients respectively for $\boldsymbol{S}$ and $\boldsymbol{A}$. The sum-to-one constraint for the abundances has been introduced as a regularization term in the objective function. To insure the positivity of the spectra and abundances matrices, the projected gradient algorithm [17]
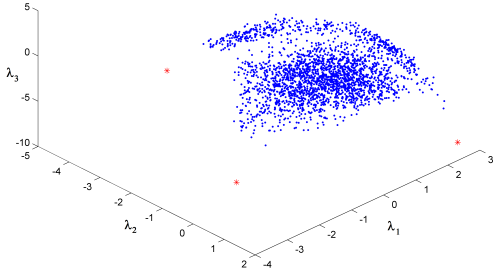
corresponding regularization parameter δ depends on the data (noise and level of mixture) and should be chosen by taking the values that minimize the final error[1].

Due to the model structure (1), it is not possible to apply NMF method to our model unless undergoing some reformulations as in [8], [13], i.e. by noting $\boldsymbol{S}^* = [\boldsymbol{S}, \boldsymbol{S}_b]$ and $\boldsymbol{A}^* = \left[\boldsymbol{A}^T, \boldsymbol{A}_b^T\right]^T$. Thus, the model can be rewritten as $\boldsymbol{R} = \boldsymbol{S}^*\boldsymbol{A}^* + \boldsymbol{N}$. From this reformulated model, contrary to the work in [13], the complete matrices $\boldsymbol{S}$ and $\boldsymbol{A}$ will be estimated. Indeed, this approach allows the non-linear part to be integrated in the computation structure of the NMF method, and therefore proposing solution closer to the theoretic model.

The complete objective function to be minimized is

$$\boldsymbol{f}(\boldsymbol{S},\boldsymbol{A}) = \|\boldsymbol{R} - \boldsymbol{S}^*\boldsymbol{A}^*\|_F^2 + \delta(\boldsymbol{1}_{1J}\boldsymbol{A} - \boldsymbol{1}_{1I})^T(\boldsymbol{1}_{1J}\boldsymbol{A} - \boldsymbol{1}_{1I}).$$

B. Updating rules

In order to guarantee positivity, we will apply the alternated Projected Gradient (PG) algorithm that has already been successfully applied in [22] and [13] for unmixing hyperspectral data. The PG algorithm iteratively updates $\boldsymbol{S}$ and $\boldsymbol{A}$ as a classical gradient descent method and project the updated values on $\mathbb{R}^+$ to ensure the strong positivity constraint. Thus, the gradients of $\boldsymbol{f}$ with respect to $\boldsymbol{S}$ and $\boldsymbol{A}$, i.e. $\nabla_{\boldsymbol{S}}\boldsymbol{f}$ and $\nabla_{\boldsymbol{A}}\boldsymbol{f}$, must be determined. Note that the sum-to-one regularization term can be included in $\boldsymbol{R}$ and $\boldsymbol{S}^*$ by letting $\tilde{\boldsymbol{S}}^* = \left[\boldsymbol{S}^{*T}, \delta \times \boldsymbol{1}_{J^*1}\right]^T$, $\tilde{\boldsymbol{R}} = \left[\boldsymbol{R}^T, \delta \times \boldsymbol{1}_{I1}\right]^T$ where $J^* = J + \frac{J(J-1)}{2}$.

The computations lead to the following expressions for the gradients

$$\begin{aligned}
\nabla_{\boldsymbol{S}}\boldsymbol{f} &= -\left[(\boldsymbol{R} - \boldsymbol{S}^*\boldsymbol{A}^*) \odot (\boldsymbol{S}\boldsymbol{A} + 1)\right]\boldsymbol{A}^T \\
&\quad + \left[(\boldsymbol{R} - \boldsymbol{S}^*\boldsymbol{A}^*)(\boldsymbol{A} \odot \boldsymbol{A})^T\right] \odot \boldsymbol{S},
\end{aligned}$$

$$\begin{aligned}
\nabla_{\boldsymbol{A}}\boldsymbol{f} &= -\tilde{\boldsymbol{S}}^T\left[\left(\tilde{\boldsymbol{R}} - \tilde{\boldsymbol{S}}^*\boldsymbol{A}^*\right) \odot \left(\tilde{\boldsymbol{S}}\boldsymbol{A} + 1\right)\right] \\
&\quad + \boldsymbol{A} \odot \left[\left(\tilde{\boldsymbol{S}} \odot \tilde{\boldsymbol{S}}\right)^T \left(\tilde{\boldsymbol{R}} - \tilde{\boldsymbol{S}}^*\boldsymbol{A}^*\right)\right].
\end{aligned}$$

The updating rules are then given by

$$\begin{aligned}
\boldsymbol{S}^{(t+1)} &= P\left[\boldsymbol{S}^{(t)} - \alpha_{\boldsymbol{S}}^{(t)}\nabla_{\boldsymbol{S}}\boldsymbol{f}\right], \\
\boldsymbol{A}^{(t+1)} &= P\left[\boldsymbol{A}^{(t)} - \alpha_{\boldsymbol{A}}^{(t)}\nabla_{\boldsymbol{A}}\boldsymbol{f}\right],
\end{aligned}$$

where $P[\cdot]$ represents the projection on $\mathbb{R}^+$ and $\alpha_{\boldsymbol{S}}^{(t)}$, $\alpha_{\boldsymbol{A}}^{(t)}$ are respectively the updating steps of $\boldsymbol{S}$ and $\boldsymbol{A}$ that guarantee convergence of the PG algorithm. They are estimated at each iteration using a modified Armijo rule described in [12].

III. SIMULATION RESULTS

In this section, the proposed method (FAN-NMF) is compared with another NMF approach developed for linear unmixing model named Maximum Dispersion Minimum Dispersion-NMF (MDMD-NMF) [22], using only the sum-to-one regularization constraint. We also compare the proposed method

---

[1]Note this does not give unique solution to NMF method, but only restricts the solution subset.

Fig. 1. Cluster representation of data mixtures (blue) generated by a Fan model with 3 endmembers (red), and maximum abundance a_{max}=0.7

is employed with varying step sizes, involving a modified Armijo-Lin rule [12]. The proposed method has been successively tested on synthetic and real hyperspectral images, and compared with other non-linear unmixing algorithms proposed in the literature.

II. NMF FOR NON-LINEAR UNMIXING WITH FAN MODEL

In this section, after recalling the principle of NMF, the updating rules for the parameters of interests are detailed.

A. Principle of NMF

The basic NMF formulation consists of approximating a nonnegative matrix $\boldsymbol{R}$ by a two matrices product [9], [17]

$$\boldsymbol{R} \simeq \hat{\boldsymbol{S}}\hat{\boldsymbol{A}}. \tag{3}$$

These matrices can be found by solving the resulting optimization problem

$$\min_{\boldsymbol{S},\boldsymbol{A}} D(\boldsymbol{R}|\boldsymbol{S}\boldsymbol{A}) \text{ s. c. } \boldsymbol{S} \geq 0, \boldsymbol{A} \geq 0 \tag{4}$$

where $D(\cdot|\cdot)$ is a divergence measure. The Frobenius norm is considered as the divergence measure by default in this work, i.e. $D(\boldsymbol{R}|\boldsymbol{S}\boldsymbol{A}) = \|\boldsymbol{R} - \boldsymbol{S}\boldsymbol{A}\|_F^2$. The strong positivity constraints are fundamental since they ensure the existence of stationary point, which is necessary to obtain approximate solutions. A wide range of methods have been developed to solve this problem. Lee and Seung in [18], [19] proposed a multiplicative approach. Then, in order to improve the convergence speed, a projected gradient method has been proposed in [20]. Solutions obtained by the NMF methods are not guaranteed to be unique. Indeed, for any invertible matrix $\boldsymbol{M}$, the equality $\hat{\boldsymbol{S}}\hat{\boldsymbol{A}} = (\hat{\boldsymbol{S}}\boldsymbol{M})(\boldsymbol{M}^{-1}\hat{\boldsymbol{A}})$ is still valid. It is possible to alleviate this problem by appropriate initialisations and adding regularization constraints as in [21] where a sparsity constraint has been proposed. In this contribution, since the abundances are subject to the sum-to-one constraint (2), a corresponding regularization term is added. In [15], the sum-to-one constraint is forced by estimating only the $J-1$ abundance coefficients while the remaining coefficient is deduced using the sum-to-one relation. However, in this work, a smooth constraint has been chosen for the Fan model. The
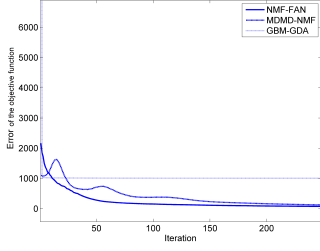
Fig. 2. Objective function of the proposed algorithm Fan-NMF, the linear MDMD-NMF and the gradient descent algorithm GBM-GDA, for $a_{max} = 0.7$.
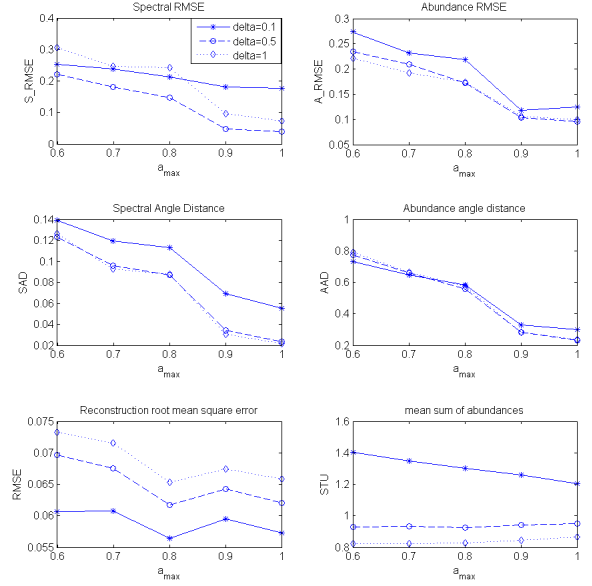

Fig. 3. Errors on abundances and spectra using the proposed method, for $\delta = 0.1$ (*), 0.5 (o), 1 (diamond). Bottom: reconstruction RMSE and final sum-to-one after estimation.

with the gradient descent algorithm for unmixing bilinear models proposed in [15] (GBM-GDA), which is based on a generalized bilinear model, including Fan model as a particular case, but is focused only on abundance estimations. The generalized bilinear model is given by $\mathbf{r} = \sum_{l=1}^{J} a_l \boldsymbol{s}_l + \sum_{i=1}^{J-1} \sum_{j=i+1}^{J} \gamma_{ij} a_i a_j \boldsymbol{s}_i \odot \boldsymbol{s}_j + \mathbf{n}$, where $0 \leq \gamma_{ij} \leq 1$ is a parameter to be estimated that tunes the interactions terms between the different endmembers, and a_i the ith endmember abundance.

A. Synthetic data

The synthetic data set, composed of $N = 1000$ pixels and $L = 224$ spectral bands, is generated using $J = 7$ endmembers randomly selected among the USGS spectral library [23] at each run of the algorithm, in order to obtain generalizable results. The abundances are obtained from Dirichlet J-variate random generations, which guarantees the sum-to-one and the positivity constraints. In order to select data with a fixed maximum abundance level, we keep only the vectors whose maximal value is lower than a chosen threshold a_{max}, which controls the mixing level. In order to test the influence of the mixing level, we make the parameter a_{max} vary from 0.6 (max purity of the data) to 1. Then we construct data from Fan model (1), and add white gaussian noise with a signal to noise ratio $SNR = 40dB$. During the experiments, we stop the algorithms after 1000 iterations, the execution time being around 70 seconds for both NMF algorithms on a basic computer.

The spectra and abundances are respectively initialized with VCA [16] and FCLS [24] for the three methods. The initialized spectra are also employed for the GBM-GDA method since it assumes the endmembers are known. The sum-to-one regularization parameter δ is empirically fixed in these experiments. We compare the convergence rates on 250 iterations, showing in Fig. 2 the objective function of each algorithm in function of the iteration number for one simulation, thus showing our proposed method is converging faster than MDMD-NMF. It is worth noting that GBM-GDA is converging faster than the proposed method. However, the error values remain constant and higher than the error of the proposed method. This reveals the GBM-GDA may be trapped in a local minimum.

a) Regularization parameters: Original MDMD-NMF depends on three regularization parameters, respectively on sum-to-one (or sum-to-unity, STU) parameter δ, on spectra dispersion, and on abundances dispersion [22]. For the purpose of brevity, we only consider here STU, setting the other two regularization parameters to 0. The specificity of this NMF is the optimization algorithm, which is a projected gradient with a modified Armijo/Lin rule [12]. We study the performance of Fan-NMF towards the sum-to-one constraint. Indeed, many authors contest the opportunity of such a constraint. There are two reasons for this: in a purely algorithmic point of view, it may disturb the convergence, whereas in a physical point of view it could be not relevant. We test here the first alternative, as the generated data obey to the sum-to-one constraint. It is possible to tune the sum-to-one constraint influence using the regularization parameter δ inherent to the two NMF-based algorithms.

The results, obtained after averaging over 10 Monte-Carlo runs, are given Fig. 3 for three values of δ, as a function of the maximum purity coefficient a_{max}. As comparison criterion, the root mean squared error (RMSE) has been computed for the abundances and spectra (where "reconstruction RMSE" is the RMSE based on the whole data), as well as their mean angle distance. The final RMSE and STU after 1000 iterations are also given. We notice that :

- final reconstruction error slightly increases with δ, since the objective function is $MSE + STU^2$,
- final value of STU depends on δ,
- abundances estimations are notably better when δ increases, especially for highly mixed cases,
- spectra estimations are better for intermediate values of δ.

As a consequence, we conclude that a soft STU constraint improve the results, especially for abundances estimation, but it must be controlled for spectra estimation. As a compromise, we choose $\delta = 0.6$ in the next simulations, knowing that if we were only interested in abundances estimation, a higher value should be taken.
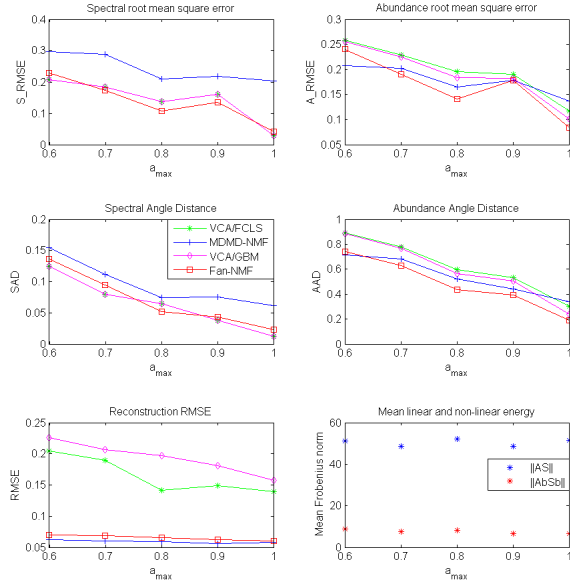
Fig. 4. Comparison between MDMD-NMF, VCA+FCLS, GBM-GDA and Fan-NMF, and norm of the linear (resp. non-linear) part of the data.

b) Comparison with other algorithms: In this section, we test:

- Two non-linear methods together
- The improvement brought by the integration of the non-linear model towards linear model in the algorithms

The results are presented in Fig. 4. The two NMF based algorithms obtain very close final reconstruction error, around 0.06, while the VCA/GBM and VCA/FCLS error is approximately 0.2, and the error slightly increases with the mixing level. It is clearly shown that the Fan-NMF gives generally the best results. The performance on spectra estimation is not very different from VCA, whereas abundances are better estimated than with GBM-GDA. If we compare the linear NMF versus non-linear one, we see that modeling the non-linearity is powerful, except for highly mixed data ($a_{max} = 0.6$).

The two linear and non-linear abundance estimators, FCLS and GBM-GDA respectively, give close results, a little better for GBM-GDA for reasonable mixing levels ($a_{max} > 0.6$). For very high level of mixing, the performances decrease for all algorithms. Furthermore, the non-linear models seem to be no more efficient. Indeed, for $a_{max} = 0.6$, there is no difference between FCLS and GBM-GDA, while MDMD-NMF outperforms Fan-NMF for abundances estimation. Note that for $a_{max} = 0.6$, the mean value of bilinear abundance term $mean(Ab) = 14 \times 10^{-3}$, while the noise standard deviation was $\sigma = 7 \times 10^{-3}$. In Fig.4, the linear and non linear Frobenius norm of the data are plotted, showing a nearly constant ratio for the various mixing levels.

B. Real data

The real image considered here (Fig. 5) has been acquired by Actimar company [2] with HySpex [3] sensor in 2010, during

the data campaign of the HYPLITT project [25], funded by the French Defence Agency (DGA). The image is composed of 81 spectral bands in the range of $400nm - 700nm$ because of the absorption by the water column in infra-red bands, and 150×150 pixels. The scene represents the Quiberon bay, and we expect some non linear mixing due to the complex structure and water column.

The same initialization methods as in the previous section have been employed, and we set $\delta = 0.4$ and $J = 8$. After a single simulation we compute the reconstruction error for the four algorithms previously tested on simulated images.

Fig. 5. Real hyperspectral data: coast of Brittany acquired by Actimar society (left) and the region of interest shown in true colors (right).

Table I gives the respective RMSE obtained with the four algorithms for this data ("spectral RMSE" is the RMSE based on the spectra). The two NMF based algorithms have the lowest error. To complete with another criterion, we assume that VCA has correctly identified the endmembers, and employ them as spectral references, in order to compare between the linear and the non-linear NMF. Best results are obtained with Fan-NMF, thus indicating that our proposed method could be well adapted for unmixing this environment.

TABLE I
RECONSTRUCTION ERROR OF THE FOUR DIFFERENT ALGORITHMS FOR HYPLITT DATA AND SPECTRAL RMSE.

	Fan-NMF	MDMD-NMF	VCA/GBM-GDA	VCA/FCLS
RMSE	0.02	0.02	0.24	0.24
S-RMSE	0.01	0.07	-	-

As we do not have the ground truth we only show in Fig. 6 the abundance maps obtained for Fan-NMF, and the endmembers estimated respectively with Fan-NMF and VCA. Many spectra have very low reflectance, due to the water column attenuation. The low-reflectance spectra are better discriminated with Fan-NMF than with VCA.

IV. CONCLUSION

A new unmixing NMF method has been presented in this contribution. It is based on a non-linear mixing model, given as a second order polynomial development of a non-linear function. The proposed method is bilinear for spectra and abundances, contrary to NMF based preliminary works. Using a projected gradient algorithm with adaptive step size, results
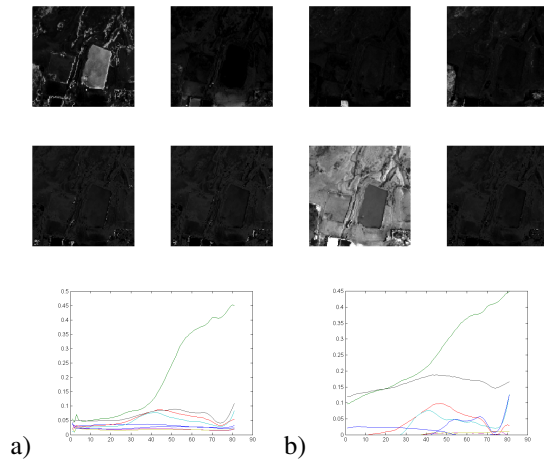
Fig. 6. Top: abundance maps obtained with fan-NMF, bottom: endmembers estimated with a) VCA, b) Fan-NMF.

on synthetic and real data show good convergence and good unmixing performances. A discussion on sum to one constraint shows that, whenever the data fulfils it, we take advantage in imposing it as a soft constraint. Comparison with an other unmixing algorithm, using similar model for the data, but estimating only the abundances, show the interest of the NMF approach for non-linear unmixing. On real data, the proposed algorithm seems to be able to obtain interesting results. However, in the case of highly mixed pixels, the approach becomes limited, and the difference between linear and nonlinear algorithms gets weak. Some more systematic experiments on real data should be carried out to explore the domain of efficiency of Fan-NMF.

V. Acknowledgements

References

[1] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, and P. Gader, "Hyperspectral unmixing overview: geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Applied Earth Observations and Remote Sensing*, vol. 5, no. 2, pp. 354–379, Apr. 2012.
[2] 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.
[3] B. Hapke, "Bidirectional reflectance spectroscopy 1. theory," *J. Geophys. Res.*, vol. 86, pp. 3039–3054, 1981.
[4] J. Broadwater and A. Banerjee, "Mapping intimate mixtures using an adaptive kernel-based technique," in *Proc. IEEE GRSS Workshop on Hyperspectral Image and SIgnal Processing: Evolution in Remote Sensing (WHISPERS)*, Lisbon, Portugal, June 2011.
[5] J. Plaza, A. Plaza, P. Martinez, and R. Perez, "Nonlinear mixture models for analyzing laboratory simulated-forest hyperspectral data," *Proc. SPIE Image and Signal Processing for Remote Sensing IX*, vol. 5238, pp. 480–487, 2004.
[6] P. Comon, C. Jutten, and J. Herault, "Blind separation of sources. part ii: problems statement," *Signal Processing*, vol. 24, no. 1, pp. 11–20, July 1991.
[7] S. Hosseini and Y. Deville, "Blind maximum likelihood separation of a linear-quadratic mixture," in *Proceedings of the Fifth International Workshop on Independent Component Analysis and Blind Signal Separation (ICA)*, 2004, pp. 694–701.
[8] L. T. Duarte, C. Jutten, and S. Moussaoui, "Bayesian source separation of linear and linear-quadratic mixtures using truncated priors," *Journal of Signal Processing Systems*, vol. 11, no. 3, pp. 311–323, 2011.
[9] P. Paatero and U. Tapper, "Positive matrix factorization: a non-negative factor model with optimal utilization of error estimates of data values," in *Fourth International Conference on Statistical methods for the Environmental Sciences (Environmetrics)*, vol. 5, Aug. 1994, pp. 111–126.
[10] P. Sajda, S. Du, T. R. Brown, R. Stoyanova, D. C. Shungu, M. Xiangling, and L. Parra, "Nonnegative matrix factorization for rapid recovery of constituent spectra in magnetic resonance chemical shift imaging of the brain," *IEEE Transactions on Medical Imaging*, vol. 23, no. 12, pp. 1453–1465, Dec. 2004.
[11] L. Miao and H. Qi, "Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization," *IEEE Trans. Geosci. and Remote Sensing*, vol. 45, no. 3, pp. 765–776, March 2007.
[12] A. Huck, M. Guillaume, and J. Blanc-Talon, "Minimum dispersion constrained nonnegative matrix factorization to unmix hyperspectral data," *IEEE Trans. Geoscience and Remote Sensing*, vol. 48, no. 6, pp. 2590–2600, June 2010.
[13] I. Meganem, Y. Deville, S. Hosseini, P. Déliot, X. Briottet, and L. T. Duarte, "Linear-quadratic and polynomial non-negative matrix factorization; application to spectral unmixing," in *19th European Signal Processing Conference (EUSIPCO 2011)*, Barcelona, Spain, Sept. 2011.
[14] A. Halimi, Y. Altmann, N. Dobigeon, and J.-Y. Tourneret, "Nonlinear unmixing of hyperspectral images using a generalized bilinear model," *IEEE Trans. Geoscience and Remote Sensing*, vol. 49, no. 11, pp. 4153–4162, Nov. 2011.
[15] ——, "Unmixing hyperspectral images using the generalized bilinear mixing model," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Vancouver, Canada, July 2011, pp. 1886–1889.
[16] J. M. Nascimento and J. M. Bioucas-Dias, "Vertex component analysis: a fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. and Remote Sensing*, vol. 43, no. 4, pp. 898–910, April 2005.
[17] A. Cichocki, R. Zdunek, A. H. Phan, and S.-I. Amari, *Nonnegative Matrix and Tensor Factorizations*. Chichester, UK: Wiley, 2009.
[18] D. D. Lee and H. S. Seung, "Learning the parts of objects by nonnegative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, Oct. 1999.
[19] ——, "Algorithms for non-negative matrix factorization," in *Advances in Neural Information Processing Systems*, vol. 13. Cambridge: MIT Press, 2001, pp. 556–562.
[20] C.-J. Lin, "Projected gradient methods for non-negative matrix factorization," *Neural Computation*, vol. 19, no. 10, pp. 2756–2779, 2007.
[21] P. O. Hoyer, "Non-negative matrix factorization with sparseness constraints," *J. Mach. Learn. Res.*, vol. 5, no. 37, pp. 1457–1469, 2004.
[22] A. Huck and M. Guillaume, "Robust hyperspectral data unmixing with spatial and spectral regularized NMF," in *Proc. IEEE GRSS Workshop on Hyperspectral Image and SIgnal Processing: Evolution in Remote Sensing (WHISPERS)*, Reykjavik, Iceland, June 2010.
[23] Http://speclab.cr.usgs.gov/spectral-lib.html.
[24] 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. 39, no. 3, pp. 529–545, March 2001.
[25] S. Smet, G. Sicot, and M. Lennon, "Evaluation des capacités de la télédétection HYPerspectrale et développement de méthodes innovantes de traitement d'images pour des applications défense en zone LITTorales (HYPLITT)." DGA-RF-280310, Tech. Rep., 2010.