Fast Hyperspectral Unmixing in Presence of Sparse Multiple Scattering Nonlinearities

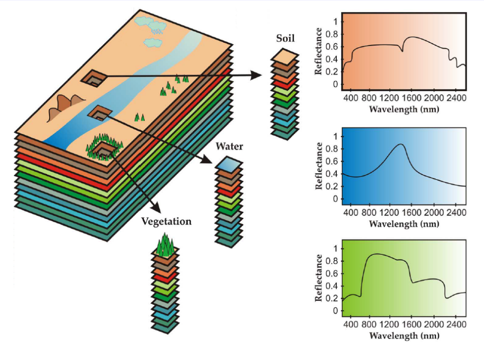
Abderrahim Halimi [1], Jose Bioucas-Dias [2], Nicolas Dobigeon [3], Gerald S. Buller [1], Stephen McLaughlin [1]

(1) Heriot-Watt University, Sch. of Eng. and Physical Sciences, Edinburgh, U.K.
(2) Instituto de Telecomunicacões and Instituto Superior Técnico, Portugal.
(3) University of Toulouse, IRIT/INP-ENSEEIHT, Toulouse, France.

ICASSP 2017, New Orleans, USA

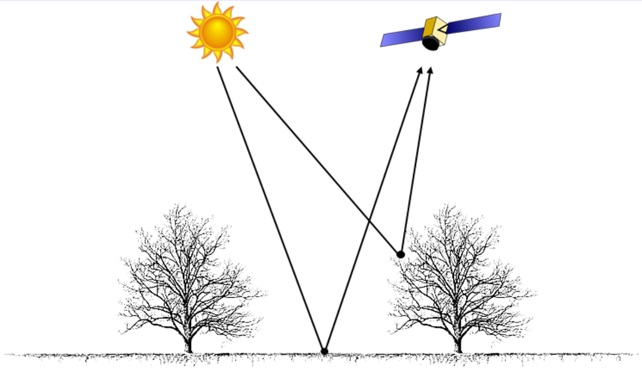
Hyperspectral images

- Each pixel is represented in several spectral bands
- The pixel spectrum is related to its physical elements



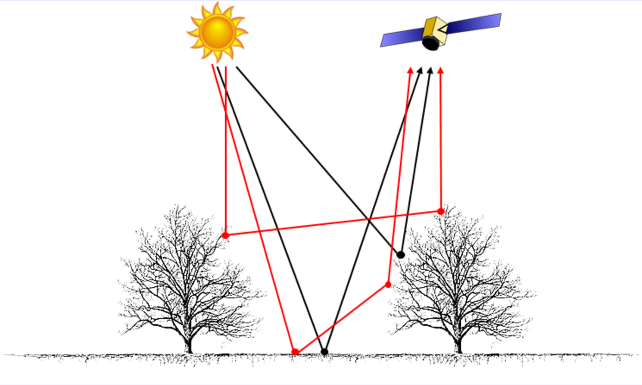
- N pixels
- L spectral bands
- R endmembers

Mixture models



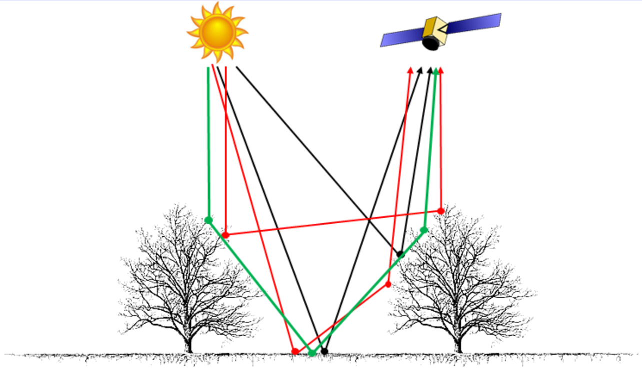
Linear model

Mixture models



Bilinear models

Mixture models



Multilinear models

Problem statement

- Supervised unmixing: estimate the abundances while assuming known endmembers
- Take into account the effect of multiple scattering interactions
- Take into account the known properties/constraints of the abundances and the other parameters of interest
- Propose an algorithm with a reduced computational complexity

Summary

Summary

Linear mixture model (LMM)

$$\boldsymbol{y}_n = \boldsymbol{M}\boldsymbol{a}_n + \boldsymbol{e}_n$$

- $\boldsymbol{y}_n$ is the nth pixel spectrum of size $(L \times 1)$
- $\boldsymbol{M} = [\boldsymbol{m}_1, ..., \boldsymbol{m}_R]$ is the $(L \times R)$ endmember matrix
- $\boldsymbol{e}_n \sim \mathcal{N}\left(\boldsymbol{0}_L, \sigma^2 \mathbb{I}\right)$ is an $(L \times 1)$ white Gaussian noise
- $\boldsymbol{a}_n = [a_{1,n}, \ldots, a_{R,n}]^T$ is the abundance vector satisfying:

  $a_{r,n} \geq 0, r = 1, \cdots, R$ and $\sum_{r=1}^{R} a_{r,n} = 1$

Bilinear models

$$\boldsymbol{y}_n = \boldsymbol{M}\boldsymbol{a}_n + \sum_{r,r'} x_n^{r,r'} \boldsymbol{m}_r \odot \boldsymbol{m}_{r'} + \boldsymbol{e}_n \text{ with } x_n^{r,r'} \geq 0.$$

- $\odot$ denotes the Hadamard (term-wise) product.

Different constraints on $x_n^{r,r'}$

- Nascimento model[1], FM[2], GBM[3], PPNMM[4], RCA-MCMC[5]

---

[1] J. M. Bioucas-Dias and al., "Nonlinear mixture model for hyperspectral unmixing", *Proc. SPIE ISPRS XV*, vol. 7477, no. 1, 2009, p. 74770I.

[2] W. Fan, and al., "Comparative study between a new nonlinear model and common linear model for analysing laboratory simulated-forest hyperspectral data," *Int. Journal of Remote Sens.*, vol. 30, no. 11, pp. 2951-2962, June 2009.

[3] A. Halimi and al., "Nonlinear unmixing of hyperspectral images using a generalized bilinear model," *IEEE TGRS*, vol. 49, no. 11, pp. 4153-4162, 2011.

[4] Y. Altmann, and al., "Supervised nonlinear spectral unmixing using a postnonlinear mixing model for hyperspectral imagery," *IEEE TIP*, vol. 21, no. 6, pp. 3017-3025, 2012.

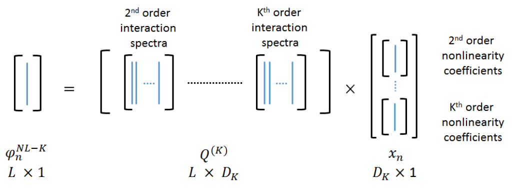
[5] Y. Altmann, and al., "Bayesian nonlinear hyperspectral unmixing with spatial residual component analysis," *IEEE TCI*, vol. 1, no. 3, pp. 174-185, 2015.

Proposed residual component model: RC-NL-K

$$\boldsymbol{y}_n = \boldsymbol{M}\boldsymbol{a}_n + \phi_n^{\text{NL-}K}(\boldsymbol{M}, \boldsymbol{x}_n) + \boldsymbol{e}_n$$

with $\phi_n^{\text{NL-}K}(\boldsymbol{M}, \boldsymbol{x}_n) = \boldsymbol{Q}^{(K)}(\boldsymbol{M})\,\boldsymbol{x}_n$, and $K \geq 2$

- $\boldsymbol{Q}^{(K)}(\boldsymbol{M})$ is the $(L \times D_K)$ matrix gathering the interaction spectra of the form $\boldsymbol{m}_i \odot \boldsymbol{m}_j \odot \cdots \odot \boldsymbol{m}_s$,
- $\boldsymbol{x}_n \geq 0$ is the nth vector of non-negative nonlinearity coefficients of size $(D_K \times 1)$,



Proposed residual component model: RC-NL-K

$$\boldsymbol{y}_n = \boldsymbol{M}\boldsymbol{a}_n + \boldsymbol{\phi}_n^{\text{NL-}K}(\boldsymbol{M}, \boldsymbol{x}_n) + \boldsymbol{e}_n$$

with $\boldsymbol{\phi}_n^{\text{NL-}K}(\boldsymbol{M}, \boldsymbol{x}_n) = \boldsymbol{Q}^{(K)}(\boldsymbol{M})\,\boldsymbol{x}_n$, and $K \geq 2$

Special case: linear model

- Reduces to LMM for $\boldsymbol{\phi}_n^{\text{NL-}K}(\boldsymbol{M}, \boldsymbol{x}_n) = 0$, i.e., $\boldsymbol{x}_n^{(d)} = 0, \forall n, d$

Proposed residual component model: RC-NL-K

$$\boldsymbol{y}_n = \boldsymbol{M}\boldsymbol{a}_n + \boldsymbol{\phi}_n^{\text{NL-}K}(\boldsymbol{M}, \boldsymbol{x}_n) + \boldsymbol{e}_n$$

with $\boldsymbol{\phi}_n^{\text{NL-}K}(\boldsymbol{M}, \boldsymbol{x}_n) = \boldsymbol{Q}^{(K)}(\boldsymbol{M})\,\boldsymbol{x}_n$, and $K \geq 2$

Special case: bilinear models

- $K = 2,\ D_2 = \frac{R(R+1)}{2}$,
- $\boldsymbol{Q}^{(2)}(\boldsymbol{M}) = \left(\sqrt{2}\boldsymbol{m}_{1,2}, \cdots, \sqrt{2}\boldsymbol{m}_{R-1,R}, \boldsymbol{m}_{1,1}, \cdots, \boldsymbol{m}_{R,R}\right)$,
- $\boldsymbol{x}_n = \left(x_n^{(1,2)}, \cdots, x_n^{(R-1,R)}, x_n^{(1,1)}, \cdots, x_n^{(R,R)}\right)^T, \forall n$,
- $\boldsymbol{m}_{r,r'} = \boldsymbol{m}_r \odot \boldsymbol{m}_{r'}$.
- Relation to bilinear models: RCA-MCMC, NM, FM, GBM, PPNMM

Proposed residual component model: RC-NL-K

$$\boldsymbol{y}_n = \boldsymbol{M}\boldsymbol{a}_n + \boldsymbol{\phi}_n^{\text{NL-}K}\left(\boldsymbol{M}, \boldsymbol{x}_n\right) + \boldsymbol{e}_n$$

with $\boldsymbol{\phi}_n^{\text{NL-}K}\left(\boldsymbol{M}, \boldsymbol{x}_n\right) = \boldsymbol{Q}^{(K)}(\boldsymbol{M})\,\boldsymbol{x}_n$, and $K \geq 2$

Special case: multilinear model

- $K > 2$,
- $\boldsymbol{Q}^{(K)} = \left[\boldsymbol{Q}_2^{(K)}, \boldsymbol{Q}_3^{(K)}, \cdots, \boldsymbol{Q}_K^{(K)}\right]$,
- $\boldsymbol{x}_n$ is the nth vector of nonlinearity coefficients of size $(D_K \times 1)$,
- $D_K = \sum_{i=2}^{K} \frac{(R+i-1)!}{i!(R-1)!}$, where $z!$ denotes the factorial of z

Summary

General formulation

$$\mathcal{C}(\boldsymbol{Z}) = \mathcal{L}_{\boldsymbol{Q}}(\boldsymbol{Z}) + \psi(\boldsymbol{Z})$$

- $\mathcal{L}_{\boldsymbol{Q}}(\boldsymbol{Z}) = \frac{1}{2}||\boldsymbol{Y} - [\boldsymbol{M}, \boldsymbol{Q}]\boldsymbol{Z}||_F^2$ due to the Gaussian noise properties
- $\boldsymbol{Z} = \left[\boldsymbol{A}^\top, \boldsymbol{X}^\top\right]^\top$ the parameters of interest
- $\psi(\boldsymbol{Z})$: regularization term to account for the known properties/constraints on $\boldsymbol{Z}$
- $\boldsymbol{A}$ matrix of abundances of size $(R \times N)$
- $\boldsymbol{X}$ matrix of coefficients of size $(D \times N)$

Prior knowledge/hypotheses on $\boldsymbol{Z}$

Abundances: $\boldsymbol{A}$

- Non-negativity and sum-to-one constraints

Nonlinearity coefficients: $\boldsymbol{X}$

- Non-negativity of the coefficients (a widely used assumption)
- The nonlinearity appears in some pixels of the image (as in [6], [7])
- In a nonlinear pixel, only a few interactions are active (implicitly assumed by bilinear models).

[6] C. Fevotte, and al., "Nonlinear hyperspectral unmixing with robust nonnegative matrix factorization," *IEEE TIP*, vol. 24, no. 12, 2015.

[7] Y. Altmann, and al., "Residual component analysis of hyperspectral images: Application to joint nonlinear unmixing and nonlinearity detection," *IEEE TIP*, vol. 23, no. 5, 2014.

Cost function

$$\mathcal{C}(\boldsymbol{Z}) = \mathcal{L}_{\boldsymbol{Q}}(\boldsymbol{Z}) + i_{\mathbb{R}_+}(\boldsymbol{A}) + i_{\{\mathbf{1}_{(1,R)}\}}\left(\mathbf{1}_{(1,R)}\boldsymbol{A}\right) + \tau_1 ||\boldsymbol{X}||_1 + \tau_2 ||\boldsymbol{X}||_{2,1} + i_{\mathbb{R}_+}(\boldsymbol{X})$$

- $\tau_1 > 0, \tau_2 > 0$ are fixed regularization parameters

Cost function

$$\mathcal{C}(\boldsymbol{Z}) = \mathcal{L}_{\boldsymbol{Q}}(\boldsymbol{Z}) + i_{\mathbb{R}_+}(\boldsymbol{A}) + i_{\{\mathbf{1}_{(1,R)}\}}(\mathbf{1}_{(1,R)}\boldsymbol{A}) + \tau_1||\boldsymbol{X}||_1 + \tau_2||\boldsymbol{X}||_{2,1} + i_{\mathbb{R}_+}(\boldsymbol{X})$$

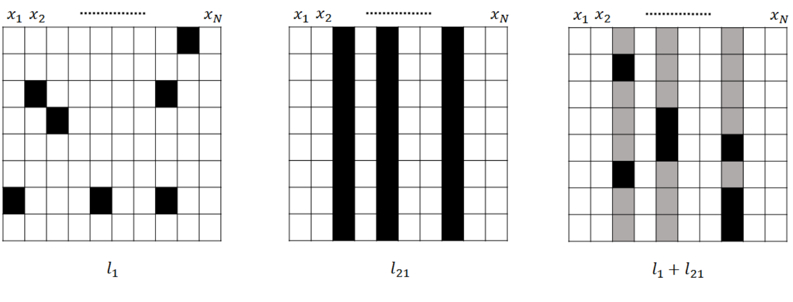
Abundance contraints

- $i_{\mathbb{R}_+}(\boldsymbol{A})$ Abundances non-negativity constraint
- $i_{\{\mathbf{1}_{(1,R)}\}}(\mathbf{1}_{(1,R)}\boldsymbol{A})$ Abundances sum-to-one constraint

Cost function

$$\mathcal{C}\left(\boldsymbol{Z}\right) = \mathcal{L}_{\boldsymbol{Q}}\left(\boldsymbol{Z}\right) + i_{\mathbb{R}_+}\left(\boldsymbol{A}\right) + i_{\left\{\mathbf{1}_{(1,R)}\right\}}\left(\mathbf{1}_{(1,R)}\boldsymbol{A}\right)$$
$$+ \tau_1 ||\boldsymbol{X}||_1 + \tau_2 ||\boldsymbol{X}||_{2,1} + i_{\mathbb{R}_+}\left(\boldsymbol{X}\right)$$

Nonlinearity coefficients



Summary

Description of the NUSAL-K algorithm (1)

Nonlinear Unmixing by variable Splitting and Augmented Lagrangian (with order K)

$$\underset{\boldsymbol{Z}}{\operatorname{argmin}}\, \mathcal{C}(\boldsymbol{Z}) = \underset{\boldsymbol{Z}}{\operatorname{argmin}}\, \mathcal{L}_{\boldsymbol{Q}}(\boldsymbol{Z}) + i_{\mathbb{R}_+}(\boldsymbol{A}) + i_{\{\mathbf{1}_{(1,R)}\}}\left(\mathbf{1}_{(1,R)}\boldsymbol{A}\right) + \tau_1||\boldsymbol{X}||_1 + \tau_2||\boldsymbol{X}||_{2,1} + i_{\mathbb{R}_+}(\boldsymbol{X})$$

with $\boldsymbol{Z} = \left[\boldsymbol{A}^\top, \boldsymbol{X}^\top\right]^\top$.

Equivalent formulation

$$\underset{\boldsymbol{z}}{\operatorname{argmin}}\, \mathcal{C}(\boldsymbol{Z}) = \underset{\boldsymbol{Z}}{\operatorname{argmin}} \sum_{j=1}^{J} g_j\left(\boldsymbol{H}^{(j)}\boldsymbol{Z}\right)$$

- $g_j : \mathbb{R}^{p_j \times N} \to \mathbb{R}$ are proper and convex functions
- $\boldsymbol{H}^{(j)} \in \mathbb{R}^{p_j \times (R+D)}$ are selection matrices
- $\boldsymbol{U}_j = \boldsymbol{H}_j \boldsymbol{Z}, \in \mathbb{R}^{p_j \times N}$

Description of the NUSAL-K algorithm (2)

Nonlinear Unmixing by variable Splitting and Augmented Lagrangian (with order K)

$$g_1(\boldsymbol{U}_1) = \mathcal{L}_{\boldsymbol{Q}}(\boldsymbol{U}_1), \qquad \boldsymbol{H}_1 = \mathbb{I}_{(R+D_K)}$$

$$g_2(\boldsymbol{U}_2) = i_{\mathbb{R}_+}(\boldsymbol{U}_2), \qquad \boldsymbol{H}_2 = \mathbb{I}_{(R+D_K)}$$

$$g_3(\boldsymbol{U}_3) = i_{\{\boldsymbol{1}^\top\}}\left(\boldsymbol{1}^\top \boldsymbol{U}_3\right), \qquad \boldsymbol{H}_3 = \left[\mathbb{I}_R, \boldsymbol{0}_{(R,D_K)}\right]$$

$$g_4(\boldsymbol{U}_4) = \tau_1 ||\boldsymbol{U}_4||_1, \qquad \boldsymbol{H}_4 = \left[\boldsymbol{0}_{(D_K,R)}, \mathbb{I}_{D_K}\right]$$

$$g_5(\boldsymbol{U}_5) = \tau_2 ||\boldsymbol{U}_5||_{2,1}, \qquad \boldsymbol{H}_5 = \left[\boldsymbol{0}_{(D_K,R)}, \mathbb{I}_{D_K}\right]$$

Description of the NUSAL-K algorithm (3)

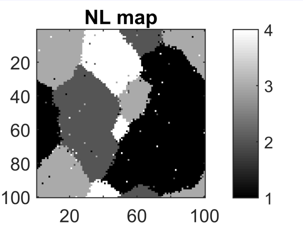
Initialize $\boldsymbol{U}_j^{(0)}, \boldsymbol{F}_j^{(0)}, \forall j$, $\mu > 0$. Set $k \leftarrow 0$, conv$\leftarrow 0$
while conv$= 0$ **do**
  for j=1:J **do**
    $\xi_j^{(k)} \leftarrow \boldsymbol{U}_j^{(k)} + \boldsymbol{F}_j^{(k)}$,
  end for
  Linear system of equations
  $\boldsymbol{Z}^{(k+1)} \leftarrow \left[\sum_{j=1}^{J} (\boldsymbol{H}_j)^\top \boldsymbol{H}_j\right]^{-1} \sum_{j=1}^{J} (\boldsymbol{H}_j)^\top \xi_j^{(k)}$,
  Moreau proximity operators
  for j=1:J **do**
    $\boldsymbol{V}_j^{(k)} \leftarrow \boldsymbol{H}_j \boldsymbol{Z}^{(k+1)} - \boldsymbol{F}_j^{(k)}$,
    $\boldsymbol{U}_j^{(k+1)} \leftarrow \underset{\boldsymbol{U}_j}{\text{argmin}} \frac{\mu}{2} ||\boldsymbol{U}_j - \boldsymbol{V}_j^{(k)}||^2 + g_j(\boldsymbol{U}_j)$,
  end for
  Update Lagrange multipliers
  for j=1:J **do**
    $\boldsymbol{F}_j^{(k+1)} \leftarrow \boldsymbol{U}_j^{(k+1)} - \boldsymbol{V}_j^{(k)}$,
  end for
  $k = k + 1$
end while

Summary

Synthetic image (1)

Considered image

- A synthetic image ($N = 100 \times 100$ pixels, $R = 3$ endmembers, $L = 207$ bands)
- $K = 4$ spatial classes (obtained using a Potts-MRF) whose pixels are generated according to LMM, RCA-NL3, GBM and PPNMM
- Abundance uniformly generated in the simplex of positivity and sum-to-one constraints.



Synthetic image (2)

Comparison algorithms

- SUNSAL[8], SKhype[9], CDA-NL[10], RNMF[11] and the proposed NUSAL-2 and NUSAL-3.

Evaluation criteria

$$\mathrm{RMSE}\,(\boldsymbol{A}) = \sqrt{\frac{1}{N\,R}\sum_{n=1}^{N}\|\boldsymbol{a}_n - \hat{\boldsymbol{a}}_n\|^2}$$

$$\mathrm{SAM} = \frac{1}{N}\sum_{n=1}^{N}\arccos\left(\frac{\hat{\boldsymbol{y}}_n^T\boldsymbol{y}_n}{\|\boldsymbol{y}_n\|\ \|\hat{\boldsymbol{y}}_n\|}\right)$$

---

[8] J. Bioucas-Dias and al., "Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing,", *WHISPERS*, 2010.
[9] J. Chen, and al., "Nonlinear unmixing of hyperspectral data based on a linear-mixture/nonlinear fluctuation model," *IEEE TIP*, vol. 61, no. 2, 2013.
[10] A. Halimi, and al., "Hyperspectral unmixing in presence of endmember variability, nonlinearity or mismodelling effects," *IEEE TIP*, vol. 25, no. 10, 2016.
[11] C. Fevotte, and al., "Nonlinear hyperspectral unmixing with robust nonnegative matrix factorization," *IEEE TIP*, vol. 24, no. 12, 2015.

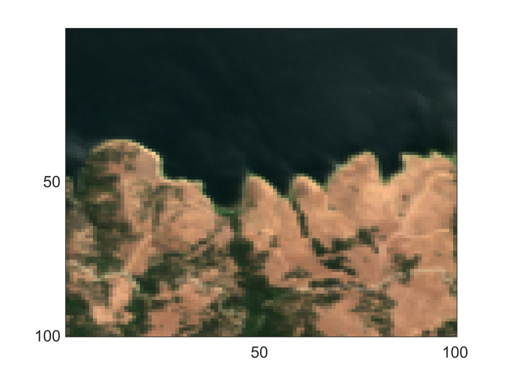
Synthetic image (3): performance

	RMSE				RMSE	SAM	Time (s)
	$\mathcal{C}_1$ LMM	$\mathcal{C}_2$ NL-3	$\mathcal{C}_3$ GBM	$\mathcal{C}_4$ PPNMM			
SUNSAL	1.4	20.3	5.8	11.9	10.8	7.6	0.1
SKhype	2.2	11.7	3.0	3.9	6.0	–	466
CDA-NL	1.4	4.5	2.1	4.2	2.9	5.8	182
RNMF	1.5	12.8	2.5	5.2	6.4	6.8	110
NUSAL-2	1.4	3.9	2.0	5.0	2.8	5.8	7
NUSAL-3	1.4	2.9	2.0	4.9	2.6	5.7	19

Results on synthetic data.

Green: best, Red: second best.

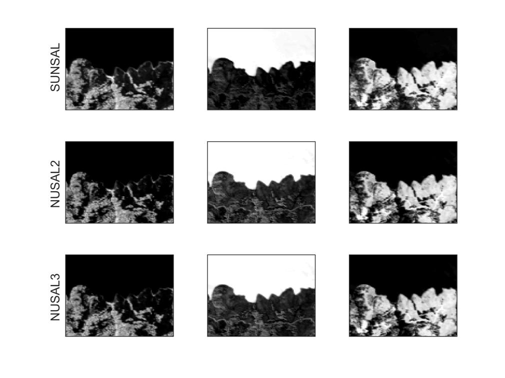
Real image (1): Moffett image



Considered subimage

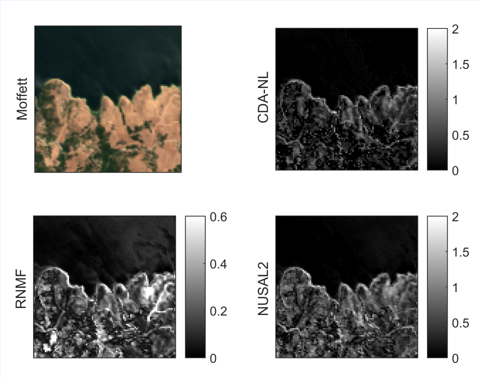
100×100 pixels, $L = 152$ spectral bands, $R = 3$ endmembers.

Real image (2): abundance estimation



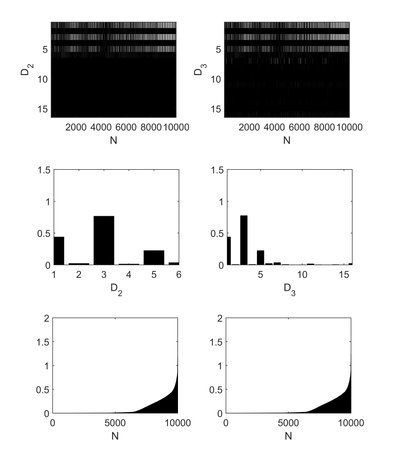
SKhype (177 s), CDA-NL (317 s), RNMF (278 s), NUSAL-2 (13 s), and NUSAL-3 (29 s)

Real image (3): Residuals



Square root of the energies of the difference between the reconstructed signal and the linear model obtained with $||\hat{\boldsymbol{y}}_{i,j} - \boldsymbol{M}\hat{\boldsymbol{a}}_{i,j}||$

Real image (4): nonlinearity coefficients



The nonlinearity coefficients are active for some pixels +
Most interactions are captured by the bilinear terms

Summary

Conclusions & future work

Conclusions

- Generalization of the existing bilinear models by accounting for multiple interactions
- Introduction of a fast estimation algorithm called NUSAL-K
- Good performance for synthetic and real images

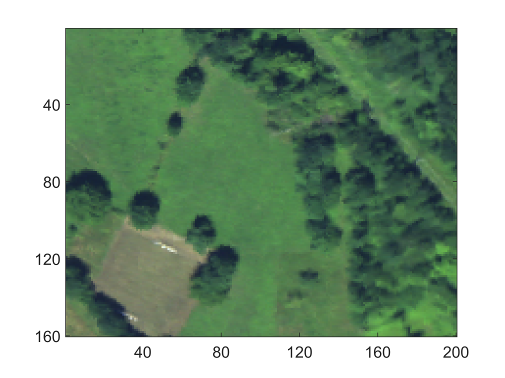
Future work

- Generalizing the model to include other prior information regarding the estimated coefficients
- Estimation of the hyperparameters

End

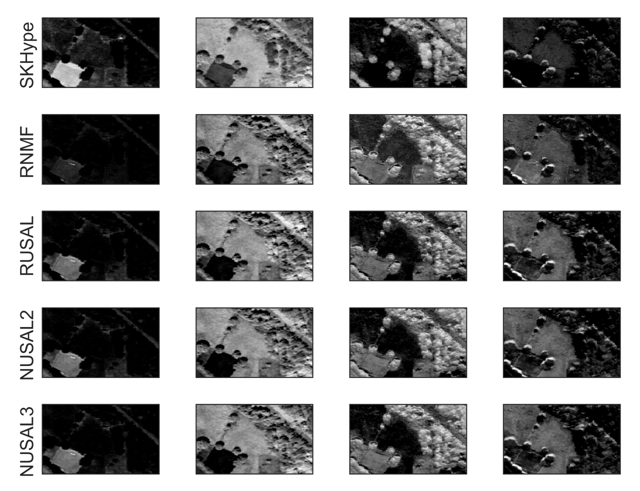
Thank you for your attention

Real image (1): Madonna image

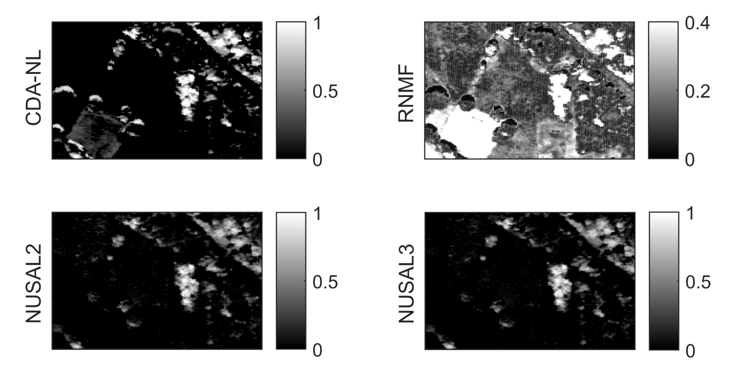


Considered subimage
160×200 pixels, $L = 160$ spectral bands, $R = 4$ endmembers.

Real image (2): abundance estimation

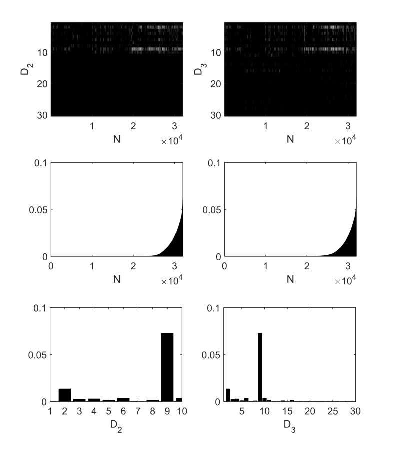


Real image (3): Residuals



Square root of the energies of the difference between the reconstructed signal and the linear model obtained with $||\hat{\boldsymbol{y}}_{i,j} - \boldsymbol{M}\hat{\boldsymbol{a}}_{i,j}||$

Real image (4): nonlinearity coefficients



The nonlinearity coefficients are active for some pixels +
Most interactions are captured by the bilinear terms