Successive Nonnegative Projection Algorithm for Linear Quadratic Mixtures (EUSIPCO 2020)

Christophe Kervazo, Nicolas Gillis, Nicolas Dobigeon

Successive Nonnegative Projection Algorithm for Linear Quadratic Mixtures

Christophe Kervazo
University of Mons
Mons, Belgium
christophe.kervazo@umons.ac.be

Nicolas Gillis
University of Mons
Mons, Belgium
nicolas.gillis@umons.ac.be

Nicolas Dobigeon
Univ. of Toulouse, IRIT/INP-ENSEEIHT
Toulouse, France
nicolas.dobigeon@enseeiht.fr

***Abstract*—In this work, we tackle the problem of hyperspectral unmixing by departing from the usual linear model and focusing on a linear-quadratic (LQ) one. The algorithm we propose, coined Successive Nonnegative Projection Algorithm for Linear Quadratic mixtures (SNPALQ), extends the Successive Nonnegative Projection Algorithm (SNPA), specifically designed to address the unmixing problem under a linear non-negative model and the pure-pixel assumption (a.k.a. near-separable assumption). By explicitly modeling the product terms inherent to the LQ model along the iterations of the SNPA scheme, the nonlinear contributions of the mixing are mitigated, thus improving the separation quality. The approach is shown to be relevant in realistic numerical experiments, which further highlight that SNPALQ is robust to noise.**

***Index Terms*—Nonnegative Matrix Factorization, Non-linear Hyperspectral Unmixing, Linear-Quadratic Models, Separability and Pure-Pixel Assumption, Non-linear Blind Source Separation.**

I. Introduction

Hyperspectral (HS) imaging is a powerful tool in a wide range of fields such as remote sensing [1], biomedical and pharmaceutical imaging [2], and astronomy [3], to only name a few. While such data sets are composed of a high number of spectral bands, HS images generally suffer from a limited spatial resolution. Therefore, several materials generally contribute to each pixel, and thus the acquired spectra correspond to mixtures of the spectra of the different pure materials in the pixel, called endmembers. Many works on HS imaging [4] have focused on the *linear* mixing model (LMM) which states that the spectral signature of the ith observed pixel $\mathbf{x}_i \in \mathbb{R}^m$ for $i \in [\![1, n]\!]$ can be written as

$$\mathbf{x}_i = \sum_{k=1}^{r} h_{ik}\mathbf{w}_k + \mathbf{n}_i,$$

where $\mathbf{w}_k$ ($k \in [\![1, r]\!]$) corresponds to the spectral signature of the kth endmember, h_{ik} to the spatial contribution (abundance) of the kth endmember in the ith pixel, and $\mathbf{n}_i$ accounts for any additive noise in the ith pixel. In a matrix form, the LMM can be written as $\mathbf{X} = \mathbf{WH} + \mathbf{N}$, with $\mathbf{X} \in \mathbb{R}^{m\times n}$, $\mathbf{W} \in \mathbb{R}^{m\times r}$, $\mathbf{H} \in \mathbb{R}^{r\times n}$ and $\mathbf{N} \in \mathbb{R}^{m\times n}$. Recovering $\mathbf{W}$ and $\mathbf{H}$ from the sole knowledge of $\mathbf{X}$ is referred to as *spectral unmixing* in the HS literature and can be cast as a blind source separation (BSS) problem [5]–[7]. As the problem is generally ill-posed, additional physical non-negativity constraints are imposed on the unknown matrices $\mathbf{W}$ and $\mathbf{H}$, akin to nonnegative matrix factorization (NMF) [8]. Although NMF is NP[1]-hard in general [9], the authors of [10] have introduced the subclass of near-separable non-negative matrices for which it can be solved in a polynomial time. This subclass corresponds to the pure-pixel assumption in HS imaging [11], [12]: for each endmember there exists a pixel in which only this endmember appears. Building on near-separable NMF, several provably robust algorithms have been proposed [10], [13], [14]. Among them, one can cite the Successive Projection Algorithm (SPA) [15], which is a fast greedy algorithm provably robust to noise [11], or an enhanced version, the Successive Nonnegative Projection Algorithm (SNPA) [16], which is particularly more efficient when $\mathbf{W}$ is either rank-deficient or ill-conditioned.

In various applicative contexts, LMM may however suffer from some limitations since it only consists in a first-order approximation. In particular, when the light arriving on the sensor interacts with several materials, nonlinear mixing effects may occur [4], [17], [18]. Specifically, this is often the case when the scene is not flat, for instance in the presence of large geometric structures, such as in urban [19] or forest [20] scenes. To take into account multiple scatterings, bilinear or linear-quadratic (LQ) models[2] include termwise products of the endmembers [20], [23], i.e.,

$$\mathbf{x}_i = \sum_{k=1}^{r} h_{ik}\mathbf{w}_k + \sum_{p=1}^{r}\sum_{l=p+1}^{r} \beta_{ipl}(\mathbf{w}_p \odot \mathbf{w}_l) + \mathbf{n}_i, \quad (1)$$

where $\odot$ denotes the Hadamard product and β_{ipl} adjusts the contribution of the quadratic term $\mathbf{w}_p \odot \mathbf{w}_l$ in the ith pixel. Despite source identifiability issues in the general context

CK and NG acknowledge the support by the European Research Council (ERC starting grant no 679515), and NG by the Fonds de la Recherche Scientifique - FNRS and the Fonds Wetenschappelijk Onderzoek - Vlanderen (FWO) under EOS Project no O005318F-RG47. ND is partly supported by the AI Interdisciplinary Institute ANITI funded by the French "Investing for the Future – PIA3" program under the Grant agreement n°ANR-19-PI3A-0004.

[1]Nondeterministic Polynomial time.

[2]While it is possible to include higher-order terms, most of the works neglect the interactions of order larger than two since they are expected to be of significantly lower magnitudes [21], [22].

of non-linear BSS problem [5], [24], [25], it was recently showed[3] that the non-linearity inherent to bilinear mixtures leads to an *essentially unique* solution in the noiseless case, provided that products of the sources up to order four are linearly independent [26]. Such an assumption thus requires the family

$$\big(\mathbf{w}_i, \mathbf{w}_i \odot \mathbf{w}_j, \mathbf{w}_i \odot \mathbf{w}_j \odot \mathbf{w}_k, \mathbf{w}_i \odot \mathbf{w}_j \odot \mathbf{w}_k \odot \mathbf{w}_l\big)_{\substack{i,j,k,l \in [\![1,r]\!], \\ l<k<j<i}} \tag{2}$$

to be linearly independent. As the size of this familly is $\frac{r(r+1)}{24}\big((r-1)(r-2)+12\big)$, this requirement might not be fulfilled in real-world scenarios since the number of spectral bands m should then also increase at least as $\mathcal{O}(r^4)$. For example, for $r = 10$, which is relatively small, we need $m \geq 385$ which is typically not satisfied for HS images. To overcome this issue, we propose to tackle problems of the form (1) under an NMF paradigm. The rationale is to convert the linear independence condition on the family (2) into a non-negative independence condition, which is significantly less restrictive in general.

In this work, we specifically focus on the so-called Nascimento model. Beyond its good capability to model some HS mixtures [20], it is an extension of the LMM, as the quadratic terms $\mathbf{w}_p \odot \mathbf{w}_l$ can be considered as additional *virtual* endmembers. Writing $\Pi_{\odot}(\mathbf{W}) = [\mathbf{w}_i, \mathbf{w}_i \odot \mathbf{w}_j]_{i,j \in [\![1,r]\!],\ j<i} \in \mathbb{R}^{m \times \tilde{r}}$ the matrix containing the endmembers and their second-order products with $\tilde{r} = r(r+1)/2$ (*cf.* notations at the end of this section), the model becomes

$$\mathbf{X} = \Pi_{\odot}(\mathbf{W})\tilde{\mathbf{H}} + \mathbf{N} \tag{3}$$

with $\tilde{\mathbf{H}} \in \mathbb{R}^{\tilde{r} \times n}$ the matrix of mixing coefficients associated with the linear h_{ik} and nonlinear β_{ipl} contributions in (1). This model is accompanied by the following constraints

$$\begin{aligned} &\forall i \in [\![1,n]\!], \forall k \in [\![1,\tilde{r}]\!], \tilde{h}_{ki} \geq 0, \\ &\forall i \in [\![1,n]\!], \sum_{k=1}^{\tilde{r}} \tilde{h}_{ki} \leq 1, \\ &\alpha = \min_{\substack{j \in [\![1,r]\!] \\ x \in \Delta}} \|\mathbf{w}_j - \Pi_{\odot}(\mathbf{W})_{\mathcal{J}} x\|_2 > 0,\ \mathcal{J} = [\![1,\tilde{r}]\!] \setminus \{j\}. \end{aligned} \tag{4}$$

The last condition ensures that no endmember lies within the convex hull formed by the other ones, the virtual ones and the origin. Lastly, extending the subclass of r near-separable mixing of [27] to the LQ model, we will assume the mixing to be r-LQ near-separable, as stated below.

Assumption I.1. *The matrix $\mathbf{X}$ is r-LQ near-separable if it can be written as:*

$$\mathbf{X} = \Pi_{\odot}(\mathbf{W}) \underbrace{\begin{bmatrix} \mathbf{I_r} & \tilde{\mathbf{H}}' \\ \mathbf{0}_{\frac{r(r-1)}{2} \times r} & \end{bmatrix}}_{\tilde{\mathbf{H}}} \mathbf{P} + \mathbf{N},$$

where $\mathbf{W} \in \mathbb{R}^{m \times r}$ satisfies the last condition of (4), $\mathbf{I_r}$ is the r-by-r identity matrix, $\mathbf{0}_{p \times q}$ is the p-by-q matrix of zeros, $\mathbf{P}$ is a permutation matrix and $\tilde{\mathbf{H}}' \in \mathbb{R}^{\tilde{r} \times m-r}$ is a matrix satisfying the two first conditions of (4).

The aim of this work is to introduce an algorithm which, given a r-LQ near separable mixture fulfilling Assumption I.1, recovers the factors $\mathbf{W}$ and $\tilde{\mathbf{H}}$ up to a permutation. To do so, we generalize the SNPA [16] by explicitly modeling the quadratic products along the greedy search process. The resulting sucessive nonnegative projection algorithm for linear quadratic mixtures (SNPALQ) is presented in the next section. The effectiveness of the algorithm is attested through extensive numerical experiments in Section III.

Notations – In the following, we denote $[\![1,r]\!] = \{1,2,..,r\}$, $|\mathcal{K}|$ the number of elements in the set $\mathcal{K}$ whose ith element is denoted $\mathcal{K}(i)$. Matrices are written as $\mathbf{A} \in \mathbb{R}^{m \times r}$, a column indexed by $i \in [\![1,r]\!]$ as $\mathbf{a}_i$, and a row indexed by $j \in [\![1,m]\!]$ as $\mathbf{a}^j$. The submatrix formed by the columns (resp. rows) indexed by $\mathcal{K}$ is denoted $\mathbf{A}_{\mathcal{K}}$ (resp. $\mathbf{A}^{\mathcal{K}}$). The set Δ^r, for which the superscript is omitted when clear from the context, is $\Delta^r = \{x \in \mathbb{R}^r | x \geq 0, \sum_{i=1}^r x_i \leq 1\}$. Furthermore, we write as $\Pi_{\odot}(\mathbf{W})$ the matrix containing all the columns of $\mathbf{W}$ and their quadratic products $\Pi_{\odot}(\mathbf{W}) = [\mathbf{w}_1, \mathbf{w}_2, .., \mathbf{w}_r, \mathbf{w}_2 \odot \mathbf{w}_1, \mathbf{w}_3 \odot \mathbf{w}_1, \mathbf{w}_3 \odot \mathbf{w}_2, ..., \mathbf{w}_r \odot \mathbf{w}_{r-1}] = [\mathbf{w}_i, \mathbf{w}_i \odot \mathbf{w}_j]$, for $i,j \in [\![1,r]\!]$ and $j < i$.

II. Proposed algorithm

To handle LQ near-separable HS unmixing (see Assumption I.1), a first (naive) approach would be to use a pure-pixel search algorithm based on the LMM and to look for $\tilde{r}$ endmembers: as the quadratic terms $(\mathbf{w}_i \odot \mathbf{w}_j)_{\substack{i,j \in [\![1,r]\!] \\ j<i}}$ can be considered as *virtual* endmembers, they could be retrieved along with the columns of $\mathbf{W}$, provided that they appear as pure pixels. One could for instance choose to use SNPA [16], which has shown to yield very good separation performances compared to other algorithms such as VCA and SPA [15]. Due to mathematical guarantees, if the LQ mixing follows SNPA requirements, then the algorithm would extract the $\tilde{r}$ endmembers defining $\Pi_{\odot}(\mathbf{W})$. It would then be easy to check from this solution which ones are the columns of $\mathbf{W}$. Nevertheless, for practical scenarios, the pure-pixel assumption on the *virtual* endmembers might be a too strong requirement, as one cannot expect all virtual endmembers to be present in the data set. As such, the recovery by SNPA is not guaranteed, calling for an algorithm better designed for LQ mixtures. The rationale of the proposed SNPALQ algorithm is that we are mostly interested by recovering the *linear* endmembers and thus the virtual endmembers $\mathbf{w}_i \odot \mathbf{w}_j$ are only taken into account in order to improve the estimation of $\mathbf{W}$. Therefore, we propose to adapt SNPA to the LQ mixing model, reducing at each iteration of the unmixing process the weight of the contribution of the terms $\mathbf{w}_i \odot \mathbf{w}_j$ for $i \neq j$.

[3]More precisely, it is shown that, under the additional assumption that $\text{rowrank}(\mathbf{X}) = \frac{r(r+1)}{2}$, if $\hat{\mathbf{W}}$ and $\hat{\mathbf{H}}$ can be found such that $\mathbf{X} = \Pi_{\odot}(\hat{\mathbf{W}})\hat{\mathbf{H}}$, then $\hat{\mathbf{W}} = \mathbf{W}$ and $\hat{\mathbf{H}} = \tilde{\mathbf{H}}$ up to a scaling and permutation of the columns of $\hat{\mathbf{W}}$ and the rows of $\hat{\mathbf{H}}$. Note that this is not anymore true when squared terms are added, calling for additional priors such as non-negativity.

A. SNPALQ algorithm

The pseudocode[4] of the proposed SNPALQ is described in Algo. 1. Similarly to SNPA, it is a greedy algorithm. At each iteration, the column of the residual matrix $\mathbf{R}$ with the largest ℓ_2 norm is extracted[5] [16, Assumption 2]. SNPALQ and SNPA however differ by their respective projection steps:

- In SNPA, each column of $\mathbf{X}$ is projected onto the convex hull formed by the origin and all the columns extracted so far;
- In SNPALQ, we propose to perform the projection of each column of $\mathbf{X}$ onto the convex hull formed by the origin, the columns extracted so far *and their second order products*.

Therefore, if two endmembers $\mathbf{w}_i$ and $\mathbf{w}_j$ for $i \neq j$ are extracted during the iterative process of SNPALQ, the contribution of the virtual endmember $\mathbf{w}_i \odot \mathbf{w}_j$ is removed. Beyond the advantage that $\mathbf{w}_i \odot \mathbf{w}_j$ will not be extracted in subsequent steps, the non-linear contribution is reduced in the unmixing process, giving more weight to the linear part. Thus, the endmembers $\mathbf{W}$ are more likely to be extracted in the early steps of the iterative process, and SNPALQ will empirically extract much less mixed pixels than SNPA; see Section III.

Algorithm 1 SNPALQ: SNPA for Linear Quadratic mixtures

Input: A r-LQ r-near-separable matrix $\mathbf{X} \in \mathbb{R}^{m \times \tilde{r}}$ satisfying constraints (4), the number r of endmembers.
Initialization: $\mathbf{R} = \mathbf{X}$, $\mathcal{K} = \{\}$
% Greedy search
while $|\mathcal{K}| \leq r$ **do**
  $p = \operatorname{argmax}_{j \in [\![1,n]\!]} \|\mathbf{r}_j\|_2$
  $\mathcal{K} = \mathcal{K} \cup \{p\}$
  for $j \in [\![1,n]\!]$ **do**

$$\hat{\mathbf{h}}_j = \underset{\mathbf{h} \in \Delta^{\frac{|\mathcal{K}|(|\mathcal{K}|+1)}{2}}}{\operatorname{argmin}} \|\mathbf{x}_j - \Pi_{\odot}(\mathbf{X}_{\mathcal{K}})\mathbf{h}\|_2 \qquad (5)$$

$$\mathbf{r}_j = \mathbf{x}_j - \Pi_{\odot}(\mathbf{X}_{\mathcal{K}})\hat{\mathbf{h}}_j$$

  end for
end while
Output: Set $\mathcal{K}$ of indices such that $\mathbf{X}_{\mathcal{K}} \simeq \mathbf{W}$ up to a permutation.

Let us make a few other remarks concerning Algo. 1. First, the projection step (5) is a convex optimization problem, solved using the fast gradient method of [28] (see [16, Appendix A]). Second, in case of ties for finding p, the same rule as in SNPA is used, that is, the index j for which $\|\mathbf{x}_j\|_2$ is maximum is selected.

B. Computational complexity

The complexity of the lth iteration of SNPALQ is dominated by computing the projection step (5), which requires the projection of an m-by-n matrix onto a convex hull with $l(l+1)/2+1$ vertices, yielding to $\mathcal{O}\left(mnl^2\right)$ operations when using a first-order method. This corresponds to an asymptotic complexity of the order of $\mathcal{O}\left(mnr^2\right)$. Note that SNPA requires $\mathcal{O}\left(mnr\right)$ operations hence SNPALQ is r times more expensive.

III. Numerical results

We here study the algorithm behavior on the realistic data sets described in subsection III-A. While subsection III-B dwells on noiseless mixtures, in subsection III-C the robustness of SNPALQ in the presence of noise is studied, first on linear mixtures and then depending on the non-linearity level. The results of SNPALQ are compared with the ones of SNPA [16] and SPA [15].

A. Experimental setting and metrics

1) Data sets: Experiments reported in this paper are conducted on realistic data sets $\mathbf{X}$ on the form described in Assumption I.1:

- The endmembers signatures $\mathbf{W}$ are extracted from the USGS database[4,6]. They correspond to spectra from diverse origins (minerals, soils, plants...) and naturally follow $0 \leq \mathbf{W} \leq 1$, ensuring the first point extracted by SNPALQ, SNPA and SPA to be an endmember since this implies $\mathbf{w}_i \geq \mathbf{w}_i \odot \mathbf{w}_j$ for all i, j.
- The matrix $\tilde{\mathbf{H}}'$ is generated in the following way:
  - The columns of a first matrix $\acute{\mathbf{H}}$ of the same dimension as $\tilde{\mathbf{H}}'$ are generated randomly using a realistic [29] Dirichlet distribution $\mathcal{D}(\alpha, ..., \alpha)$ with $\alpha = 0.5$.
  - The r first rows (corresponding to the linear contribution) are multiplied by $1-\nu$, with ν a non-linearity parameter, while the remaining rows (corresponding to the virtual endmembers) are multiplied by ν to enable various non-linearity levels:

$$\tilde{\mathbf{H}}' = \begin{bmatrix} \acute{\mathbf{H}}^{[\![1,r]\!]} \times (1-\nu) \\ \acute{\mathbf{H}}^{[\![r+1,\tilde{r}]\!]} \times \nu \end{bmatrix}.$$

  - Since the columns are assumed to sum to (at most) one, the last step divides each column of $\tilde{\mathbf{H}}'$ by its ℓ_1 norm.
- The matrix $\mathbf{N}$ is generated using a centered Gaussian distribution with a variance corresponding to a given signal-to-noise ratio (SNR).
- The matrix $\mathbf{X}$ is created ensuring all the entries to be non-negative: $\mathbf{X} = \left[\Pi_{\odot}(\mathbf{W})\tilde{\mathbf{H}} + \mathbf{N}\right]_+$, where $[.]_+$ is the elementwise projection on the non-negative orthant. We denote $\bar{\mathbf{X}} = \Pi_{\odot}(\mathbf{W})\tilde{\mathbf{H}}$ the noiseless mixing.

2) Metrics: The quality is assessed using the minimum Spectral Angle Distance (SAD) over the endmembers:

$$\theta_{\min} = \min_{i \in [\![1,r]\!]} \text{SAD}(\mathbf{w}_i, \bar{\mathbf{x}}_{\mathcal{K}(i)})$$

[4]The algorithm and data used will be made available online at https://sites.google.com/site/nicolasgillis/code

[5]Note that any strongly convex function $f(\cdot)$ can be used instead of the ℓ_2 norm, provided that $f(0) = 0$

[6]https://www.usgs.gov/

with $\text{SAD}(\mathbf{u}, \mathbf{v}) = \frac{\mathbf{u}^T\mathbf{v}}{\|\mathbf{u}\|_2\|\mathbf{v}\|_2}$ and where the set of indices $\mathcal{K}$ is permuted to minimize θ_{min}. We consider a perfect separation is achieved if $\theta_{\text{min}} > 0.999$.

B. Numerical results on noiseless mixtures

We here assess the behavior of SNPALQ as a function of the number of endmembers r in a noiseless setting. There are $n = 1000$ pixels observed in $m = 50$ spectral bands. The non-linearity parameter is chosen as $\nu = 0.5$. We conducted 100 Monte-Carlo experiments, each time generating a new data set. The percentage of perfect separation by the different algorithms is displayed in Fig. 1.

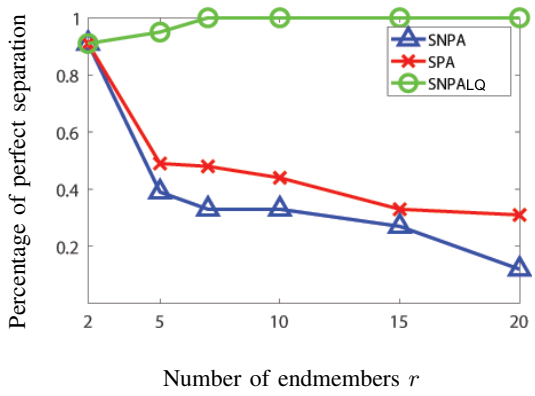

Fig. 1. Percentage of experiments in which a perfect separation is achieved.

In this experiment, SNPALQ clearly shows its interest, as it obtains much better results than SNPA and SPA, and achieves in more than 90% of the experiments a perfect separation. The improvement when r increases is linked to the use of a Dirichlet distribution with $\alpha = 0.5$ when randomly generating the mixing coefficient matrix $\tilde{\mathbf{H}}'$. When r is small, the data points are more spread within the convex hull formed by the origin, the endmembers and the virtual endmembers, leading to a higher probability for a virtual endmember to be extracted. SNPALQ and SNPA obtain the same results for $r = 2$, as both algorithms are identical in this case. On the other hand, SNPA and SPA results deteriorate quickly when r increases.

C. Numerical results on noisy mixtures

We now investigate SNPALQ robustness to noise in case of linear and linear-quadratic mixings.

Linear mixing – If the input matrix follows the LMM, it also satisfies Assumption I.1 (with $\tilde{\mathbf{H}}'^{[\![r+1,\tilde{r}]\!]} = 0$). However, while SNPALQ can thus be applied on linear mixings, it is expected to perform worse than SNPA since it projects the residual onto non-existing virtual endmembers, which leads to a loss of information (the norm of the residual will decrease faster), especially in the presence of noise. The goal of this experiment is to assess the performance of SNPALQ against SNPA under the LMM model for various noise levels.

The number of spectral bands is set to $m = 50$, there are $n = 1000$ pixels and $r = 10$ endmembers. To obtain linear mixtures, ν was set to 0 in the 100 Monte-Carlo experiments. The average θ_{min} value as a function of the SNR is displayed in Figure 2.

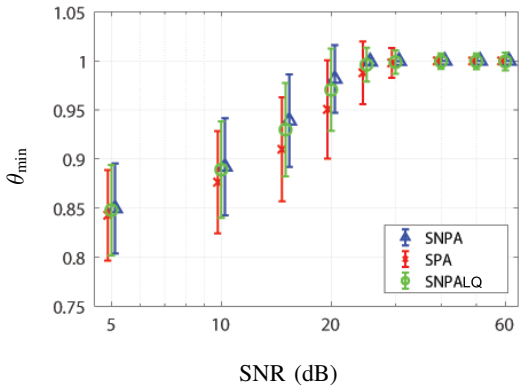

Fig. 2. Avergage θ_{min} in the linear case. The error bars correspond to the standard deviation.

This figure shows that SNPALQ, SNPA and SPA achieve similar results for high SNR values (namely, SNR $\geq$ 40 dB), all reaching a perfect separation. The main discrepancies appear in the range SNR $\in [10\text{ dB}, 30\text{ dB}]$, in which SNPA obtains better results than SNPALQ, which was expected (see above). The difference between the two algorithms is however relatively mild, showing that even in this setting SNPALQ obtains competitive results compared to SNPA.

Linear-quadratic mixing – The impact of the noise and non-linearity levels is now studied: we generated data sets $\mathbf{X}$ with 12 different SNR levels and 12 values for the non-linearity parameter ν. For each SNR and each ν, 100 Monte-Carlo experiments were performed on nonlinear mixtures characterized by $m = 50$ spectral bands, $r = 10$ endmembers and $n = 1000$ pixels. The recovery performance of SNPALQ and SNPA is depicted as a 2-dimensional map in Figure 3.

SNPALQ mostly shows its interest over SNPA when the non-linearity level increases and more specifically when the noise is relatively small (upper-right corner of the figures). More precisely, when $\nu \geq 0.3$ and SNR $\geq$ 30 dB, SNPALQ always obtains a perfect recovery, which represents an improvement over SNPA, up to 12%. In the lowest right corner of the figures, when the SNR decreases, the results of both algorithms deteriorate as the problem is highly difficult. Lastly, we mention that SPA results are similar to that of SNPA and are not shown here.

CONCLUSION

To tackle the problem of linear-quadratic hyperspectral unmixing, we introduced SNPALQ, an extension of SNPA which takes into account the presence of quadratic terms in the projection step. SNPALQ was shown to obtain good results on non-linear realistic data sets. Furthermore, SNPALQ empirically exhibited a good robustness to noise, including when it is applied on linear mixtures, in which case it obtains competitive results with respect to SNPA. Future works include theoretical

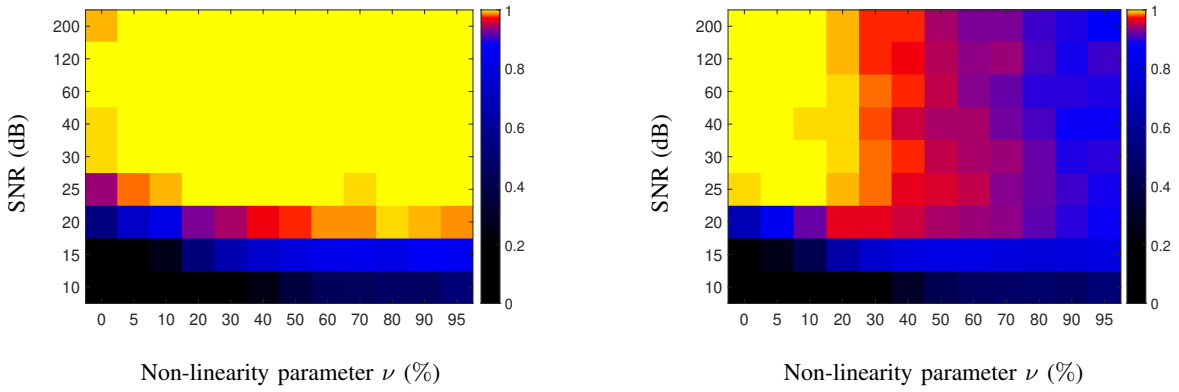

Fig. 3. Percentage of perfect separation using: SNPALQ *(Left)*, SNPA *(Right)*.

guarantees on such a robustness and applying SNPALQ on real data sets.

REFERENCES

[1] M. E. Schaepman, S. L. Ustin, A. J. Plaza, T. H. Painter, J. Verrelst, and S. Liang, "Earth system science related imaging spectroscopy—an assessment," *Remote Sensing of Environment*, vol. 113, pp. S123–S137, 2009.

[2] H. Akbari, K. Uto, Y. Kosugi, K. Kojima, and N. Tanaka, "Cancer detection using infrared hyperspectral imaging," *Cancer science*, vol. 102, no. 4, pp. 852–857, 2011.

[3] K. E. Themelis, F. Schmidt, O. Sykioti, A. A. Rontogiannis, K. D. Koutroumbas, and I. A. Daglis, "On the unmixing of mex/omega hyperspectral data," *Planetary and Space Science*, vol. 68, no. 1, pp. 34–41, 2012.

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

[5] P. Comon and C. Jutten, *Handbook of Blind Source Separation: Independent component analysis and applications*. Academic Press, 2010.

[6] J. Bobin, J. Rapin, A. Larue, and J.-L. Starck, "Sparsity and adaptivity for the blind separation of partially correlated sources." *IEEE Transactions on Signal Processing*, vol. 63, no. 5, pp. 1199–1213, 2015.

[7] C. Kervazo, J. Bobin, C. Chenot, and F. Sureau, "Use of palm for ℓ_1 sparse matrix factorization: Difficulty and rationalization of an heuristic approach." *Digital Signal Processing*, in press.

[8] N. Gillis, "The why and how of nonnegative matrix factorization," *Regularization, optimization, kernels, and support vector machines*, vol. 12, no. 257, pp. 257–291, 2014.

[9] S. A. Vavasis, "On the complexity of nonnegative matrix factorization," *SIAM Journal on Optimization*, vol. 20, no. 3, pp. 1364–1377, 2010.

[10] S. Arora, R. Ge, R. Kannan, and A. Moitra, "Computing a nonnegative matrix factorization—provably," *SIAM Journal on Computing*, vol. 45, no. 4, pp. 1582–1611, 2016.

[11] N. Gillis and S. A. Vavasis, "Fast and Robust Recursive Algorithms for Separable Nonnegative Matrix Factorization," *arXiv*, 2012.

[12] W.-K. Ma, J. M. Bioucas-Dias, T.-H. Chan, N. Gillis, P. Gader, A. J. Plaza, A. Ambikapathi, and C.-Y. Chi, "A signal processing perspective on hyperspectral unmixing: Insights from remote sensing," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 67–81, 2013.

[13] M. Möller, E. Esser, S. Osher, G. Sapiro, and J. Xin, "A convex model for matrix factorization and dimensionality reduction on physical space and its application to blind hyperspectral unmixing," Institute for Mathematics and its Applications, University of Minnesota, Minneapolis, MN, Tech. Rep., 2010.

[14] B. Recht, C. Re, J. Tropp, and V. Bittorf, "Factoring nonnegative matrices with linear programs," in *Advances in Neural Information Processing Systems*, 2012, pp. 1214–1222.

[15] U. Araújo, B. Saldanha, R. Galvão, T. Yoneyama, H. Chame, and V. Visani, "The successive projections algorithm for variable selection in spectroscopic multicomponent analysis," *Chemometrics and Intelligent Laboratory Systems*, vol. 57, no. 2, pp. 65–73, 2001.

[16] N. Gillis, "Successive nonnegative projection algorithm for robust nonnegative blind source separation," *SIAM Journal on Imaging Sciences*, vol. 7, no. 2, pp. 1420–1450, jan 2014.

[17] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observations Remote Sens.*, vol. 5, no. 2, pp. 354–379, 2012.

[18] N. Dobigeon, Y. Altmann, N. Brun, and S. Moussaoui, "Linear and nonlinear unmixing in hyperspectral imaging," in *Data Handling in Science and Technology*. Elsevier, 2016, vol. 30, pp. 185–224.

[19] I. Meganem, Y. Deville, S. Hosseini, P. Deliot, and X. Briottet, "Linear-quadratic blind source separation using nmf to unmix urban hyperspectral images," *IEEE Trans. Signal Process.*, vol. 62, no. 7, pp. 1822–1833, 2014.

[20] N. Dobigeon, L. Tits, B. Somers, Y. Altmann, and P. Coppin, "A comparison of nonlinear mixing models for vegetated areas using simulated and real hyperspectral data," *IEEE J. Sel. Topics Appl. Earth Observations Remote Sens.*, vol. 7, no. 6, pp. 1869–1878, 2014.

[21] 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, 2012.

[22] I. Meganem, P. Déliot, 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, 2013.

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

[24] Y. Deville and L. T. Duarte, "An overview of blind source separation methods for linear-quadratic and post-nonlinear mixtures," in *International Conference on Latent Variable Analysis and Signal Separation*. Springer, 2015, pp. 155–167.

[25] C. Kervazo and J. Bobin, "Stacked sparse blind source separation for non-linear mixtures," In press.

[26] Y. Deville, "From separability/identifiability properties of bilinear and linear-quadratic mixture matrix factorization to factorization algorithms," *Digital Signal Processing*, vol. 87, pp. 21–33, apr 2019.

[27] N. Gillis, "Sparse and Unique Nonnegative Matrix Factorization Through Data Preprocessing," *arXiv*, 2012.

[28] Y. Nesterov, *Introductory lectures on convex optimization: A basic course*. Springer Science & Business Media, 2013, vol. 87.

[29] J. M. Nascimento and J. M. Bioucas-Dias, "Hyperspectral unmixing based on mixtures of dirichlet components," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 3, pp. 863–878, 2011.