Quaternion Non-negative Matrix Factorization: a new tool for spectropolarimetric imaging

Julien Flamant
CRAN, CNRS
Université de Lorraine
Nancy, France
julien.flamant@univ-lorraine.fr

Sebastian Miron
CRAN, CNRS
Université de Lorraine
Nancy, France
sebastian.miron@univ-lorraine.fr

David Brie
CRAN, CNRS
Université de Lorraine
Nancy, France
david.brie@univ-lorraine.fr

***Abstract*—Quaternion non-negative matrix factorization (QNMF) is a new tool which generalizes usual non-negative matrix factorization (NMF) to the case of polarized signals. The approach relies on two key features: *(i)* the algebraic representation of polarization information, namely Stokes parameters, thanks to quaternions and *(ii)* the exploitation of physical constraints linked to polarization generalizing non-negativity constraints. QNMF improves NMF model identifiability by revealing the key disambiguating role played by polarization information. A simple and numerically efficient algorithm is introduced for practical resolution of the QNMF problem. Numerical experiments on synthetic data validate the proposed approach and illustrate the potential of QNMF as a generic spectropolarimetric image unmixing tool.**

***Index Terms*—quaternion non-negative matrix factorization (QNMF), Stokes parameters, spectropolarimetry**

I. Introduction

Polarization information is essential to many imaging applications ranging from astrophysics [1] to biology [2]. It simultaneously carries a strong natural discriminative power as well as numerous physical and morphological attributes of the observed scene which are unaccessible to conventional intensity imaging. Recent years have seen a booming interest in pairing polarization diversity together with conventional hyperspectral imaging systems [3]. Such *spectropolarimetric imaging* technique consists in acquiring, for every pixel u and wavelength λ, a Stokes vector $\mathbf{x}(\lambda, u) \in \mathbb{R}^4$ describing the polarization properties of the measured light. Importantly, these energetic parameters obey to structural constraints that reveal the physics of polarization. These constraints generalize classical non-negativity: only one Stokes parameter is non-negative, whereas the 4 Stokes parameters are linked all together by an inequality constraint.

Spectropolarimetric images can be represented as a 3D data array gathering the spatial, spectral and polarization diversities. For narrow-band sources with constant polarization, usual low-rank tensor approximation techniques (*e.g.* CPD with constraints [4]) – adapted to the specificities of polarization – permit an efficient unmixing of spectropolarimetric images. However, the general wideband sources case is much more challenging, notably due to wavelength-dependent polarization properties. The latter case thus requires the development of new, generic and interpretable signal processing tools for spectropolarimetric data.

To this aim, we introduce quaternion non-negative matrix factorization (QNMF). This new tool extends to the case of polarized signals the concept of non-negative matrix factorization (NMF), widely used for hyperspectral unmixing [5]. QNMF relies on two ingredients: *(i)* the algebraic representation of the Stokes vector $\mathbf{x}(\lambda, u)$ using quaternions and *(ii)* the exploitation of physical constraints related to polarization, which generalize the usual non-negativity constraint. Focusing on practical purposes, we give a simple yet numerically efficient algorithm that solves the QNMF problem using a quaternion-domain alternated least square optimization strategy. A first theoretical analysis of QNMF can be found in [6], where it is shown that QNMF greatly improves NMF in terms of model identifiability, by having less strict uniqueness conditions on the sources. Numerical experiments on synthetic data demonstrate the potential of QNMF as a generic low-rank approximation tool for polarized signals.

II. Background

A. Polarization

Polarization is a fundamental property of waves (electromagnetic, elastic, gravitational) which describes the geometric nature of oscillations. Consider for instance light in vacuum: in the transverse plane, *i.e.* perpendicular to the direction of propagation, the electric field vector usually describes an elliptical trajectory; hence one says that the wave is elliptically polarized. Intensity and polarization properties of light are characterized by 4 Stokes parameters S_0, S_1, S_2, S_3. This set of energetic, experimentally measurable quantities is widely used in many applications, including optics [7]. The first Stokes parameter $S_0 \geq 0$ measures the *total* intensity, *i.e.* the sum of intensities of the *polarized* part and *unpolarized* part of light. The relative contribution of these two parts is ruled by the *degree of polarization* Φ

$$\Phi = \frac{\text{polarized intensity}}{\text{total intensity}} = \frac{\sqrt{S_1^2 + S_2^2 + S_3^2}}{S_0}\,, \qquad (1)$$

where by definition $\Phi \in [0, 1]$. When $\Phi = 1$, light is said to be *fully polarized*, whereas for $\Phi = 0$ light is said to

Research supported by Région Grand Est and project ANR-15-CE10-0007 OPTIFIN.

be *unpolarized*. For $0 < \Phi < 1$, it is said to be *partially polarized*. The three remaining Stokes parameters S_1, S_2, S_3 define the polarization state fo the wave: S_1 and S_2 encode linear polarization states, while S_3 gives circularly polarized states.

B. Quaternion representation

A Stokes vector $(S_0, S_1, S_2, S_3)^\top \in \mathbb{R}^4$ can be algebraically represented by the quaternion

$$w = S_0 + \boldsymbol{i}S_3 + \boldsymbol{j}S_1 + \boldsymbol{k}S_2 \in \mathbb{H}\,, \tag{2}$$

where $\mathbb{H} = \mathrm{Span}\{1, \boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$ is the set of quaternions, $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ are roots of -1 ($\boldsymbol{i}^2 = \boldsymbol{j}^2 = \boldsymbol{k}^2 = -1$) such that $\boldsymbol{ij} = \boldsymbol{k}$ and $\boldsymbol{ij} = -\boldsymbol{ji}$. Eq. (2) can be geometrically rewritten as:

$$w = I + I\Phi\boldsymbol{\mu} \tag{3}$$

where $I = S_0 \geq 0$ is the total intensity, $\Phi \in [0, 1]$ is the degree of polarization. The *polarization axis* $\boldsymbol{\mu}$ is a pure quaternion (*i.e.* such that $\boldsymbol{\mu}^2 = -1$) which can be identified with a unit vector of $\mathbb{R}^3$ on the Poincaré sphere of polarization states [8]. Eq. (3) illustrates one of the benefits of the quaternion approach, as it enables a natural separation between purely energetic information (the real part of w, $\mathrm{Re}w = I$) and polarization / geometric content (the imaginary part of w, $\mathrm{Im}w = I\Phi\boldsymbol{\mu}$).

C. Polarization constraint and non-negativity

A Stokes vector $(S_0, S_1, S_2, S_3)^\top \in \mathbb{R}^4$ is said to be admissible when the following constraints ($\mathcal{S}$) are satisfied:

$$S_0 \geq 0 \quad \text{and} \quad S_1^2 + S_2^2 + S_3^2 \leq S_0^2\,. \tag{$\mathcal{S}$}$$

The first constraint $S_0 \geq 0$ is classical and indicates that the total intensity is a non-negative real quantity. The second constraint is specific to polarization: it stipulates that the intensity of the polarized part cannot exceed that of the total intensity. From a mathematical viewpoint, ($\mathcal{S}$) extends the usual non-negativity constraint of univariate signals to the case of bivariate signals. To see this, consider the *coherency* matrix $\mathbf{J}$ [9, Section 1.4], *i.e.* the spectral covariance matrix of the bivariate electric field such that

$$\mathbf{J} = \frac{1}{2}\begin{bmatrix} S_0 + S_1 & S_2 + \boldsymbol{i}S_3 \\ S_2 - \boldsymbol{i}S_3 & S_0 - S_1 \end{bmatrix} \in \mathbb{C}^{2\times 2}. \tag{4}$$

By definition, $\mathbf{J}$ is a positive semidefinite (or simply non-negative) Hermitian matrix. Non-negativity of $\mathbf{J}$ is characterized by $\operatorname{tr}\mathbf{J} \geq 0$ and $\det\mathbf{J} \geq 0$, two conditions which yield directly to ($\mathcal{S}$). The proposed approach relies on exploiting the constraint ($\mathcal{S}$) through the quaternion representation (2) of Stokes parameters. By extension, we define the set of non-negative quaternions $\mathbb{H}_\mathcal{S} \subset \mathbb{H}$ such that

$$\mathbb{H}_\mathcal{S} \triangleq \{q \in \mathbb{H} | \mathrm{Re}q \geq 0 \text{ et } |\mathrm{Im}q| \leq \mathrm{Re}q\}\ . \tag{5}$$

This formal identification between a quaternion $q \in \mathbb{H}_\mathcal{S}$ and a Stokes vector is used throughout the paper.

III. Quaternion non-negative matrix factorization

This section introduces quaternion non-negative matrix factorization (QNMF), a new tool which generalizes NMF to the case of polarized signals. QNMF relies on two key distinctives aspects: *(i)* the polarization constraints ($\mathcal{S}$), which encodes non-negativity for polarized signals and *(ii)* the algebraic representation (2) of Stokes parameters using quaternions.

A. Definition

Consider for simplicity the case of spectropolarimetric imaging, where the four Stokes parameters $x(\lambda, u) \in \mathbb{H}_\mathcal{S}$ are measured for every wavelength λ and pixel u. The following linear mixing model permits to decompose spectropolarimetric data into P sources like

$$x(\lambda, u) = \sum_{p=1}^{P} w_p(\lambda) h_p(u) \tag{6}$$

where $h_p(u) \geq 0$ is the activation coefficient of the p^{th} source at pixel u, and $w_p(\lambda) \in \mathbb{H}_\mathcal{S}$ is the corresponding Stokes vector at wavelength λ. Eq. (6) describes a blind wideband polarized source separation problem. This model is very general: no hypotheses on the frequency dependence of polarization properties are introduced, unlike other existing approaches based on narrow-band modeling [10].

Solving (6) can be formulated as the resolution of the following quaternion non-negative matrix factorization (QNMF) problem

$$\mathbf{X} = \mathbf{W}\mathbf{H} \tag{7}$$

where $\mathbf{X} \in \mathbb{H}_\mathcal{S}^{M\times N}$ is the spectropolarimetric data matrix, *i.e.* $\mathbf{X}_{mn} = x(\lambda_m, u_n)$. The source matrix $\mathbf{W} \in \mathbb{H}_\mathcal{S}^{M\times P}$ has coefficients $\mathbf{W}_{mp} = w_p(\nu_m)$, while $\mathbf{H} \in \mathbb{R}_+^{P\times N}$ is the activation matrix such that $\mathbf{H}_{pn} = h_p(u_n)$. Given $\mathbf{X}$, the QNMF (7) relies on exploiting the polarization onstraint ($\mathcal{S}$) for the sources $\mathbf{W}$, together with the usual non-negativity constraint for the activations $\mathbf{H}$. As a result, for a given $P \leq M, N$, the product $\mathbf{WH}$ defines a structured low-rank approximation of $\mathbf{X}$.

B. Relation with NMF

The QNMF problem (7) can be rewritten using the decomposition of any quaternion into the sum of its real and imaginary part like

$$\mathbf{X} = \mathbf{WH} \Leftrightarrow \begin{cases} \mathrm{Re}\mathbf{X} = [\mathrm{Re}\mathbf{W}]\,\mathbf{H} & \text{(NMF)} \\ \mathrm{Im}\mathbf{X} = [\mathrm{Im}\mathbf{W}]\,\mathbf{H} & \text{(polarization)} \end{cases}. \tag{8}$$

Eq. (8) shows that QNMF can be interpreted as a matrix cofactorization problem, where activation matrix $\mathbf{H}$ stands as a common factor. The first problem $\mathrm{Re}\mathbf{X} = [\mathrm{Re}\mathbf{W}]\,\mathbf{H}$ is a standard NMF one on the real part of $\mathbf{X}$, *i.e.* on intensity data only (first Stokes parameter $S_0 \geq 0$). The second problem $\mathrm{Im}\,\mathbf{X} = [\mathrm{Im}\,\mathbf{W}]\,\mathbf{H}$ searches for a factorization of the imaginary part of $\mathbf{X}$, which describes polarization properties of the sources (Stokes parameters S_1, S_2, S_3). Importantly, these two matrix factorization problems are not independent, for two reasons: *(i)* the activation matrix $\mathbf{H}$ is a common factor and *(ii)* the very nature of the polarization constraint ($\mathcal{S}$) links the real and imaginary parts of the source matrix $\mathbf{W}$. In full generality, QNMF greatly improves NMF model identifiablity by taking advantage of supporting polarization information – fundamental properties that can be measured in most optical imaging setups. From a theoretical perspective, we derive in [6] sufficient conditions (for $P = 2$ sources) and necessary uniqueness conditions (for $P \geq 2$ sources). Compared to their NMF counterparts [11]–[14], these conditions appear far less restrictive – in particular, one may recover QNMF uniqueness even when sources never vanish, a case where NMF is known to be not unique.

IV. Algorithm

A. Optimization problem formulation

The practical resolution of the QNMF problem can be formulated as an optimization problem, which aims at minimizing the following Euclidean quadratic cost

$$\min_{\substack{\mathbf{W}\in\mathbb{H}_{\mathcal{S}}^{M\times P}\\ \mathbf{H}\in\mathbb{R}_{+}^{P\times N}}} \|\mathbf{X} - \mathbf{W}\mathbf{H}\|_F^2 \tag{9}$$

where $\|\mathbf{X} - \mathbf{Y}\|_F^2 = \sum_{m,n} |\mathbf{X}_{mn} - \mathbf{Y}_{mn}|^2$ is the Frobenius norm between quaternion matrices $\mathbf{X}$ and $\mathbf{Y}$. Note the similarity between (9) and the standard Euclidean cost NMF. However, the QNMF setting raises two specific issues that subtend any algorithm attempting to solve (9), namely: *1)* can we implement the constraint ($\mathcal{S}$) on $\mathbf{W}$ in an efficient way? and *2)* can we optimize directly with respect to the quaternion matrix $\mathbf{W}$? Fortunately, the key link between ($\mathcal{S}$) and the set of non-negative Hermitian 2-by-2 matrices allows to answer positively to *1)*. For *2)*, recent advances in the theory of quaternion-domain derivatives [15]–[18] (also called *generalized* $\mathbb{HR}$ *calculus*), enable to perform optimization of (9) directly in the quaternion domain, as described below.

B. Quaternion alternating least squares

Since the objective function (9) is biconvex, *i.e.* it is convex in each $\mathbf{W}$ and $\mathbf{H}$, but not in $(\mathbf{W}, \mathbf{H})$, we adopt a popular strategy based on the alternated constrained minimization of (9) w.r.t. $\mathbf{H}$ and $\mathbf{W}$. To this aim, a first approach consists in the following. At a given iteration r, for each factor, one solves the unconstrained least-squares problem and project the obtained solution onto the corresponding constraint. This approach is called *quaternion alternating least squares* (QALS) due to its resemblance with the usual ALS algorithm [19] for the NMF. QALS is written symbolically as

$$\mathbf{H}_{r+1} \leftarrow \Pi_{\mathbb{R}_+}\left[\arg\min_{\mathbf{H}} \|\mathbf{X} - \mathbf{W}_r\mathbf{H}\|_F^2\right] \tag{10}$$

$$\mathbf{W}_{r+1} \leftarrow \Pi_{\mathbb{H}_{\mathcal{S}}}\left[\arg\min_{\mathbf{W}} \|\mathbf{X} - \mathbf{W}\mathbf{H}_{r+1}\|_F^2\right] \tag{11}$$

where $\Pi_{\mathbb{R}_+}$ and $\Pi_{\mathbb{H}_{\mathcal{S}}}$ denote element-wise projections onto constraints $\mathbb{R}_+$ and ($\mathcal{S}$), respectively. Projection onto $\mathbb{R}_+$ is simply obtained by keeping non-negative values, *i.e.* $\Pi_{\mathbb{R}_+}(\mathbf{H}_{pn}) = \max(0, \mathbf{H}_{pn})$. The projection of elements of $\mathbf{W}$ onto ($\mathcal{S}$) is performed by projecting the associated 2-by-2 Hermitian matrix onto the subspace spanned by its non-negative eigenvalues, see [6] for details.

Detailed derivations of explicit updates for unconstrained least-squares problems are provided in [6]. It requires special care due to the quaternion nature of $\mathbf{X}$ and $\mathbf{W}$. For $\mathbf{H}$, it involves a cautious handling of quaternion non-commutativity. Optimization over $\mathbf{W}$ is performed thanks to the recently introduced theory of quaternion derivatives [15]–[18]. It yields updates expressions that are very much alike the standard ALS algorithm for the NMF [20]. At iteration $r > 0$, one gets

$$\mathbf{H}_{r+1} \leftarrow \Pi_{\mathbb{R}_+}\left[\left(\mathrm{Re}\left[\mathbf{W}_r^\top \overline{\mathbf{W}}_r\right]\right)^{-1} \mathrm{Re}\left[\mathbf{W}_r^\top \overline{\mathbf{X}}\right]\right] \tag{12}$$

$$\mathbf{W}_{r+1} \leftarrow \Pi_{\mathbb{H}_{\mathcal{S}}}\left[\mathbf{X}\mathbf{H}_{r+1}^\top \left(\mathbf{H}_{r+1}\mathbf{H}_{r+1}^\top\right)^{-1}\right] \tag{13}$$

where $\mathbf{X}^\top$ and $\overline{\mathbf{X}}$ are the transpose and element-wise conjugate of the quaternion matrix $\mathbf{X}$, respectively. Note that in (12)–(13), matrix inverses are computed for real matrices only. The proposed QALS algorithm is thus remarkably simple and cheap. Its numerical complexity is similar to that of the ALS algorithm, which scales as $\mathcal{O}(MNP)$ for small values of P. This heuristic algorithm provides a first baseline for the resolution of QNMF, and paves the way to the development of more sophisticated algorithms in the near future.

V. Numerical validation

A. Illustration

To illustrate the potential of QNMF, we consider a data example $\mathbf{X}_0 = \mathbf{W}_0\mathbf{H}_0$ constructed from synthetic quaternion non-negative sources $\mathbf{W}_0$ and activation maps $\mathbf{H}_0$. We generate $P = 3$ sources with realistic spectral signatures (S_0 parameter) for $M = 64$ wavelengths together with $N = 16 \times 16 = 256$ spatial locations. For simplicity, each source is fully polarized ($\Phi = 1$) with constant polarization properties distinct from one another. Activation maps are chosen such that the pure pixel condition is satisfied. Such conditions guarantee that the QNMF $\mathbf{X}_0 = \mathbf{W}_0\mathbf{H}_0$ satisfies the necessary uniqueness condition given in [6, Prop. 3].

Additive noise was added to $\mathbf{X}_0$, resulting in observations $\mathbf{X} = \mathbf{X}_0 + \mathbf{N}$ where $\mathbf{N}$ has i.i.d. quaternion circular Gaussian entries, *i.e.* $\mathbf{N}_{mn} = n^0_{mn} + n^1_{mn}\boldsymbol{i} + n^2_{mn}\boldsymbol{j} + n^3_{mn}\boldsymbol{k}$ with i.i.d. $n^i_{mn} \sim \mathcal{N}(0, \sigma^2)$. Signal-to-noise ratio is defined as $\mathrm{SNR} = \|\mathbf{X}_0\|_F^2/(4MN\sigma^2)$ and we note $\mathrm{SNR}_{\mathrm{dB}} = 10\log\mathrm{SNR}$.

Fig. 1 depicts reconstruction results for sources and corresponding activation maps obtained using the QALS algorithm for $\mathrm{SNR}_{\mathrm{dB}} = 0$ dB. We performed 100 independent runs of the QALS algorithm. Convergence of the algorithm is monitored by the relative reconstruction error at iteration r, $\varepsilon_r = \|\mathbf{X} - \mathbf{W}_r\mathbf{H}_r\|_F^2/\|\mathbf{X}\|_F^2$. The algorithm stops whenever $|\varepsilon_r - \varepsilon_{r-1}|$ goes below some predefined threshold. Out of 100 runs, it took on average $\simeq 8$ iterations for the QALS algorithm to converge. Fig. 1 displays the best of all runs in terms of final reconstruction error. Overall, the excellent

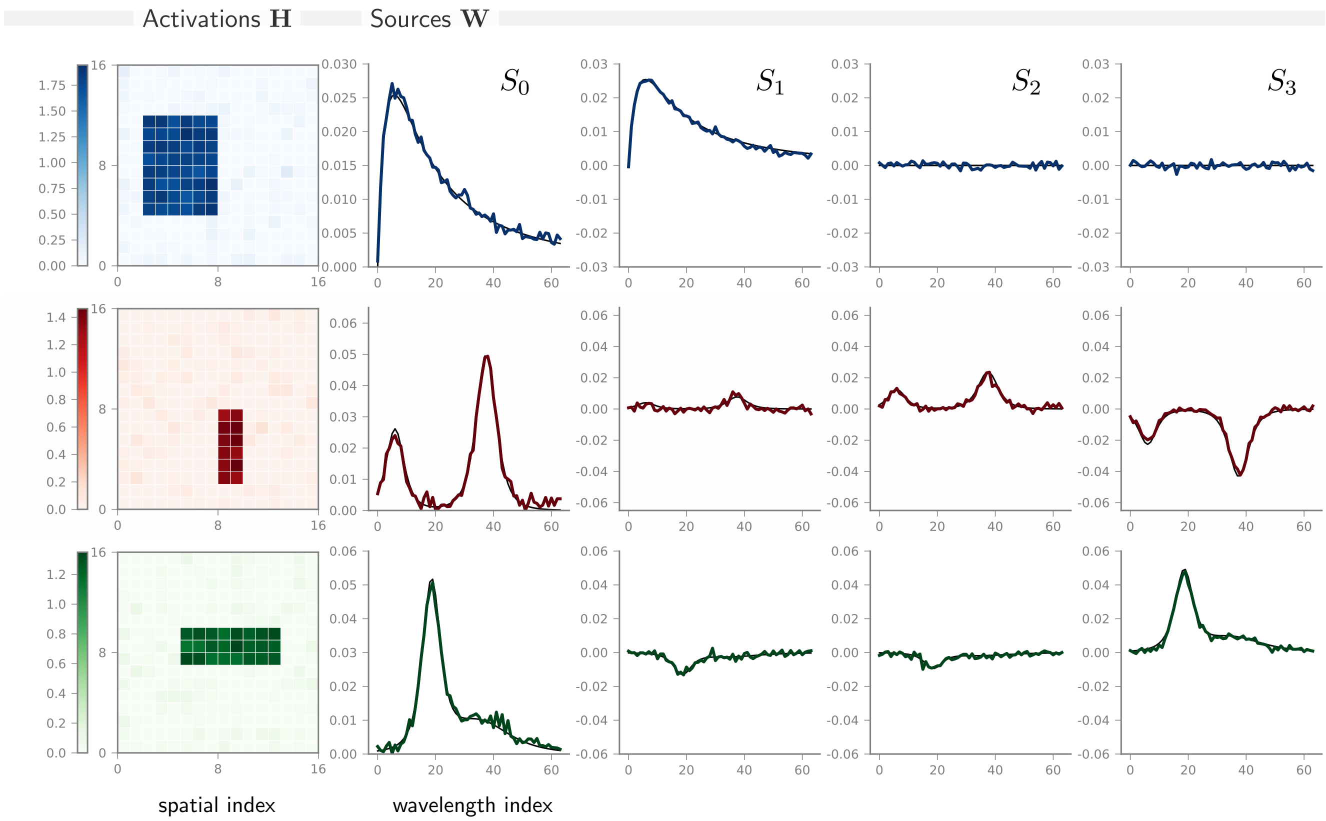

Fig. 1. QALS reconstruction example for the $P = 3$ sources case with $\mathrm{SNR}_{\mathrm{dB}} = 0$ dB. Reconstructed activations and Stokes parameters are given row-wise. Results display excellent match with theoretical values (solid black lines).

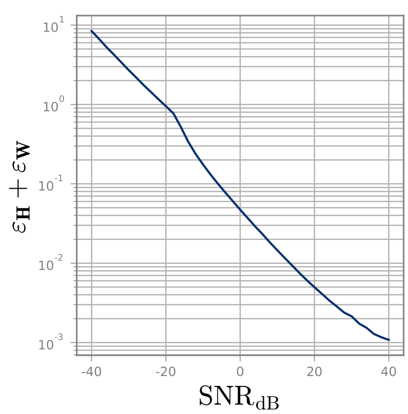

Fig. 2. Evolution of the total reconstruction error $\varepsilon_{\mathbf{H}} + \varepsilon_{\mathbf{W}}$ with respect to the SNR for the sources and activations factors shown in Fig. 1.

match between theoretical values and reconstructed sources and activations indicates the good practical performances of the proposed QALS algorithm.

B. Reconstruction performance

To quantify the performances of QALS in terms of reconstruction accuracy, we consider a second experiment where we monitor the total reconstruction error $\|\mathbf{H}_0 - \mathbf{H}\|_F^2 + \|\mathbf{W}_0 - \mathbf{W}\|_F^2 = \varepsilon_{\mathbf{H}} + \varepsilon_{\mathbf{W}}$ for SNR values ranging from $\mathrm{SNR}_{\mathrm{dB}} = -40$ dB to $\mathrm{SNR}_{\mathrm{dB}} = 40$ dB. For each SNR value, 100 independent noisy observations were simulated and QALS reconstruction was performed.

Fig. 2 displays the evolution of the average total reconstruction error with respect to the SNR. As expected, the total reconstruction error $\varepsilon_{\mathbf{H}} + \varepsilon_{\mathbf{W}}$ decreases as the SNR increases. These results illustrate the good numerical properties of QALS, which provides, despite the lack of convergence guarantees, a good baseline for practical resolution of the QNMF problem.

VI. Conclusions

This paper has introduced QNMF, a new low-rank approximation tool which generalizes the usual NMF to the case of polarized signals. A first algorithm named QALS provides an efficient and simple resolution of the QNMF problem. These first results, together with unicity results provided in [6] appear very promising and herald many further theoretical and methodological developments on QNMF. These future works will bring fundamental tools essential to the generalization of spectropolarimetric imaging modalities.

REFERENCES

[1] R. Antonucci, "Optical spectropolarimetry of radio galaxies," *Astrophys. J.*, vol. 278, pp. 499–520, 1984.

[2] Y. Zhao, L. Zhang, and Q. Pan, "Spectropolarimetric imaging for pathological analysis of skin," *Appl. Opt.*, vol. 48, no. 10, pp. 236–246, 2009.

[3] T. Mu *et al.*, "Static hyperspectral imaging polarimeter for full linear stokes parameters," *Opt. Exp.*, vol. 20, no. 16, pp. 18 194–18 201, 2012.

[4] M. A. Veganzones, J. E. Cohen, R. C. Farias, J. Chanussot, and P. Comon, "Nonnegative tensor cp decomposition of hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 5, pp. 2577–2588, 2015.

[5] W.-K. Ma *et al.*, "A signal processing perspective on hyperspectral unmixing: Insights from remote sensing," *IEEE Sig. Proc. Mag.*, vol. 31, no. 1, pp. 67–81, 2014.

[6] J. Flamant, S. Miron, and D. Brie, "Quaternion non-negative matrix factorization: definition, uniqueness and algorithm," 2019, arXiv:1903.10593.

[7] M. Born and E. Wolf, *Principles of Optics*. Pergamon Press, 1980.

[8] J. Flamant, N. Le Bihan, and P. Chainais, "Spectral analysis of stationary random bivariate signals," *IEEE Trans. Sig. Proc.*, vol. 65, no. 23, pp. 6135–6145, 2017.

[9] J. Gil and R. Ossikovski, *Polarized Light and the Mueller matrix approach*. CRC Press, 2016.

[10] F. E. Raimondi *et al.*, "Wideband multiple diversity tensor array processing," *IEEE Trans. Sig. Proc.*, vol. 65, no. 20, pp. 5334–5346, 2017.

[11] K. Huang, N. D. Sidiropoulos, and A. Swami, "Non-negative matrix factorization revisited: Uniqueness and algorithm for symmetric decomposition," *IEEE Trans. Sig. Proc.*, vol. 62, no. 1, pp. 211–224, 2014.

[12] H. Laurberg *et al.*, "Theorems on positive data: On the uniqueness of nmf," *Computational intelligence and neuroscience*, vol. 2008, 2008.

[13] S. Moussaoui, D. Brie, and J. Idier, "Non-negative source separation: range of admissible solutions and conditions for the uniqueness of the solution," in *IEEE ICASSP'05*, vol. 5, 2005, pp. v–289.

[14] X. Fu, K. Huang, N. D. Sidiropoulos, and W.-K. Ma, "Nonnegative matrix factorization for signal and data analytics: Identifiability, algorithms, and applications," *IEEE Signal Processing Magazine*, vol. 36, pp. 59–80, 2019.

[15] D. Xu and D. P. Mandic, "The theory of quaternion matrix derivatives." *IEEE Trans. Sig. Proc.*, vol. 63, no. 6, pp. 1543–1556, 2015.

[16] D. Xu, Y. Xia, and D. P. Mandic, "Optimization in quaternion dynamic systems: gradient, hessian, and learning algorithms," *IEEE transactions on neural networks and learning systems*, vol. 27, no. 2, pp. 249–261, 2016.

[17] D. Xu, C. Jahanchahi, C. C. Took, and D. P. Mandic, "Enabling quaternion derivatives: the generalized hr calculus," *Royal Society open science*, vol. 2, no. 8, p. 150255, 2015.

[18] D. P. Mandic, C. Jahanchahi, and C. C. Took, "A quaternion gradient operator and its applications," *IEEE Signal Processing Letters*, vol. 18, no. 1, pp. 47–50, 2011.

[19] P. Paatero and U. Tapper, "Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values," *Environmetrics*, vol. 5, no. 2, pp. 111–126, 1994.

[20] M. W. Berry, M. Browne, A. N. Langville, V. P. Pauca, and R. J. Plemmons, "Algorithms and applications for approximate nonnegative matrix factorization," *Computational statistics & data analysis*, vol. 52, no. 1, pp. 155–173, 2007.