Quaternion Non-negative Matrix Factorization: definition, uniqueness and algorithm

Julien Flamant, Sebastian Miron, David Brie

Abstract—This article introduces the notion of quaternion non-negative matrix factorization (Q-NMF), which extends the usual non-negative matrix factorization (NMF) to the case of bivariate or polarized signals. The Q-NMF relies on two key ingredients: (i) the algebraic representation of polarization information thanks to quaternions and (ii) the exploitation of physical polarization constraints that generalize non-negativity. Uniqueness conditions for the Q-NMF are presented. The relationship between Q-NMF and NMF highlights the key disambiguating role played by polarization information. A simple yet efficient algorithm called quaternion alternating least squares (Q-ALS) is introduced to solve the Q-NMF problem in practice. Numerical experiments on synthetic data demonstrate the relevance of the approach, which appears very promising, notably for blind source separation problems arising in spectro-polarimetric imaging.

Index Terms—polarization, quaternion non-negative matrix factorization, Stokes parameters, blind source separation.

I. INTRODUCTION

Polarization information is essential to many fields ranging from seismology [1], optics [2] to gravitational astronomy [3], among others. Polarization encodes the geometry of wave oscillations. It thus carries many crucial morphological and physical insights about the medium which transmitted or reflected the polarized wave. The importance of polarization has been acknowledged for a long time in polarimetric imaging [4], [5], since it provides distinctive features of the observed scene — unaccessible to conventional intensity imaging. Polarization diversity features a natural disambiguating power, which can be exploited in the context of blind source separation [6].

Recent years have seen an increased interest in exploiting polarization diversity in hyperspectral imaging systems [7], [8]. This modality, called spectro-polarimetric imaging was first popularized in astronomy [9]. In full generality, it consists in acquiring, for a collection of wavelengths and locations, a 4-dimensional real vector gathering 4 Stokes parameters. These energetic parameters are widely used in optics [10] to describe the complete properties of light (intensity and polarization). The conjoint acquisition of spatial, spectral and polarization diversities then raises important and original challenges, in particular regarding blind polarized source problems.

Spectro-polarimetric data can be represented as a third order real-valued tensor, where the rows, columns and tube fibers correspond to frequency, space and polarization diversities, respectively. Note that, since polarization is described here by Stokes parameters, the tube fiber is of size 4. Solving the blind polarized source separation problem using usual low-rank approximation techniques raises two major issues. First, Stokes parameters being energetic quantities, they obey structural constraints [10] which need to be taken into account to guarantee the physical interpretation of the recovered rank-1 terms. Second, typical rank-1 decompositions such as the Canonical Polyadic Decomposition (CPD) [11] necessarily yield to rank-1 terms with constant polarization. It thus restricts the physical validity of usual tensor decompositions to specific settings, e.g. narrow band sources with constant polarization. The general case of wideband sources with frequency-dependent polarization, as observed in spectro-polarimetric data, requires the development of new low-rank approximation tools to deal with the specificities of polarization.

To this aim, this paper introduces the notion of quaternion non-negative matrix factorization (Q-NMF). It extends the concept of non-negative matrix factorization (NMF) to polarized signals. It relies on two key ingredients: (i) the algebraic representation of polarization information using quaternions and (ii) the exploitation of physical constraints on Stokes parameters, which generalize non-negativity to the case of polarized signals. Just like NMF offers a standard approach towards hyperspectral image unmixing [12], the Q-NMF establishes a general framework for the unmixing of spectro-polarimetric data.

Our contributions are threefold. First, we introduce the Q-NMF problem from a natural spectro-polarimetric mixing model and demonstrate how it generalizes the usual NMF. Second, we provide a study of uniqueness conditions for the Q-NMF problem that reveal the key disambiguating role played by polarization. In addition we give a sufficient condition for unicity in the 2 source case and a necessary condition for unicity in the general case. Finally, we provide a simple yet efficient algorithm in the quaternion domain that effectively solves the Q-NMF problem. We hope to convince the reader that quaternions, beyond offering an elegant way to handle polarization diversity, permit numerous theoretical (e.g. unicity conditions for the Q-NMF) and methodological (e.g. algorithmic design to solve the Q-NMF in practice) developments that would have been otherwise cumbersome to obtain.

This paper is organized as follows. Section II collects necessary material regarding quaternions, polarization and associated constraints. Section III defines the concept of quaternion non-negative matrix factorization (Q-NMF). Section IV studies uniqueness conditions for Q-NMF, starting with the simpler 2 sources case, followed by a discussion of the general case. Section V introduces the quaternion alternating least squares (Q-ALS) algorithm for solving the Q-NMF problem in practice. Section VI provides numerical evidence to the relevance of...
the approach, while Section VII presents concluding remarks. Appendices gather technical details and proofs.

II. Preliminaries

This paper relies on the algebraic treatment of polarization thanks to the quaternion formalism. We briefly review quaternions and polarization in Sections II-A and II-B respectively. Section II-C then describes the quaternion formalism for generic bivariate or polarized signals.

A. Quaternions

The set of quaternions $\mathbb{H}$ forms a 4-dimensional normed division algebra over the real numbers $\mathbb{R}$. Its canonical basis is $\{1, i, j, k\}$ where $i, j, k$ are imaginary units $i^2 = j^2 = k^2 = -1$ such that

$$ij = k, \quad ij = -ji, \quad ijk = -1 \quad (1)$$

Importantly, quaternion multiplication is noncommutative, i.e. for $q, q' \in \mathbb{H}$, $qq' \neq q'q$ in general. Any quaternion $q \in \mathbb{H}$ reads in Cartesian form

$$q = a + bi + cj + dk \quad (2)$$

where $a, b, c, d \in \mathbb{R}$. The real or scalar part of $q$ is $\text{Re}(q) = a$ and its imaginary or vector part is $\text{Im}(q) = bi + cj + dk$. When $\text{Re}(q) = 0$, $q$ is said to be pure (imaginary). The quaternion conjugate of $q$ is denoted by $\bar{q} = \text{Re}(q) - \text{Im}(q)$. The modulus of $q$ is $|q| = \sqrt{\bar{q}q} = \sqrt{\bar{q}q}$. More about quaternions can be found in dedicated textbooks, see e.g. [13].

A quaternion matrix $Q \in \mathbb{H}^{M \times N}$ has elements $Q_{mn} = q_{mn} \in \mathbb{H}$. Its conjugate-transpose is $\bar{Q}$ such that $(\bar{Q})_{mn} = \bar{q}_{mn}$. The conjugate-transpose of $Q$ is denoted by $Q^\dagger = \bar{Q}^\top$, i.e. such that $(Q^\dagger)_{mn} = \bar{q}_{nm}$. Note that usual matrix operations require special care due to non-commutativity of quaternion multiplication. For instance, given two matrices $Q_1 \in \mathbb{H}^{M \times N}$ and $Q_2 \in \mathbb{H}^{N \times P}$, one has $Q_1Q_2^\dagger \neq Q_2Q_1^\dagger$ and $(Q_1Q_2^\dagger)^\top \neq Q_2^\dagger Q_1^\top$ in general, whereas $(Q_1Q_2^\dagger)^\top = Q_1^\top Q_2^\dagger$ is always true. The reader is referred to [13], [14], [15], [16], [17], [18] and references therein for extensive results on quaternion matrix algebra.

B. Polarization and Stokes parameters

In optics [2], [19], [10], the polarization state of light is commonly described by four Stokes parameters $S_0, S_1, S_2, S_3$. These real-valued parameters are energetic quantities, and thus are experimentally measurable. The first Stokes parameter $S_0 \geq 0$ is classical and measures the total intensity of light, i.e. the sum of intensities from the polarized and unpolarized parts of light. The relative contribution of these two parts to the total intensity is ruled by the degree of polarization $\Phi$:

$$\Phi = \frac{\text{intensity of polarized part}}{\text{total intensity}} = \frac{\sqrt{S_1^2 + S_2^2 + S_3^2}}{S_0}. \quad (3)$$

By definition, $\Phi \in [0, 1]$. When $\Phi = 1$, light is said to be fully polarized, whereas for $\Phi = 0$ it is said unpolarized. For $\Phi \in (0, 1)$, light is said to be partially polarized.

The three remaining Stokes parameters $S_1, S_2, S_3$ encode the polarization properties of the polarized part of the light, as explained below. Fig. 1 depicts the Poincaré sphere of polarization states [20], a convenient graphical tool for interpreting Stokes parameters. Assuming $S_0 > 0$, normalized Stokes parameters $S_1/S_0, S_2/S_0, S_3/S_0$ encode Cartesian coordinates of a point inside the sphere of unit radius. This point is parameterized by its spherical coordinates $(\Phi, 2\theta, 2\chi)$, where $\theta \in [0, \pi]$ and $\chi \in [-\pi/4, \pi/4]$ are the orientation and ellipticity of the polarization ellipse, respectively. Thus to any point on the surface of the sphere correspond a unique (fully) polarization state. The North ($S_3/S_0 = 1$) and South ($S_3/S_0 = -1$) pole describe counter-clockwise and clockwise circular polarized light, respectively. The Equator ($S_3/S_0 = 0$) encodes linear polarization states with orientation $\theta$ of the ellipse major axis.

C. Quaternion power spectral density of polarized signals

Recently, a general framework for the analysis and filtering of bivariate signals has been proposed [21], [22], [23]. The proposed framework enables natural and straightforward generalizations of many concepts of signal processing to the case of bivariate signals, such as spectral densities, linear filters or time-frequency representations. It relies on two key ingredients: (i) the natural embedding of bivariate signals – viewed as complex-valued signals – into the set of quaternions $\mathbb{H}$ and (ii) the use of a dedicated quaternion Fourier transform (QFT) to provide a physically meaningful spectral representation of bivariate signals. In particular, unlike existing approaches, polarization features appear naturally in the quaternion framework. For the purposes of this paper, we only review here briefly the notion of quaternion PSD, introduced in [23]. A complete description of the quaternion formalism can be found in [21], [22], [23], [24].

A wide range of bivariate signals can be viewed as second-order stationary (simply referred to as stationary in the sequel) random signals. For such signals, the notion of quaternion power spectral density (PSD) can be rigourously defined [23].
It gives, in the spectral domain, the complete second-order structure of a stationary bivariate signal. The quaternion PSD is denoted by $w$ and directly reads in terms of frequency-dependent Stokes parameters \[ 4 \]

$$w(\nu) = S_0(\nu) + iS_1(\nu) + jS_2(\nu) + kS_3(\nu).$$  

This expression shows that Stokes parameters, originally introduced in the context of light polarization, are a natural reparameterization for the second-order properties of stationary bivariate signals. Moreover, as shown in [23], the quaternion PSD \[ 4 \] exhibits a symmetry between positive and negative frequencies. This ensures that only positive frequencies can be considered in \[ 4 \], a crucial step for physical interpretability.

Eq. \[ 4 \] can be rewritten for convenience as

$$w(\nu) = I(\nu) + I(\nu)\Phi(\nu)\mu(\nu)$$  

where at a given frequency $\nu$, $I(\nu) := S_0(\nu) \geq 0$ is the total intensity, $\Phi(\nu)$ the degree of polarization and $\mu(\nu)$ is the polarization axis – a pure unit quaternion such that $\mu(\nu)^2 = -1$. With this notation, the quantity $\Phi(\nu)\mu(\nu)$ is a pure quaternion that can be identified with a vector of $\mathbb{R}^3$ – which completely encodes the polarization state at frequency $\nu$, see Fig. [1]. The expression \[ 5 \] also emphasizes one of the benefits of the quaternion formalism: it permits a natural separation between pure energetic information – conveyed by $I(\nu)$, the real part of $w(\nu)$ – and the geometric vector information – conveyed by $I(\nu)\Phi(\nu)\mu(\nu)$, the imaginary part of $w(\nu)$.

D. Polarization constraint and quaternion non-negativity

The set of Stokes parameters does not span the entirety of $\mathbb{R}^4$. Instead to be physically admissible, a vector $(S_0, S_1, S_2, S_3)^T \in \mathbb{R}^4$ should satisfy the following constraints \[ 8 \]

$$S_0 \geq 0 \quad \text{and} \quad \sqrt{S_1^2 + S_2^2 + S_3^2} \leq S_0. \quad (S)$$  

Those two constraints bear a strong physical interpretation. The first one is classical and indicates that the total intensity $S_0$ is a non-negative real quantity. The second one means that the intensity of the polarization part cannot be larger than the total intensity. As a result, these constraints equivalently encode the range $0 \leq \Phi \leq 1$ of the degree of polarization \[ 4 \].

From a mathematical perspective, \[ 8 \] generalizes the usual non-negativity constraint to the case of bivariate signals. To see this, consider the polarization coherency matrix $J$ [10, Section 1.4]

$$J = \frac{1}{2} \begin{bmatrix} S_0 + S_1 & S_2 + iS_3 \\ S_2 - iS_3 & S_0 - S_1 \end{bmatrix} \in \mathbb{C}^{2 \times 2}. \quad (6)$$  

The matrix $J$ can be statistically defined as a covariance matrix [10, Section 1.4.1]. It is thus Hermitian positive semi-definite – one also says that $J$ is Hermitian non-negative. It is then easy to check that constraints \[ 8 \] directly follow from imposing non-negativity on the Hermitian matrix $J$ i.e.

$$\text{(S)} \Leftrightarrow \begin{cases} \text{tr} \ J & \geq 0 \\ \text{det} \ J & \geq 0 \end{cases}. \quad (7)$$

This relation between the admissible Stokes vectors and Hermitian non-negative 2-by-2 matrix will prove to be useful later on for imposing \[ 8 \] in practice, see Section \[ 3 \].

The core of the proposed approach is the quaternion representation \[ 4 \] of the Stokes parameters. Thus, we define by extension the set of non-negative quaternions $\mathbb{H}_S \subseteq \mathbb{H}$ such that

$$\mathbb{H}_S := \{ q \in \mathbb{H} | \text{Re} q \geq 0 \text{ and } |\text{Im} q| \leq \text{Re} q \}.$$

Clearly, any $q \in \mathbb{H}_S$ corresponds to a unique Stokes vector satisfying \[ 8 \]. This formal mapping will be used throughout this paper.

III. Quaternion Non-negative Matrix Factorization

This Section introduces Quaternion Non-negative Matrix Factorization (Q-NMF), a new tool which generalizes the notion of Non-negative Matrix Factorization (NMF) to the case of polarized signals. It exploits two key features of such signals: (i) the polarization constraint \[ 8 \], which extends the notion of non-negativity for bivariate signals and (ii) the algebraic representation \[ 4 \] of Stokes vectors using quaternions. The Q-NMF establishes a new and generic tool for low-rank approximations of polarized signals, with many potential applications in source separation or data completion.

A. Definition

Without loss of generality, let us consider the setting of a typical spectro-polarimetric experiment, where for each sampled frequency (or wavelength) $(\nu_m)_{1 \leq m \leq M}$ and location $(u_n)_{1 \leq n \leq N}$ the four Stokes parameters $S_i(\nu_m, u_n)$, $i = 1, 2, 3, 4$ are acquired. Using \[ 8 \], results of such experiment can be written in quaternion form as

$$x(\nu_m, u_n) = S_0(\nu_m, u_n) + iS_1(\nu_m, u_n) + jS_2(\nu_m, u_n) + kS_3(\nu_m, u_n) \in \mathbb{H}_S$$

for $1 \leq m \leq M$, $1 \leq n \leq N$. These, a sensible approach is to consider that the data $x$ results from a superposition of $P$ elementary polarized sources such that, for every frequency $\nu_m$ and location $u_n$,

$$x(\nu_m, u_n) = \sum_{p=1}^{P} w_p(\nu_m)h_p(u_n)$$

For a given source $p$, $w_p(\nu_m) \in \mathbb{H}_S$ encodes its four frequency-dependent Stokes parameters and $h_p(\cdot) \geq 0$ denotes the corresponding spatial activation coefficients.

Rewriting \[ 8 \] into matrix form introduces Quaternion Non-Negative Matrix Factorization (Q-NMF) like

$$X = WH$$

where $X \in \mathbb{H}_S^{M \times N}$ is the data matrix with coefficients $(X)_{mn} = x(\nu_m, u_n)$. The matrix $W \in \mathbb{H}_S^{M \times P}$ is the source matrix with coefficients $(W)_{mp} = w_m(\nu_m)$, whereas $H \in \mathbb{R}_+^{P \times N}$ is called the activation matrix such that $(H)_{pn} = h_p(u_n)$. 

\[ 8 \]
B. Relation with NMF

The Q-NMF extends the well-known NMF problem to the case of bivariate or polarized signals. Compared to the NMF, the Q-NMF features a quaternion-valued sources factor $W$ which exploits the polarization constraint (5) instead of the usual non-negativity constraint. Thanks to quaternion algebra, the Q-NMF (11) can also be rewritten as

$$X = WH \iff \begin{cases} \text{Re } X = [\text{Re } W]H \quad \text{(NMF)} \\ \text{Im } X = [\text{Im } W]H \quad \text{(polarization)} \end{cases}$$

Eq. (11) shows that the Q-NMF can be seen as a co-factorization problem with common activation factor $H$. The first factorization problem is an usual NMF on the real part of $X$, i.e. on intensity data (Stokes parameter $S_0$ ≥ 0) only. The second one corresponds to a factorization problem on the imaginary part of $X$ encoding polarization properties (Stokes parameters $S_1, S_2, S_3$). These two factorization problems are not independent for two reasons: (i) the activation factor $H$ appears in both and (ii) for each coefficient $(m, p)$ of the source factor $W$, the constraint (5) connects the modulus of the imaginary part (polarization factorization problem) to its real part (NMF problem).

The relationship (11) provides another illustration of how Q-NMF extends NMF to account for polarization diversity. It allows a precise quantification of the role played by the polarization information and its associated constraint (5). For each coefficient $m, p$, the constraint (15) enables for the NMF case [25], [26], [27], [28]. The close relationship (11), just as in the standard NMF case.

IV. Unicity of Q-NMF

This Section deals with a fundamental question: upon which conditions on the source $W$ and activation $H$ factors is the Q-NMF $X = WH$ unique? The uniqueness question is central in any factorization or decompositions problems, in particular for the NMF case [25], [26], [27], [28]. The close relationship between NMF and Q-NMF detailed in Section III-B enables a precise quantification of the disambiguiting role played by the polarization information.

Let $X \in \mathbb{H}^{M \times N}_S$ and suppose that there exists two matrices $W \in \mathbb{H}^{M \times P}_S$ and $H \in \mathbb{R}^{P \times N}_+$ such that $X = WH$ holds. Given a nonsingular matrix $T \in \mathbb{H}^{P \times P}$, any pair $(\tilde{W}, \tilde{H})$ defined as

$$\tilde{W} = WT \quad (13)$$
$$\tilde{H} = T^{-1}H \quad (14)$$

leaves the data matrix unchanged $X = WH = \tilde{W} \tilde{H}$. However, to be admissible the linear transformation $T$ should yield matrices $\tilde{W}$ and $\tilde{H}$ such that

$$\tilde{W} \in \mathbb{H}^{M \times P}_S \text{ and } \tilde{H} \in \mathbb{R}^{P \times N}_+. \quad (15)$$

The real-valuedness constraint on $\tilde{H}$ directly imposes that $T$ is a real matrix. Indeed, if the entries $T$ are quaternion-valued, so are the entries of $T^{-1}$, and thus $H = T^{-1}H$ is also quaternion-valued. As a result, only linear transformations $T \in \mathbb{R}^{P \times P}$ may yield alternative factors $\tilde{W}$ and $\tilde{H}$ satisfying (15), just as in the standard NMF case.

The Q-NMF then exhibits the same trivial ambiguities as the NMF, as stated by Proposition 1 below.

**Proposition 1.** The Q-NMF $X = WH$ exhibits two intrinsic ambiguities, namely,

- **scale indeterminacy:**

$$T = \text{diag} \left( t_1, t_2, \ldots, t_P \right), t_i > 0, 1 \leq i \leq P \quad (16)$$

- **order indeterminacy:**

$$T \text{ is a permutation matrix} \quad (17)$$

for which $\tilde{W} = WT$ and $\tilde{H} = T^{-1}H$ define equally valid Q-NMF factors.

**Proof.** Start by the scale indeterminacy. Let $t_i > 0, 1 \leq i \leq P$ and define $T = \text{diag} \left( t_1, t_2, \ldots, t_P \right)$. One has $T^{-1} = \text{diag} \left( t_1^{-1}, t_2^{-1}, \ldots, t_P^{-1} \right)$ a strictly positive diagonal matrix. As $\mathbb{H}_S$ and $\mathbb{R}^+$ are stable under positive scalings, then $\tilde{W} = TW$ and $\tilde{H} = T^{-1}H$ define valid factors such that (15) holds. For the order indeterminacy, suppose that $T$ is a permutation matrix. This implies that its inverse $T^{-1}$ is also a permutation matrix, and that entries of $T$ and $T^{-1}$ are either 1 or 0. It directly follows that $\tilde{W} = TW$ and $\tilde{H} = T^{-1}H$ define valid factors such that (15) holds.

These inevitable indeterminacies motivate the following definition for the identifiability of the Q-NMF.

**Definition 1.** For a given data matrix $X$, the Q-NMF $X = WH$ is said to be essentially unique if the only indeterminacies are the scale indeterminacy (16) and the scale indeterminacy (17).

The Q-NMF features the same trivial ambiguities as the standard NMF, and thus can be handled through standard techniques, e.g. by imposing normalization on the sources coefficients to get rid of the scale indeterminacy. The remaining of this Section, we explore in detail how the polarization constraint (5) for $W$ and the non-negativity constraint on $H$ impact Q-NMF identifiability.

A. Two sources case: range of admissible solutions and sufficient unicity condition

We first consider the simpler case of two sources ($P = 2$), which is particularly instrumental in the understanding of the impact of the polarization constraint (5). The approach follows closely the one presented in [25] for the NMF case.

For $1 \leq m, n \leq M$ and $1 \leq p, n \leq N$, corresponding entries of each factors $W, H$ are given by

$$(W)_{mp} = w_{mp} \in \mathbb{H}_S \text{ and } (H)_{pn} = h_{pn} \geq 0 \quad (18)$$

where $p = 1, 2$ denote the source index. We use the convenient quaternion form (5) for $w_{mp}$, i.e.

$$w_{mp} = I_{mp} + I_{mp}\Phi_{mp} u_{mp} \quad (19)$$

Model indeterminacies are handled through the explicit form of the matrix $T$:

$$T(\alpha, \beta) = \begin{bmatrix} 1 - \alpha & \alpha \\ \beta & 1 - \beta \end{bmatrix} \quad (20)$$
with inverse
\[ T^{-1}(\alpha, \beta) = \frac{1}{1 - \alpha - \beta} \begin{bmatrix} 1 - \beta & -\alpha \\ -\beta & 1 - \alpha \end{bmatrix}. \] (21)
We assume that \( \alpha + \beta < 1 \) so that \( T \) is invertible. This also ensures to get rid of the order indeterminacy [25].

Now consider the linear transformed factors \( \tilde{W} = T(\alpha, \beta)W \) and \( \tilde{H} = T^{-1}(\alpha, \beta)H \). Using (20), the mixing of sources explicitly reads for \( 1 \leq m \leq M \)
\[ \tilde{w}_{m1} = (1 - \alpha)w_{m1} + \alpha w_{m2} \quad (22) \]
\[ \tilde{w}_{m2} = \beta w_{m1} + (1 - \beta)w_{m2} \quad (23) \]
With (21) one gets the new activations for \( 1 \leq n \leq N \)
\[ \tilde{h}_{1n} = \frac{(1 - \alpha)h_{1n} + \alpha h_{2n}}{1 - \alpha - \beta} \quad (24) \]
\[ \tilde{h}_{2n} = \frac{\beta h_{1n} + (1 - \beta)h_{2n}}{1 - \alpha - \beta} \quad (25) \]
Non-negativity of \( \tilde{H} \) imposes that \( \tilde{h}_{1n} \geq 0 \) and \( \tilde{h}_{2n} \geq 0 \) for every \( n \), just like in the standard NMF case [25]. The polarization constraints [5] then impose that \( \tilde{w}_{m1}, \tilde{w}_{m2} \in \mathbb{R}_+ \). In other words, for every \( m \)
\[ \begin{cases} \text{(NMF)} & \Re \tilde{w}_{m1} \geq 0 \\ \Re \tilde{w}_{m2} \geq 0 \end{cases} \]
\[ \begin{cases} \text{(polarization)} & |\Im \tilde{w}_{m1}| \leq \Re \tilde{w}_{m1} \\ |\Im \tilde{w}_{m2}| \leq \Re \tilde{w}_{m2} \end{cases} \quad (26) \quad (27) \]
Eq. (26) corresponds also to the usual NMF case [25]. However, condition (27) reveals the specicity of Q-NMF.

In order to find the range of admissible solutions for the Q-NMF, we search for intervals \( T_{\alpha}^{Q-\text{NMF}} \) and \( T_{\beta}^{Q-\text{NMF}} \) such that, for every \( (\alpha, \beta) \in T_{\alpha}^{Q-\text{NMF}} \times T_{\beta}^{Q-\text{NMF}} \) the linear transform \( T(\alpha, \beta) \) yields valid \( \tilde{W} \) and \( \tilde{H} \) factors. Remark that, since Eqs. (24) - (26) are identical to the NMF case, we can express the intervals \( T_{\alpha}^{Q-\text{NMF}} \) and \( T_{\beta}^{Q-\text{NMF}} \) as
\[ T_{\alpha}^{Q-\text{NMF}} = \bigcap_m T_{\alpha,m}^{Q-\text{pol}} \quad (28) \]
\[ T_{\beta}^{Q-\text{NMF}} = \bigcap_m T_{\beta,m}^{Q-\text{pol}} \quad (29) \]
where \( T_{\alpha}^{Q-\text{NMF}} \) (resp. \( T_{\beta}^{Q-\text{NMF}} \)) is the interval obtained for \( \alpha \) (resp. \( \beta \)) using non-negativity constraints only, as explained below. For every \( m \), \( T_{\alpha,m}^{Q-\text{pol}} \) and \( T_{\beta,m}^{Q-\text{pol}} \) indicate intervals for \( \alpha \) and \( \beta \) such that the polarization constraint (27) holds. Thanks to this rewriting, Eqs. (28)-(29) permit to separate the contributions of the non-negativity and polarization constraints, respectively. This allows to precisely quantify the role played by polarization constraint in Q-NMF to improve NMF identifiability.

**Computation of \( T_{\alpha}^{Q-\text{NMF}} \) and \( T_{\beta}^{Q-\text{NMF}} \)** These intervals encode 2 non-negativity conditions: the real part of the sources and of the activations coefficients. Their expression follow directly from NMF results [25], that is:
\[ T_{\alpha}^{Q-\text{NMF}} = [\alpha_{\text{min}}^{Q-\text{NMF}}, \alpha_{\text{max}}^{Q-\text{NMF}}] \quad \text{and} \quad T_{\beta}^{Q-\text{NMF}} = [\beta_{\text{min}}^{Q-\text{NMF}}, \beta_{\text{max}}^{Q-\text{NMF}}] \quad (30) \]
Lower bounds depend on the ratio of the real parts of the sources like
\[ \alpha_{\text{min}}^{\text{NMF}} = \min_{m \in M_1} \frac{I_{m1}}{I_{m2} - I_{m1}}, \quad \beta_{\text{min}}^{\text{NMF}} = \min_{m \in M_2} \frac{I_{m2}}{I_{m1} - I_{m2}} \quad (31) \]
where \( M_1 = \{ m | I_{m2} > I_{m1} \} \), \( M_2 = \{ m | I_{m1} > I_{m2} \} \) and for convenience \( \Re w_{mp} = I_{mp} \) as in (19). Activations coefficients control the upper bounds of \( T_{\alpha}^{Q-\text{NMF}} \) and \( T_{\beta}^{Q-\text{NMF}} \):
\[ \alpha_{\text{max}}^{\text{NMF}} = \max_{m} \frac{h_{1n}}{h_{1n} + h_{2n}}, \quad \beta_{\text{max}}^{\text{NMF}} = \max_{m} \frac{h_{2n}}{h_{1n} + h_{2n}} \quad (32) \]

**Computation of \( T_{\text{pol}}^{Q-\text{NMF}} \)** Let us now exploit the polarization condition (27), on which relies the Q-NMF. Fix \( 1 \leq m \leq M \) and start with parameter \( \alpha \). Plugging (19) into (22) ones gets
\[ \Re \tilde{w}_{m1} = (1 - \alpha)I_{m1} + \alpha I_{m2} \quad (33) \]
\[ \Im \tilde{w}_{m1} = (1 - \alpha)I_{m2} \Phi_{m1} \mu_{m1} + \alpha I_{m2} \Phi_{m2} \mu_{m2} \quad (34) \]
Remark that the polarization constraint (27) is equivalent to the condition \( |\Im \tilde{w}_{m1}|^2 \leq (\Re \tilde{w}_{m1})^2 \). Thus, developing all terms and reorganizing the expression as a second-order polynomial in \( \alpha \), we get the following inequality
\[ \alpha^2 \left[ I_{m1}^2 (1 - \Phi_{m1}^2) + I_{m2}^2 (1 - \Phi_{m2}^2) - 2I_{m1}I_{m2} (1 - \Phi_{m1} \Phi_{m2}) \langle \mu_{m1}, \mu_{m2} \rangle \right] \]
\[ + [20 (I_{m1} I_{m2} (1 - \Phi_{m1} \Phi_{m2}) \langle \mu_{m1}, \mu_{m2} \rangle) - I_{m1}^2 (1 - \Phi_{m1}^2)] + I_{m2}^2 (1 - \Phi_{m2}^2) \quad (35) \]
where \( \langle \mu_1, \mu_2 \rangle = -\Re(\mu_1 \mu_2) \) can be identified with the usual inner product of \( \mathbb{R}^3 \). Starting from (23), the same approach is used for \( \beta \), leading to the following second-order polynomial inequality
\[ \beta^2 \left[ I_{m2}^2 (1 - \Phi_{m2}^2) + I_{m1}^2 (1 - \Phi_{m1}^2) - 2I_{m1}I_{m2} (1 - \Phi_{m1} \Phi_{m2}) \langle \mu_{m1}, \mu_{m2} \rangle \right] \]
\[ + [20 (I_{m1} I_{m2} (1 - \Phi_{m1} \Phi_{m2}) \langle \mu_{m1}, \mu_{m2} \rangle) - I_{m1}^2 (1 - \Phi_{m1}^2)] + I_{m2}^2 (1 - \Phi_{m2}^2) \quad (36) \]
For a given \( m \), intervals \( T_{\text{pol}}^{Q-\text{pol}} \) and \( T_{\text{pol}}^{Q-\text{pol}} \) are obtained by solving the corresponding polynomial inequalities (35) and (36). It involves finding the roots of the associated second-order polynomials in \( \alpha \) and \( \beta \), respectively. An easy but lengthy computation of polynomial discriminants shows that (35) and (36) always admit two real-valued solutions (possibly degenerate). Unfortunately, the amount of parameters involved (at least 4: \( \Phi_{m1}, \Phi_{m2}, (\mu_{m1}, \mu_{m2}) \)) and the ratio \( I_{m1}/I_{m2} \) prevents from performing a general theoretical study of the roots behaviour. Such a study would require the treatment of numerous particular cases, notably due to the fact that the sign of the second-order term is not constant and can even cancel out for some values of parameters. Nonetheless, in practice where values of sources parameters are given, Eqs. (35) and (36) can be solved very efficiently numerically to yield desired intervals \( T_{\text{pol}}^{Q-\text{pol}} \) and \( T_{\text{pol}}^{Q-\text{pol}} \). See Section VI for numerical illustrations on synthetic and real-world data.

**Sufficient uniqueness condition** Despite the apparent complexity of the polarization conditions (35) and (36), a simple
and interpretable sufficient condition for the unicity of Q-NMF with \( P = 2 \) can be formulated.

**Proposition 2.** If the following conditions are satisfied:

- \( \exists m_1, m_2 \in \{1, 2, \ldots, M\} \) s.t.
  \[
  \begin{cases}
  \Phi_{m_1} = 1, \Phi_{m_2} \mu_{m_2} \neq \mu_{m_1} \\
  I_{m_1} \geq \frac{1}{2} I_{m_2} - \Phi_{m_2}^2 \mu_{m_2} \\
  \Phi_{k_2} = 1, \Phi_{m_2} \mu_{m_2} \neq \mu_{m_2} \\
  I_{m_2} \geq \frac{1}{2} I_{m_2} - \Phi_{m_2}^2 \mu_{m_2} 
  \end{cases}
  \quad (C1)
  \]

- \( \exists n_1, n_2 \in \{1, 2, \ldots, N\}, n_1 \neq n_2 \) s.t.
  \[
  \begin{cases}
  h_{1n_1} > 0 \text{ and } h_{2n_1} = 0 \\
  h_{2n_2} > 0 \text{ and } h_{1n_2} = 0
  \end{cases}
  \quad (C2)
  \]

then the Q-NMF \( X = WH \) is essentially unique.

**Proof.** See Appendix A-A

On the one hand, condition (C2) is identical to the one found for the activation factor in the standard NMF case [25]. On the other hand, condition (C1) illustrates the key role played by polarization information. Compared to the usual NMF sufficient conditions for the 2 sources case [25], it does not require each source to vanish alternatively. Condition (C1) shows that it is sufficient that there exists two indices \( m_1, m_2 \) such that the first source is fully polarized at \( m_1 \) (and the second one exhibits a different polarization state at \( m_1 \)) and the second one is fully polarized at \( m_2 \) (and the first one exhibits a different polarization state at \( m_2 \)), and that respective intensities of each source should be larger than the other one by a factor in \([0,1]\).

Note that (C1) does not require at all \( m_1 \neq m_2 \). In fact, when \( m_1 = m_2 = m \), it becomes (CS1):

\[
\begin{cases}
  \Phi_{m_1} = \Phi_{m_2} = 1, \mu_{m_1} \neq \mu_{m_2} \\
  I_{m_1} > 0, I_{m_1} > 0
  \end{cases}
  \quad (CS1')
  \]

In other words, if there exists \( m \) such that both sources exhibit some arbitrary energy and are fully polarized with different polarization axes, then the Q-NMF \( X = WH \) is essentially unique.

The sufficient condition given in Proposition 2 for the uniqueness of Q-NMF in the two source case appears remarkably broad compared to the NMF case [25]. By incorporating polarization information and its associated constraints, the Q-NMF permits to recover unicity even when the sources never vanish – a typical case where NMF is known to be non-identifiable [28], [25].

**B. General case \((P \geq 2)\)**

The study of the unicity of the Q-NMF \( X = WH \) for an arbitrary number of \( P \) sources is much more cumbersome than the \( P = 2 \) sources case since the Q-NMF problem can be shown to be NP-hard. However, the relationship between Q-NMF and NMF (see Section III-B) allows to illustrate the role played by polarization information in cases where standard NMF would fail.

Consider the case where sources intensities never vanish, i.e. \( \text{Re } w_{mp} > 0 \) for every \( m, p \). In that case, the associated NMF problem \( \text{Re } X = [\text{Re } W]H \) is known [25] to be non-unique. Proposition 3 shows gives a necessary condition for uniqueness of the Q-NMF, which is far less restrictive than the corresponding NMF one.

**Proposition 3 (Necessary condition for unicity).** Suppose that the Q-NMF \( X = WH \) is essentially unique. Then the following conditions are satisfied:

- \( \forall(p, q), p \neq q, \)
  \[
  \exists m \text{ s.t. } \Phi_{mp} = 1, \Phi_{mq} \mu_{mq} \neq \mu_{mp} \quad (A1)
  \]

- \( \forall(p, q), p \neq q, \)
  \[
  \exists n \text{ s.t. } h_{pn} = 0 \text{ and } h_{qn} > 0 \quad (A2)
  \]

**Proof.** See Appendix A-B

Proposition 3 shows that a necessary condition for unicity of the Q-NMF \( X = WH \) with non-vanishing sources is that, for any distinct pair of sources \((p, q)\), two criteria are satisfied: (A1) there exists an index \( m \) such that the source \( p \) is fully polarized and the source \( q \) exhibits a different polarization state (it can be totally polarized but with a different polarization axis); (A2) there is an index \( n \) such that the source \( q \) is active while the source \( p \) is not. Finally, note that condition (A2) on the activation factor is exactly the same as in the NMF case [25].

**V. AN ALGORITHM FOR Q-NMF**

This Section deals with the practical resolution of the Q-NMF \( X = WH \) problem. For a given value \( P \leq M, N \), its resolution can be seen as an optimization procedure,

\[
\min_{W \in \mathbb{H}^{R \times P} \times \mathbb{H}^{M \times N}} D(X, WH)
\]

where \( D : \mathbb{H}^{N \times M} \times \mathbb{H}^{N \times M} \rightarrow \mathbb{R}_+ \) can be an arbitrary cost function for quaternion matrices. For simplicity, we choose here the Frobenius distance between quaternion matrices \( D(X, WH) = \|X - WH\|_F^2 \), hereafter denoted Euclidean cost. The generic QMF optimization problem (37) becomes

\[
\min_{W \in \mathbb{H}^{R \times P \times \mathbb{H} \times (M \times N)}} D(X, WH) \quad (38)
\]

The formulation of the resolution of the Q-NMF problem (37) appears much alike the standard NMF with Euclidean cost. However, two fundamental questions need to be answered: (i) is the constraint (3) easy to implement? and (ii) can we optimize w.r.t. \( W \) directly in the quaternion domain?

Fortunately, answers to these two questions are affirmative. For (i), the answer relies on the key link between (3) and the set of non-negative Hermitian 2-by-2 matrices. A positive answer to (ii) is made possible by the recent advent of the theory of quaternion-domain derivatives [29], [30], [31], [32] – the so-called generalized \( \overline{\mathbb{H}} \overline{\mathbb{R}} \) calculus.
Section V-A below explains how constraints on \( W \) and \( H \) factors are implemented. Section V-B describes the proposed alternating least square strategy to solve (38). For completeness, Appendix F-A provides a quick introduction to quaternion optimization. Detailed computation of Q-NMF updates are found in Appendices B-C and B-B.

A. Projections onto constraints sets

Projections onto the constraints sets of \( W \) and \( H \) is a cornerstone of any method attempting to solve numerically the Q-NMF factorization problem (37) or (38). Projection operators onto respective constraints \( \mathbb{S} \) and \( \mathbb{R}^+ \) are denoted by \( \Pi_{\mathbb{R}^+} \) and \( \Pi_{\mathbb{S}} \), respectively.

Start by \( \Pi_{\mathbb{R}^+} \). This projection is classical in the NMF literature [33], [34], and thus reads for an arbitrary matrix \( M \in \mathbb{R}^{P \times N} \)

\[
[\Pi_{\mathbb{R}^+}(M)]_{pn} = \max(0,(M)_{pn}) \tag{39}
\]

where \( 1 \leq p \leq P, 1 \leq n \leq N \).

To compute \( \Pi_{\mathbb{S}} \), one uses the tight relationship between \( \mathbb{S} \) and the set of 2-by-2 Hermitian non-negative matrices – as described in Section V-A. Consider an arbitrary matrix \( M \in \mathbb{H}^{M \times P} \) with \((m, p)\) coefficient \((M)_{mp} = M_{mp} = \text{Re} M_{mp} + i \text{Im}_m M_{mp} + j \text{Im}_p M_{mp} + k \text{Im}_k M_{mp} \). Then define the following bijective mapping \( f : \mathbb{H} \rightarrow \mathbb{C}^{2 \times 2} \) such that

\[
f(M_{mp}) = \begin{bmatrix}
\text{Re} M_{mp} + i \text{Im}_j M_{mp} & \text{Im}_k M_{mp} + i \text{Im}_l M_{mp} \\
\text{Im}_k M_{mp} - i \text{Im}_l M_{mp} & \text{Re} M_{mp} - i \text{Im}_j M_{mp}
\end{bmatrix} \tag{40}
\]

By construction, \( f(M_{mp}) \) is an Hermitian matrix. Its projection onto the set of non-negative Hermitian matrices \( \mathbb{K}_+^2 \) is well-known [35, p. 399], i.e.

\[
\Pi_{\mathbb{K}_+^2} f(M_{mp}) = \sum_{i=1}^{2} \max(0, \eta_i) v_i v_i^\dagger \tag{41}
\]

where \( \eta_i \) and \( v_i \) are the \( i \)-th eigenvalue and eigenvector of the matrix \( f(M_{mp}) \), respectively. As a result, for an arbitrary quaternion matrix \( M \), its projection \( \Pi_{\mathbb{K}_+^2} \) reads

\[
[\Pi_{\mathbb{K}_+^2}(M)]_{mp} = f^{-1} \left\{ \Pi_{\mathbb{K}_+^2} \left[ f(M_{mp}) \right] \right\} \tag{42}
\]

where \( f^{-1} \) denotes the inverse mapping corresponding to (40).

B. Proposed algorithm: Quaternion Alternating Least Squares (Q-ALS)

The proposed algorithm adopts a popular strategy for solving the Q-NMF problem, based on the alternative constrained minimization of (35) w.r.t. \( H \) and \( W \). This choice is motivated by the fact that, whereas the Euclidean cost (38) is not convex in both \( W \) and \( H \), it is convex in either variable. After initialization of each factors, the iteration \( r > 0 \) reads

\[
H_{r+1} \leftarrow \arg\min_{H \in \mathbb{R}^+} \|X - WH_rH\|_F^2 \tag{43}
\]

\[
W_{r+1} \leftarrow \arg\min_{W \in \mathbb{H}_S} \|X - WH_{r+1}\|_F^2 \tag{44}
\]

However, constraints in (43)-(44) makes it difficult to obtain these updates directly. Instead, at a given iteration \( r \), for each factor one can solve the unconstrained least-squares problem and project the obtained solution onto the corresponding constraint:

\[
H_{r+1} \leftarrow \Pi_{\mathbb{R}^+} \left[ \arg\min_{H} \|X - WH_rH\|_F^2 \right] \tag{45}
\]

\[
W_{r+1} \leftarrow \Pi_{\mathbb{H}_S} \left[ \arg\min_{W} \|X - WH_{r+1}\|_F^2 \right] \tag{46}
\]

Due to its resemblance with the usual alternating least squares (ALS) algorithm [36] for the NMF, we call this strategy quaternion alternating least squares (Q-ALS). Derivation of explicit updates for unconstrained least-squares problems requires special care due to the quaternion nature of \( X \) and \( W \). It involves the notion of quaternion derivative as well as cautious handling of quaternion non-commutivity, see detailed computations provided in Appendices B-B and Appendix B-C.

As a result, one gets the explicit updates:

\[
H_{r+1} \leftarrow \Pi_{\mathbb{R}^+} \left[ \left( \text{Re}[W_r^\top W_r] \right)^{-1} \text{Re}[W_r^\top X] \right] \tag{47}
\]

\[
W_{r+1} \leftarrow \Pi_{\mathbb{H}_S} \left[ XH_{r+1} \left( H_{r+1}^\top H_{r+1} \right)^{-1} \right] \tag{48}
\]

Projections onto constraints sets are carried out as described in Section V-A.

The proposed algorithm is remarkably simple and cheap. Performing optimization directly in the quaternion-domain yields updates expressions that are very much alike those of the standard ALS algorithm for the NMF [37]. However, this apparent simplicity should not conceal the important underlying technical details presented in Appendix B. Such computations would not have been amenable without using the powerful theory of quaternion derivatives introduced recently [29], [30], [31], [32]. Finally, despite the lack of theoretical guarantees on its convergence, the quaternion ALS algorithm provides a good baseline for the resolution of the Q-NMF with reasonably good results in many situations.

VI. NUMERICAL EXPERIMENTS

This Section illustrates the relevance of the proposed approach by performing numerical experiments on synthetic data. Section VI-A provides numerical evidence to the key role played by polarization properties onto the range of admissible solutions for the 2 sources case. Section VI-B demonstrates the effectiveness of the Q-ALS algorithm to solve the Q-NMF problem.

A. Polarization information and identifiability

This Section illustrates the theoretical study provided in Section V-A regarding the identifiability of the Q-NMF in the \( P = 2 \) sources case. To emphasize how polarization impacts the range of admissible solutions, we choose a setting where the Q-NMF \( X = WH \) is not unique – and so is the associated NMF \( \text{Re} X = \text{Re} [W] H \), see [28].

Both sources and associated activations are non-vanishing, i.e. for all \( m, n \), one has \( w_{m1}, w_{m2}, h_{n1}, h_{n2} \neq 0 \). Intensity responses \( \text{Re} w_{m1}, \text{Re} w_{m2} \) encode realistic \( S_0 \) parameters,
obtained for instance in hyperspectral imaging of wood components [38]. To simplify, both sources are chosen to exhibit constant polarization properties:

\[ \Phi_{m1} = 0.7, \quad \mu_{m1} = 0.87i - 0.25j - 0.43k \]
\[ \Phi_{m2} = 0.5, \quad \mu_{m2} = -0.71i + 0.44j + 0.55k \]

(49) (50)

Each source being partially polarized, it ensures that the solution to the Q-NMF is not unique, since it does not fulfill the necessary conditions stated by Proposition 3.

Numerical computations yield the range of admissible solutions for the Q-NMF and its associated NMF problem. Following Section IV-A, the range of admissible parameters \( \alpha \) and \( \beta \) defining the transformation matrix \( T(\alpha, \beta) \) are found for the Q-NMF

\[ T^\text{Q-NMF}_\alpha = [-1.494 \cdot 10^{-1}, 7.692 \cdot 10^{-2}] \]
\[ T^\text{Q-NMF}_\beta = [-3.719 \cdot 10^{-1}, 1.538 \cdot 10^{-1}] \]

(51)

and for the associated NMF problem:

\[ T^\text{NMF}_\alpha = [-1.799, 7.692 \cdot 10^{-2}] \]
\[ T^\text{NMF}_\beta = [-1.303 \cdot 10^1, 1.538 \cdot 10^{-1}] \]

(52)

Note that the Q-NMF and NMF intervals share the same upper bounds on \( \alpha \) and \( \beta \), which arise from the activation factor ratios \( \langle 32 \rangle \).

Fig. 2 represents the set of admissible solutions corresponding to these NMF and Q-NMF intervals, respectively. Thick lines indicate the original sources and activations factors. Comparing respective ranges of solutions for \( S_0 \), it appears that the Q-NMF significantly improves identifiability over the standard NMF, by taking advantage of polarization information. Improvements are found on both \( S_0 \) and activations factors, and importantly, the Q-NMF permits to reconstruct Stokes parameters \( S_1, S_2, S_3 \) with limited uncertainty. These results illustrate how the Q-NMF takes advantage of the strong discriminative power of polarization in order to improve model identifiability.

B. Numerical reconstruction using Q-ALS

We provide in this Section a first numerical validation of the Q-ALS algorithm in a noise-free setting, and leave a more detailed study of algorithmic performance for future research. We consider the case of \( P = 3 \) sources, with \( M = 64 \)
Fig. 3. Recovery of sources and activations factors using the Q-ALS algorithm in the case \( P = 3 \) sources. Traces displayed were obtained for \( K = 50 \) random initializations.

and \( N = 128 \) wavelength and spatial samples, respectively. True sources and activation factors are constructed such that they satisfy the necessary conditions for Q-NMF identifiability stated in Proposition 5. In particular, there exists at least 3 wavelength indices such that each source is fully polarized whereas the two remaining ones exhibit a different polarization state. They are then combined to form the data matrix \( X = WH \).

The Q-ALS algorithm is initialized at random. More precisely, the initial activation factor \( H_0 \) is chosen as a matrix of i.i.d. entries drawn from the \( U([0,1]) \), the uniform distribution on \([0,1]\). The initial source factor \( W_0 \) has i.i.d. entries drawn from the quaternion circular unit Gaussian distribution [39] projected onto the constraint (S). Whereas this choice of initialization strategy proved here to be sufficient for our purpose, it might not be optimal with respect to algorithmic performances. Future work will focus on this important topic in more detail.

Fig. 3 depicts sources and activations factors obtained for \( K = 50 \) independent initializations of the Q-ALS algorithm. It took on average 70 iterations for the algorithm to converge, where convergence was assessed by monitoring the relative error \( \varepsilon_r = \|X - WH_r\|_F^2 / \|X\|_F^2 \) and stopping whenever the improvement was below a given threshold, i.e. \( |\varepsilon_r - \varepsilon_{r-1}| \leq 10^{-5} \) in our case. One can remark that, despite fulfilling necessary conditions of Proposition 3, recovered source and activations factors differ from one initialization to the other. This variability might be interpreted as a consequence of non-uniqueness of the Q-NMF for the specific choice of source and activation factors made here. However, this could be also attributed to the properties of the Q-ALS algorithm, due to the lack of convergence guarantees.
Nonetheless, these simulation results demonstrate the ability of the proposed Q-ALS algorithm to solve effectively the Q-NMF problem. As such, Q-ALS provides a baseline algorithm for further algorithmic developments.

VII. CONCLUSION

This paper has introduced a new powerful tool called quaternion non-negative matrix factorization (Q-NMF), which generalizes the well-known concept of non-negative matrix factorization (NMF) to the case of polarized signals. The algebraic representation of Stokes parameters using quaternions, together with the generalization of the non-negativity constraint on Stokes parameters made it possible to formulate and prove uniqueness results that may have been cumbersome to obtain otherwise. These results, stated by Proposition 2 for the two sources case and by Proposition 3 for the general case, illustrate the key disambiguating role played by polarization information. Furthermore, taking advantage of recent results in quaternion optimization, we have proposed a simple yet efficient algorithm for solving the Q-NMF problem in practice. These first results seem very promising and pave the way to future work on the theoretical and methodological aspects of the Q-NMF.

APPENDIX A

UNICITY CONDITIONS FOR Q-NMF

A. Proof of Proposition 2

Suppose that $P = 2$ and that the Q-NMF $X = WH$ exists. We suppose that conditions (C1) and (C2) are satisfied. To show that the Q-NMF is identifiable, we prove that intervals $I_{\alpha}^\text{Q-NMF}$ and $I_{\beta}^\text{Q-NMF}$ defined in (28)-(29) are restricted to $\{0\}$.

Condition (C1) permits to simplify (35) and (36) for $m_1, m_2$ like

$$\alpha^2 \left[ I_{m_2}^1(1 - \Phi_{m_2}) - 2I_{m_1}^1(1 - \Phi_{m_2} \langle \mu_{m_1}, \mu_{m_2} \rangle) \right] + 2\alpha \left[ I_{m_1}^1(1 - \Phi_{m_2} \langle \mu_{m_1}, \mu_{m_2} \rangle) \right] \geq 0 \tag{53}$$

$$\beta^2 \left[ I_{m_1}^1(1 - \Phi_{m_1}) - 2I_{m_2}^1(1 - \Phi_{m_1} \langle \mu_{m_1}, \mu_{m_2} \rangle) \right] + 2\beta \left[ I_{m_1}^1(1 - \Phi_{m_2} \langle \mu_{m_1}, \mu_{m_2} \rangle) \right] \geq 0 \tag{54}$$

Using inequalities linking $I_{m_1}, I_{m_2}$ for $m_1, m_2$ in (C1) yields the following domains of solutions to (53) and (54):

$$T_{\alpha, m_1}^\text{pol} = [0, \alpha_0], \quad T_{\beta, m_2}^\text{pol} = [0, \beta_0] \tag{55}$$

where $\alpha_0, \beta_0 \geq 1$. Condition (C2) imply that

$$T_{\alpha}^\text{Q-NMF} = [\alpha_0', 0], \quad T_{\beta}^\text{Q-NMF} = [\beta_0', 0] \tag{56}$$

where $\alpha_0', \beta_0' \leq 0$. By intersection of intervals, one gets that $T_{\alpha}^\text{Q-NMF}$ and $T_{\beta}^\text{Q-NMF}$ defined in (28)-(29) are restricted to $\{0\}$, so that the Q-NMF is unique.

B. Proof of Proposition 3

Suppose that the Q-NMF $X = WH$ exists and that sources never vanish, i.e. $w_{mp} > 0$ for every $m, p$. We obtain conditions (A1) and (A2) by contradiction: one at a time we suppose that either (A1) or (A2) is false and show that there exists a non-trivial matrix $T$ leading to different factors $\tilde{W} = WT$ and $\tilde{H} = T^{-1}H$.

Suppose that (A1) is not satisfied. Then, $\exists (p_0, q_0), p_0 \neq q_0$ such that

$$\forall m, \left( \Phi_{mp_0} \in [0, 1) \text{ or } \Phi_{mq_0} \notin \mu_{mp_0} \right), \tag{57}$$

or equivalently,

$$\forall m, \left( \Phi_{mp_0} \in [0, 1) \text{ or } \Phi_{mq_0} = 1 \text{ and } \mu_{mp_0} = \mu_{mp_0} \right). \tag{58}$$

Now consider the transformation $T_{p_{0}q_{0}}^\alpha \in \mathbb{R}^{p \times p}$ defined by

$$\forall k, \ell \in \{1, 2, \ldots, P\}, \left\{ \begin{array}{ll}
T_{p_{0}q_{0}}^\alpha k_{k} &= 1 \\
T_{p_{0}q_{0}}^\alpha p_{0} & = -\alpha \\
T_{p_{0}q_{0}}^\alpha k_{\ell} & = 0 \text{ otherwise}
\end{array} \right. \tag{59}$$

By construction $T_{p_{0}q_{0}}^\alpha$ does not correspond to a trivial ambiguity of the Q-NMF when $\alpha \neq 0$. Note that $\left( T_{p_{0}q_{0}}^\alpha \right)^{-1} = T_{p_{0}q_{0}}^{-\alpha}$ Consider the new factors $\tilde{W} = WT_{p_{0}q_{0}}^\alpha$ and $\tilde{H} = T_{p_{0}q_{0}}^{-\alpha}H$ such that $X = \tilde{W}H$.

From (59), the source indexed by $p_0$ is the only one affected by $T_{p_{0}q_{0}}^\alpha$ like

$$\forall m, \tilde{w}_{mp_0} = w_{mp_0} - \alpha w_{mq_0} \tag{60}$$

The corresponding activation coefficients read

$$\forall n, \tilde{h}_{pn} = h_{pn} + \alpha h_{qn} \tag{61}$$

Supposing $\alpha > 0, \tilde{h}_{p_0n} \geq 0$ for every $n$ by non-negativity of the matrix $H$. It remains to find at least one $\alpha > 0$ such that

$$\forall m, \tilde{w}_{mp_0} = w_{mp_0} - \alpha w_{mq_0} \in \mathbb{H}_S.$$ Imposing $\tilde{w}_{mp_0} \in \mathbb{H}_S$ for every $m$, yields

$$\forall m, \left\{ I_{mp_0} - \alpha I_{mq_0} \right\} \geq 0$$

$$\left\{ I_{mp_0} \Phi_{mp_0} \mu_{mp_0} - \alpha I_{mq_0} \Phi_{mq_0} \mu_{mq_0} \right\} \leq \left| I_{mp_0} - \alpha I_{mq_0} \right| \tag{62}$$

where we used notation (19) for convenience. Since by assumption $I_{mp_0}, I_{q_0} > 0$, the first condition implies that $\alpha \leq \min_m \left( I_{mp_0}/I_{mq_0} \right)$. The second condition related to polarization can be rewritten as, for every $m$

$$\alpha^2 I_{mq_0}^2 (1 - \Phi_{mq_0})$$

$$- 2\alpha I_{mp_0} I_{mq_0} (1 - \Phi_{mp_0} \Phi_{mq_0} \langle \mu_{mp_0}, \mu_{mq_0} \rangle) \tag{63}$$

$$+ I_{mp_0}^2 (1 - \Phi_{mp_0}^2) \geq 0 \tag{64}$$

According to (58) two cases may occur, not exclusively from each other. Fix $m$ and suppose that $\Phi_{mp_0} \in [0, 1)$. Two cases are possible: either $\Phi_{mq_0} = 1$ or $\Phi_{mq_0} \in (0, 1)$. Assume that $\Phi_{mq_0} = 1$, then (63) is true for $\alpha \leq I_{mp_0} (1 - \Phi_{mp_0}) I_{mp_0}^{-1} \Phi_{mp_0} (1 - \Phi_{mp_0} \langle \mu_{mp_0}, \mu_{mq_0} \rangle)$. For $\Phi_{mq_0} \in (0, 1)$, since $I_{mp_0} (1 - \Phi_{mq_0}^2) > 0$ in virtue of assumptions, the discriminant $\Delta$ reads

$$\Delta = 4 I_{mp_0}^2 I_{mq_0}^2 \left( (1 - \Phi_{mq_0} \Phi_{mp_0} \langle \mu_{mp_0}, \mu_{mq_0} \rangle)^2 \right)$$

$$- (1 - \Phi_{mp_0}^2) (1 - \Phi_{mq_0}^2) \tag{64}$$

Ranges of parameters implies that $\Delta \geq 0$. As a result, (63) is true for $\alpha \in (-\infty, \alpha^\text{m}] \cup [\alpha^\text{m}, +\infty)$ where $\alpha^\text{m}$ are the roots of the polynomial. Sign of polynomial coefficients in (63) together with $0 \leq \Phi_{mp_0} < 1$ imply that $\alpha^\text{m} > 0$. Suppose now that $\Phi_{mp_0} = 1$ and $\Phi_{mq_0} = \mu_{mp_0}$. Taking $\Phi_{mp_0} = 1$ Eq.
becomes trivial. The case \( \Phi_{mq_0} \in [0, 1) \) is included in the previous discussion.

Summarizing all cases, for every \( m \) such that \( \Phi_{mq_0} \in [0, 1) \) or \( \Phi_{mq_0} \mu_{mq_0} = \mu_{mq_0} \), there always exists \( \alpha > 0 \) such that conditions (62) are satisfied, meaning that the Q-NMF is not unique. This leads to the first necessary condition (A2). Repeating the same approach with \( \alpha < 0 \) yields condition (A2).

**Appendix B**

**Derivation of quaternion ALS updates**

**A.Quaternion derivatives using generalized \( \mathbb{H} \times \mathbb{R} \)-calculus**

Given a function \( f : \mathbb{H} \to \mathbb{H} \) of the variable \( q \in \mathbb{H} \), one outstanding question is: does its derivative \( \partial f/\partial q \) exists and if so, how do we compute it? Quaternion analytic functions [40], [41] are known to be differentiable, unfortunately they form a very restricted class of functions—of little interest to signal processing. In fact, cost functions \( f : \mathbb{H} \to \mathbb{R} \) are not analytic [40], so that other strategies need to be deployed. First, a pseudo-derivative approach can be used by treating \( f(q_0, q_1, q_2, q_3) \) of the four real components of \( q \). As pointed out in [29], such approach requires lengthy and cumbersome computations, thus limiting its applicability. In order to compute derivatives directly in the quaternion domain, crucial steps have been made recently with the development of the \( \mathbb{H} \times \mathbb{R} \)-calculus [32], and subsequently the generalized \( \mathbb{H} \times \mathbb{R} \)-calculus [31]. The latter provides a complete framework including chain rule and product rules, extending naturally the \( \mathbb{C} \times \mathbb{R} \)-calculus [42] of complex-valued optimization to the case of quaternion functions.

As a complete description of the generalized \( \mathbb{H} \times \mathbb{R} \)-calculus is well beyond the scope of this paper, we only point out some important results related to quaternion optimization and refer to [31], [29] for further reference. From now on, consider the case of real-valued function of quaternion matrices \( f : \mathbb{H}^{M \times N} \to \mathbb{R} \).

**Proposition 4** ([29], Theorem 4.1). Let \( f : \mathbb{H}^{M \times N} \to \mathbb{R} \) and denote by \( \nabla_Q f \) its gradient w.r.t. \( Q \). Then,

\[
Q_0 \text{ is a stationary point} \iff \nabla_Q f(Q_0) = \nabla_{Q'} f(Q_0) = 0 
\]  

**Proposition 5** ([29], Theorem 4.3). Let \( f : \mathbb{H}^{M \times N} \to \mathbb{R} \). Its gradient w.r.t. \( Q \), \( \nabla_Q f \) defines the direction of the maximum rate of change of \( f \).

These two propositions lay out the basic results for computing stationary points as well as developing gradient-descent type algorithm directly in the quaternion domain. Table I reprints from [29] some useful functions of quaternion matrix variable \( Q \) and their corresponding gradients.

**B. Updates for \( H \)**

Due to non-commutativity of the quaternion product, standard rules matrix derivatives cannot be applied directly. Instead, let us write explicitly the Euclidean cost function

\[
||X - WH||^2_F = \sum_{m,n} \left| x_{mn} - \sum_k w_{mk}h_{kn} \right|^2 
\]  

Replacing the same approach with \( f \), we get

\[
= \sum_{m,n} \left( x_{mn} - \sum_k w_{mk}h_{kn} \right) \left( x_{mn} - \sum_k w_{mk}h_{kn} \right) 
\]  

where we have used that \( \overline{w_{mk}h_{kn}} = \overline{w_{mk}}h_{kn} \) since \( h_{kn} \) is real-valued. Using the chain rule, the partial derivative of \( f \) w.r.t. \( h_{ij} \) is

\[
\frac{\partial ||X - WH||^2_F}{\partial h_{ij}} = -\sum_m \frac{\partial \sum_k w_{mk}h_{kn}}{\partial h_{ij}} \left( x_{mn} - \sum_k w_{mk}h_{kn} \right) 
\]  

\[
- \sum_m \frac{\partial w_{mi}}{\partial h_{ij}} \left( x_{mj} - \sum_k w_{mk}h_{kj} \right) 
\]  

\[
= -2\text{Re} \left[ \sum_m w_{mi} \left( x_{mj} - (WH)_{mj} \right) \right] 
\]  

As a result, one gets

\[
\nabla_H ||X - WH||^2_F = -2\text{Re} \left[ W^T(X - WH) \right]. 
\]

The Euclidean cost function is convex in \( H \); therefore, the minimizer \( H \) of this cost is obtained by cancelling out the gradient (73):

\[
-2\text{Re} \left[ W^T \tilde{X} \right] + 2\text{Re} \left[ W^T \tilde{W} \tilde{H} \right] = 0, 
\]

where again we used that \( \overline{W\tilde{H}} = \overline{W}\tilde{H} \) since \( \tilde{H} \) is a real matrix. It also implies that \( \text{Re} \left[ W^T \tilde{W} \right] = \text{Re} \left[ W^T W \right] \tilde{H} \),

| \( f(Q) \) | \( \nabla_Q f \) |
|---|---|
| \( \text{tr} A_1 Q A_2 \) | \( -\frac{1}{2} A_1^T A_2^T \) |
| \( \text{tr} A_1 Q^T A_2 \) | \( \text{Re} \left[ A_2^T A_1 \right] \) |
| \( \text{tr} A_1 Q^T Q A_2 \) | \( \text{Re} \left[ (Q A_2) A_1 - \frac{1}{2} (A_1 Q)^T A_2^T \right] \) |
leading to the following expression for $\hat{H}$:

$$\hat{H} = \arg \min_{H} \| X - WH \|^2_F$$

$$= \left( \text{Re} \left[ W^T \bar{W} \right] \right)^{-1} \text{Re} \left[ W^T X \right]$$  \hspace{1cm} (75)

\[  \hfill (76) \]

\[  \hfill (77) \]

C. Updates for $W$

Let us rewrite the Euclidean distance $\| X - WH \|^2_F$ in terms of a trace operator

$$\| X - WH \|^2_F = \text{tr} \left[ (X - WH)^\dagger (X - WH) \right]$$  \hspace{1cm} (78)

where $Q^\dagger$ is the (quaternion) conjugate-transpose of $Q$. Developing (77) one gets

$$\| X - WH \|^2_F = \text{tr} X^\dagger X - \text{tr} X^\dagger WH - \text{tr} W^\dagger X + \text{tr} W^\dagger WH$$  \hspace{1cm} (79)

According to Proposition 5, the gradient of the Euclidean cost with respect to $W$ defines the direction of maximum rate of change. Then, using results from Table 4 one gets

$$\nabla_W \| X - WH \|^2_F = \frac{1}{2} X^\dagger H^\top - \text{Re} \left[ X H^\top \right]$$

$$+ \text{Re} \left[ WH \right] H^\top - \frac{1}{2} \left( H^\top W^\top \right)^\top H^\top$$

$$= - \left( \text{Re} \left[ X \right] - \frac{1}{2} X + \frac{1}{2} WH - \text{Re} \left[ WH \right] \right) H^\top$$

$$= - \frac{1}{2} \left( X - WH \right) H^\top.$$  \hspace{1cm} (80)

The Euclidean cost function is convex in $W$. By Proposition 4 cancelling out the gradient (81) yields the global minimizer $\hat{W}$ of this cost such that

$$-XH^\top + WHH^\top = 0$$  \hspace{1cm} (82)

so that

$$\hat{W} = \arg \min_{W} \| X - WH \|^2_F = XH^\top \left( HH^\top \right)^{-1}.$$  \hspace{1cm} (83)

REFERENCES

[1] K. Aki and P. G. Richards, *Quantitative Seismology, 2nd Ed.* University Science Books, 2002.

[2] M. Born and E. Wolf, *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*, 6th ed. Pergamon Press, 1980.

[3] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation*. W. H. Freeman and Company, 1973.

[4] J.-S. Lee and E. Potter, *Polarimetric radar imaging: from basics to applications*. CRC press, 2009.

[5] J. S. Tyo, D. L. Goldstein, D. B. Chenault, and J. A. Shaw, “Review of passive imaging polarimetry for remote sensing applications,” *Applied optics*, vol. 45, no. 22, pp. 5453–5469, 2006.

[6] S. Miron, M. Dossot, C. Carteret, S. Margueron, and D. Brie, “Joint processing of the parallel and crossed polarized raman spectra and uniqueness in blind nonnegative source separation,” *Chemometrics and Intelligent Laboratory Systems*, vol. 105, no. 1, pp. 7–18, 2011.

[7] 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.

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

[9] H. Lee, “Eigenvalues and canonical forms of matrices with quaternion coefficients,” in *Proceedings of the Royal Irish Academy. Section A: Mathematical and Physical Sciences*, vol. 52. JSTOR, 1948, pp. 255–260.

[10] F. Zhang, “Quaternions and matrices of quaternions,” *Linear algebra and its applications*, vol. 251, pp. 21–57, 1997.

[11] J. H. Conway and D. A. Smith, *On quaternions and octonions: their geometry, arithmetic, and symmetry*. AK Peters, 2003.

[12] J. L. Brenner, “Matrices of quaternions.” *Pacific Journal of Mathematics*, vol. 1, no. 3, pp. 329–335, 1951.
[37] 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.

[38] M. Schwanninger, J. C. Rodrigues, and K. Fackler, “A review of band assignments in near infrared spectra of wood and wood components,” *Journal of Near Infrared Spectroscopy*, vol. 19, no. 5, pp. 287–308, 2011.

[39] N. Le Bihan, “The geometry of proper quaternion random variables.” *Signal Processing*, vol. 138, pp. 106–116, 2017.

[40] A. Sudbery, “Quaternionic analysis,” in *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 85, no. 2. Cambridge University Press, 1979, pp. 199–225.

[41] R. Watson, “The generalized cauchy-riemann-fueter equation and handedness,” *Complex Variables*, vol. 48, no. 7, pp. 555–568, 2003.

[42] K. Kreutz-Delgado, “The complex gradient operator and the cr-calculus,” *arXiv preprint arXiv:0906.4835*, 2009.