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

Julien Flamant, Sebastian Miron, David Brie

Abstract—This article introduces quaternion non-negative matrix factorization (QNMF), which generalizes the usual non-negative matrix factorization (NMF) to the case of polarized signals. Polarization information is represented by Stokes parameters, a set of 4 energetic parameters widely used in polarimetric imaging. QNMF relies on two key ingredients: *(i)* the algebraic representation of Stokes parameters thanks to quaternions and *(ii)* the exploitation of physical constraints on Stokes parameters. These constraints generalize non-negativity to the case of polarized signals, encoding positive semi-definiteness of the covariance matrix associated which each source. Uniqueness conditions for QNMF are presented. Remarkably, they encompass known sufficient uniqueness conditions from NMF. Meanwhile, QNMF further relaxes NMF uniqueness conditions requiring sources to exhibit a certain zero-pattern, by leveraging the complete polarization information. We introduce a simple yet efficient algorithm called quaternion alternating least squares (QALS) to solve the QNMF problem in practice. Closed-form quaternion updates are obtained using the recently introduced generalized HR calculus. Numerical experiments on synthetic data demonstrate the relevance of the approach. QNMF defines a promising generic low-rank approximation tool to handle polarization, notably for blind source separation problems arising in imaging applications.

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 [2], optics [3] to gravitational astronomy [4], among others. Polarization encodes the geometry of wave oscillations. It carries crucial morphological and physical insights about the medium which transmitted or reflected the polarized wave. For instance, polarization permits the detection of hidden targets in remote sensing [5], [6], it provides insights into solar magnetic fields [7] and reveals the anisotropy of biological tissues [8]. In summary, *polarimetric imaging* proves to be useful in a wide range of situations as it provides unique features of the observed scene, unaccessible to conventional intensity imaging.

Recent years have seen an increased interest in exploiting polarization diversity in hyperspectral imaging systems [9]–[11]. This modality, called *spectro-polarimetric imaging*, was first popularized in astronomy [12]. In full generality, it consists in acquiring, for a collection of wavelengths and locations, a 4-dimensional real vector gathering the 4 Stokes parameters. These energetic parameters are widely used in

optics [13] to describe the complete properties of light (intensity and polarization). The joint acquisition of spatial, spectral and polarization diversities raises important and original challenges, in particular regarding blind polarized source separation problems.

Spectro-polarimetric data can be represented as a third order real-valued tensor, where the rows, columns and fibers correspond to frequency, space and polarization diversities, respectively. Note that, since polarization is described here by Stokes parameters, the 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 [13] which need to be taken into account to guarantee the physical interpretation of the recovered rank-1 terms. Second, low-rank decompositions of spectro-polarimetric images such as the Canonical Polyadic Decomposition (CPD) [14] necessarily yield 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 quaternion non-negative matrix factorization (QNMF). It extends the concept of non-negative matrix factorization (NMF) to polarized sources. 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. Non-negativity is to be understood here in a *vectorial* sense: it encodes the non-negativity, or positive semi-definiteness, of the 2-by-2 complex covariance matrix associated with each polarized source, as will be explained later.

This *generalized* non-negativity constraint on the source factor makes QNMF fundamentally different from NMF. In fact, the three Stokes parameters describing polarization are not required to be real non-negative: instead, Stokes parameters must lie inside a particular 4D second-order convex cone. This means that QNMF cannot be simply reformulated as an augmented real NMF problem by stacking Stokes parameters in a taller matrix. Instead, QNMF preserves the intrinsic structure of Stokes parameters by leveraging their quaternion representation: it allows the associated generalized non-negativity constraints to be handled straightforwardly.

QNMF encodes a quaternion blind source separation problem. However, unlike quaternion ICA [15], [16] which relies on a standard quaternion linear mixing model and assumes

statistical independence of the sources, QNMF does not require any statistical model; it defines a quaternion linear mixing model with non-negativity constraints only. Such non-negativity constraints are essential to physical interpretation: they guarantee the general relevance of QNMF to solve blind unmixing of spectro-polarimetric data, just like NMF offers a standard approach towards hyperspectral image unmixing [17]. More generally, recent works [18]–[21] have shown that Stokes parameters provide a natural reparameterization for the second-order properties of bivariate signals. As a result, QNMF defines a new low-rank approximation tool for arbitrary bivariate signals, such as current velocities in oceanography [22] or polarized gravitational waves [23].

Our contributions are threefold. First, we introduce the QNMF problem from a spectro-polarimetric linear mixing model perspective and demonstrate how it generalizes the usual NMF. Second, we study uniqueness conditions for the QNMF problem and reveal the key disambiguating role played by polarization. QNMF is shown to encompass known NMF sufficient uniqueness conditions, while relaxing NMF necessary uniqueness conditions. Finally, we provide a simple yet efficient algorithm in the quaternion domain that effectively solves the QNMF problem. We believe that quaternions, beyond offering an elegant way to handle polarization diversity, permit numerous theoretical (*e.g.* uniqueness conditions) and methodological (*e.g.* algorithmic design) developments that would have been cumbersome to obtain otherwise.

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 (QNMF). Section IV studies uniqueness conditions for QNMF, starting with the simpler 2 sources case, followed by a discussion of the general case. Section V introduces the quaternion alternating least squares (QALS) algorithm for solving the QNMF 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 representation of polarization using quaternions together with the exploitation of structural non-negativity constraints. To this end, Section II-A briefly reviews quaternion algebra and Section II-B describes polarization using Stokes parameters. Section II-C then introduces the quaternion representation of Stokes parameters, whereas Section II-D details structural non-negative like constraints on Stokes parameters.

A. Quaternions

Quaternions $\mathbb{H}$ forms a 4-dimensional normed division algebra over the real numbers $\mathbb{R}$ with canonical basis $\{1, \boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$ where $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ are imaginary units $\boldsymbol{i}^2 = \boldsymbol{j}^2 = \boldsymbol{k}^2 = -1$ such that $\boldsymbol{ij} = \boldsymbol{k}, \boldsymbol{ij} = -\boldsymbol{ji}$ and $\boldsymbol{ijk} = -1$. 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 $q = a + b\boldsymbol{i} + c\boldsymbol{j} + d\boldsymbol{k}$ where $a, b, c, d \in \mathbb{R}$. The real or scalar part of q is $\mathrm{Re}\, q = a$ and its imaginary or vector part is $\mathrm{Im}\, q = b\boldsymbol{i} + c\boldsymbol{j} + d\boldsymbol{k}$. We note $\mathrm{Im}_{\boldsymbol{i}}\, q = b$, $\mathrm{Im}_{\boldsymbol{j}}\, q = c$ and $\mathrm{Im}_{\boldsymbol{k}}\, q = d$. When $\mathrm{Re}\, q = 0$, q is said to be *pure imaginary*, or simply *pure*. The quaternion conjugate of q is denoted by $\overline{q} = \mathrm{Re}\, q - \mathrm{Im}\, q$. The modulus of q is $|q| = \sqrt{q\overline{q}} = \sqrt{\overline{q}q}$.

A quaternion matrix $\mathbf{Q} \in \mathbb{H}^{M \times N}$ has elements $\mathbf{Q}_{mn} = q_{mn} \in \mathbb{H}$. Notation $\overline{\mathbf{Q}}$ denotes the entry-wise conjugate of $\mathbf{Q}$ whereas its conjugate-transpose is denoted by $\mathbf{Q}^\dagger = \overline{\mathbf{Q}}^\top$. Usual matrix operations require special care due to non-commutativity of quaternion multiplication: *e.g.* given two matrices $\mathbf{Q}_1 \in \mathbb{H}^{M \times N}$ and $\mathbf{Q}_2 \in \mathbb{H}^{N \times P}$, in general one has $\overline{\mathbf{Q}_1 \mathbf{Q}_2} \neq \overline{\mathbf{Q}}_1 \overline{\mathbf{Q}}_2$ and $(\mathbf{Q}_1 \mathbf{Q}_2)^\top \neq \mathbf{Q}_2^\top \mathbf{Q}_1^\top$. However, $(\mathbf{Q}_1 \mathbf{Q}_2)^\dagger = \mathbf{Q}_2^\dagger \mathbf{Q}_1^\dagger$ is always true. More details on quaternions and their properties can be found in [24]. See also [25]–[29] and references therein for more on quaternion matrix algebra.

B. Polarization and Stokes parameters

For electromagnetic (EM) waves such as light traveling in free space, the instantaneous electric field vector $\mathbf{E}(t)$ lives in the 2D plane orthogonal to the direction of propagation. In such transverse waves, the end tip of $\mathbf{E}(t) \in \mathbb{R}^2$ describes a two-dimensional elliptical trajectory whose properties can vary with time, space or frequency. This phenomenon is called *polarization*: it describes the geometric nature of wave oscillations. This behaviour is not limited to electromagnetic radiation: transverse elastic waves and gravitational waves also exhibit polarization.

In optics [3], [13], [30], 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 that can be measured in experiments [31], [32]. 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 three remaining Stokes parameters S_1, S_2, S_3 describe the polarized part of light, *i.e.* the properties of the polarization ellipse encoded by the trajectory of the electric field vector. The *degree of polarization* $\Phi \in [0, 1]$ rules the relative contribution of the polarized part to the total intensity:

$$\Phi = \frac{\text{intensity of polarized part}}{\text{total intensity}} = \frac{\sqrt{S_1^2 + S_2^2 + S_3^2}}{S_0} . \quad (1)$$

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*.

C. Quaternion representation of Stokes parameters

The Stokes vector $(S_0, S_1, S_2, S_3)^\top \in \mathbb{R}^4$ can be conveniently expressed using quaternion algebra as

$$w = S_0 + \boldsymbol{i}S_3 + \boldsymbol{j}S_1 + \boldsymbol{k}S_2 \in \mathbb{H} . \quad (2)$$

This algebraic representation of Stokes parameters has been long standing in the field of optics [33]–[37]. More recently [18], [21], it has been shown that (2) exhibits a general relevance for arbitrary bivariate signals. Our choice for the ordering of Stokes parameters in (2) adopts that of [21].

The quaternion representation of Stokes parameters enables straightforward physical and geometric interpretations while simplifying computations [18]. Eq. (2) can be rewritten for convenience as

$$w = I + I\Phi\boldsymbol{\mu} \qquad (3)$$

where $I := S_0 \geq 0$ is the total intensity and $\boldsymbol{\mu}$ is the *polarization axis* – a pure unit quaternion such that $\boldsymbol{\mu}^2 = -1$. With this notation, the quantity $\Phi\boldsymbol{\mu}$ is a pure quaternion that can be identified with a vector of $\mathbb{R}^3$ inside the Poincaré sphere [3] which completely encodes the polarization state of light. The expression (3) also emphasizes one of the benefits of the quaternion formalism: it allows a straightforward separation between pure energetic information, conveyed by $\operatorname{Re} w = I$, and the geometric vector information encoded by $\operatorname{Im} w = I\Phi\boldsymbol{\mu}$.

D. Polarization constraint and quaternion non-negativity

Stokes parameters are energetic quantities which are subject to admissibility conditions. A vector $(S_0, S_1, S_2, S_3)^\top \in \mathbb{R}^4$ is admissible as a Stokes vector if it satisfies the following constraints $(\mathcal{S})$

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

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 polarized 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 (1). Note that $(\mathcal{S})$ does not impose at all that S_1, S_2 and S_3 are non-negative.

From a mathematical perspective, $(\mathcal{S})$ defines a closed second-order convex cone in $\mathbb{R}^4$, denoted by $\mathbb{R}_{\mathcal{S}}^4$. Conditions $(\mathcal{S})$ are referred to as *non-negativity* constraints for polarized signals by the following argument. Given $(S_0, S_1, S_2, S_3)^\top \in \mathbb{R}^4$, consider the 2-by-2 Hermitian matrix[1]:

$$\mathbf{J} = \frac{1}{2} \begin{bmatrix} S_0 + S_1 & S_2 + iS_3 \\ S_2 - iS_3 & S_0 - S_1 \end{bmatrix} \in \mathbb{C}^{2\times 2}. \qquad (4)$$

This mapping is bijective. Then, a quick check shows that imposing positive semi-definiteness of $\mathbf{J}$ – or simply, *non-negativity* of $\mathbf{J}$, is equivalent to $(\mathcal{S})$, *i.e.*

$$(\mathcal{S}) \Leftrightarrow \begin{cases} \operatorname{tr} \mathbf{J} \geq 0 \\ \det \mathbf{J} \geq 0 \end{cases}. \qquad (5)$$

This shows that $\mathbb{R}_{\mathcal{S}}^4$ and the set of 2-by-2 Hermitian non-negative matrices are isomorphic. This result will prove to be essential later on for imposing $(\mathcal{S})$ in practice, see Section V-A. Since the proposed approach relies on the quaternion representation (2) of Stokes parameters, we define by extension the set of *non-negative quaternions* $\mathbb{H}_{\mathcal{S}} \subset \mathbb{H}$ such that

$$\mathbb{H}_{\mathcal{S}} \triangleq \left\{ q \in \mathbb{H} \,\middle|\, \operatorname{Re} q \geq 0 \text{ and } |\operatorname{Im} q|^2 \leq (\operatorname{Re} q)^2 \right\} \,. \qquad (6)$$

[1]In optics, the matrix $\mathbf{J}$ is known as the *polarization* or *coherency* matrix [13, Section 1.4] associated with the bivariate electric field $\mathbf{E}$. It can be statistically defined as a covariance matrix [13, Section 1.4.1], from which Hermitian positive semi-definiteness follows.

By definition, the sets $\mathbb{H}_{\mathcal{S}}$ and $\mathbb{R}_{\mathcal{S}}^4$ are isomorphic. They share the same properties: in particular, this means that $\mathbb{H}_{\mathcal{S}}$ is a closed second-order convex cone of $\mathbb{H}$.

III. QUATERNION NON-NEGATIVE MATRIX FACTORIZATION

This section introduces Quaternion Non-negative Matrix Factorization (QNMF), a new tool which generalizes the notion of Non-negative Matrix Factorization (NMF) to the case of polarized signals. It exploits two key features: *(i)* the polarization constraint $(\mathcal{S})$, which extends the notion of non-negativity to polarized signals and *(ii)* the algebraic representation (2) of Stokes parameters using quaternions. QNMF 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. In such experiment, the four Stokes parameters $S_i(\nu_m, u_n)$, $i = 0, 1, 2, 3$ are acquired for a collection of sampled frequencies (or wavelengths) $(\nu_m)_{1 \leq m \leq M}$ and spatial locations $(u_n)_{1 \leq n \leq N}$. Using (2), the collected data can be written in quaternion form as

$$\begin{aligned} x(\nu_m, u_n) &= S_0(\nu_m, u_n) + iS_3(\nu_m, u_n) \\ &\quad + jS_1(\nu_m, u_n) + kS_2(\nu_m, u_n) \in \mathbb{H}_{\mathcal{S}} \end{aligned} \qquad (7)$$

for $1 \leq m \leq M$, $1 \leq n \leq N$. Note that spectro-polarimetric imaging includes standard hyperspectral imaging as a special case, in which only intensity measurements are performed *i.e.* $x(\nu_m, u_n) = S_0(\nu_m, u_n)$.

A standard assumption in hyperspectral imaging is to consider that, at a given pixel u_n, data can be represented as a linearly weighted combination of P elementary sources (or endmembers) [17], [38]. While being simple, the linear mixing model appears reasonable for many real-world settings [39], [40]. Generalizing this model to the case of polarized sources, we decompose spectro-polarimetric data (7) as the linear superposition of P elementary polarized sources $w_p(\cdot) \in \mathbb{H}_{\mathcal{S}}$:

$$x(\nu_m, u_n) = \sum_{p=1}^{P} w_p(\nu_m) h_p(u_n) \,. \qquad (8)$$

The quantity $h_p(\cdot) \geq 0$ denotes the spatial activation coefficients attached to the source p. Considering that $h_p(\cdot)$ is a non-negative real quantity has two key implications. First, it means that $h_p(u_n)$ can be interpreted as the abundance of the source p at location u_n. Second, it ensures that, for any elementary source $w_p(\cdot) \in \mathbb{H}_{\mathcal{S}}$ the linear combination $\sum_{p=1}^{P} w_p(\nu_m) h_p(u_n)$ is in $\mathbb{H}_{\mathcal{S}}$ since $\mathbb{H}_{\mathcal{S}}$ is a convex cone.

Rewriting (8) into matrix form introduces *Quaternion Non-Negative Matrix Factorization* (QNMF) like

$$\mathbf{X} = \mathbf{W}\mathbf{H} \qquad (9)$$

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

B. Relation with NMF

The QNMF problem extends the well-known NMF problem to the case of polarized signals. Compared to NMF, QNMF features a quaternion-valued sources factor $\mathbf{W}$ which exploits the constraint $(\mathcal{S})$ instead of the usual non-negativity constraint. Using quaternion algebra rules, the QNMF (9) can also be rewritten as

$$\mathbf{X} = \mathbf{WH} \Leftrightarrow \begin{cases} \operatorname{Re} \mathbf{X} = [\operatorname{Re} \mathbf{W}] \mathbf{H} & \text{(NMF)} \\ \operatorname{Im} \mathbf{X} = [\operatorname{Im} \mathbf{W}] \mathbf{H} & \text{(polarization)} \end{cases} . \quad (10)$$

Eq. (10) shows that QNMF can be seen as a co-factorization problem with common activation factor $\mathbf{H}$. The first factorization problem is an usual NMF on the real part of $\mathbf{X}$, *i.e.* on intensity data (Stokes parameter $S_0 \geq 0$) only. The second one corresponds to a factorization problem on the imaginary part of $\mathbf{X}$ 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 $\mathbf{H}$ appears in both and *(ii)* for each coefficient (m, p) of the source factor $\mathbf{W}$, the constraint $(\mathcal{S})$ connects the modulus of the imaginary part (polarization factorization problem) to its real part (NMF problem) by an inequality constraint.

The relationship (10) provides another illustration of how QNMF extends NMF to account for polarization diversity. It allows a precise quantification of the role played by the polarization information and its associated constraint $(\mathcal{S})$. This will be particularly useful for comparing uniqueness conditions of QNMF and NMF, see Section IV.

IV. UNIQUENESS CONDITIONS FOR QNMF

This section deals with a fundamental question: upon which conditions on the source $\mathbf{W}$ and activation $\mathbf{H}$ factors is the QNMF $\mathbf{X} = \mathbf{WH}$ unique? The *uniqueness or identifiability* question is central in any factorization or decomposition problem, as it is intimately related to the interpretability of the underlying model. Many authors have thus considered this question for the NMF problem, either by an analytical study of the properties of the NMF model [41]–[44] or by the design of suitable identification criteria to guarantee identifiability [45]–[48]. In this paper, we address the QNMF uniqueness question by a theoretical study of the properties of the QNMF model. To this aim, Section IV-A starts by introducing the intrinsic ambiguities of the QNMF model. We provide sufficient uniqueness conditions for QNMF, directly inherited from known NMF ones. To reveal the key role played by polarization information in model identifiability, Section IV-B studies in detail the $P = 2$ sources case and provides a sufficient condition for QNMF uniqueness. Section IV-C considers the general $P \geq 2$ case and gives a necessary uniqueness condition that relaxes the requirement for sources $\mathbf{W}$ to have some zero entries.

A. Trivial ambiguities and NMF-based uniqueness

Let $\mathbf{X} \in \mathbb{H}_{\mathcal{S}}^{M \times N}$ and suppose that there exists two matrices $\mathbf{W} \in \mathbb{H}_{\mathcal{S}}^{M \times P}$ and $\mathbf{H} \in \mathbb{R}_{+}^{P \times N}$ such that $\mathbf{X} = \mathbf{WH}$ holds.

Given a nonsingular matrix $\mathbf{T} \in \mathbb{H}^{P \times P}$, any pair $(\widetilde{\mathbf{W}}, \widetilde{\mathbf{H}})$ defined as

$$\widetilde{\mathbf{W}} = \mathbf{WT} \quad (11)$$

$$\widetilde{\mathbf{H}} = \mathbf{T}^{-1} \mathbf{H} \quad (12)$$

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

$$\widetilde{\mathbf{W}} \in \mathbb{H}_{\mathcal{S}}^{M \times P} \text{ and } \widetilde{\mathbf{H}} \in \mathbb{R}_{+}^{P \times N} . \quad (13)$$

The real-valuedness constraint on $\widetilde{\mathbf{H}}$ directly imposes that $\mathbf{T}$ is a real matrix. Indeed, if the entries $\mathbf{T}$ are quaternion-valued, so are the entries of $\mathbf{T}^{-1}$, and thus $\widetilde{\mathbf{H}} = \mathbf{T}^{-1}\mathbf{H}$ is also quaternion-valued. As a result, only linear transformations $\mathbf{T} \in \mathbb{R}^{P \times P}$ may yield alternative factors $\widetilde{\mathbf{W}}$ and $\widetilde{\mathbf{H}}$ satisfying (13), just as in the standard NMF case. As a result, QNMF exhibits the same trivial ambiguities as NMF, summarized by Proposition 1 below.

Proposition 1. *The QNMF $\mathbf{X} = \mathbf{WH}$ exhibits two intrinsic ambiguities, namely,*

- *scale indeterminacy:*

$$\mathbf{T} = \operatorname{diag}(t_1, t_2, \ldots, t_P), t_i > 0, 1 \leq i \leq P \quad (14)$$

- *order indeterminacy:*

$$\mathbf{T} \text{ is a permutation matrix} \quad (15)$$

for which $\widetilde{\mathbf{W}} = \mathbf{WT}$ and $\widetilde{\mathbf{H}} = \mathbf{T}^{-1}\mathbf{H}$ define equally valid QNMF factors.

These inevitable indeterminacies motivate the following definition for the identifiability of the QNMF problem.

Definition 1. *The QNMF $\mathbf{X} = \mathbf{WH}$ is said to be essentially unique if $\mathbf{X} = \widetilde{\mathbf{W}}\widetilde{\mathbf{H}}$ implies $\widetilde{\mathbf{W}} = \mathbf{WDP}$ and $\widetilde{\mathbf{H}} = (\mathbf{DP})^{-1}\mathbf{H}$ where $\mathbf{D}$ is a diagonal matrix with strictly positive entries and $\mathbf{P}$ is a permutation matrix.*

QNMF features the same trivial ambiguities as NMF: those can be handled by standard techniques, *e.g.* by imposing normalization on the source coefficients to get rid of the scale indeterminacy.

As explained in Section III-B, QNMF generalizes NMF to the case of polarized signals. The co-factorization perspective given by (10) shows that QNMF inherits sufficient uniqueness conditions from NMF.

Proposition 2 (NMF-based sufficient condition for uniqueness)**.** *If the NMF $\operatorname{Re} \mathbf{X} = [\operatorname{Re} \mathbf{W}] \mathbf{H}$ is essentially unique, then the QNMF $\mathbf{X} = \mathbf{WH}$ is essentially unique.*

Proof. Suppose that $\operatorname{Re} \mathbf{X} = [\operatorname{Re} \mathbf{W}] \mathbf{H}$ is essentially unique. Thus $\mathbf{H}$ can always be written as $\mathbf{H} = \mathbf{DP}\widetilde{\mathbf{H}}$ where $\mathbf{D}$ is a diagonal matrix with strictly positive entries and $\mathbf{P}$ is a permutation matrix. Then $\mathbf{X} = \mathbf{WH} = \widetilde{\mathbf{W}}\widetilde{\mathbf{H}}$, with $\widetilde{\mathbf{W}} = \mathbf{WDP}$ and $\widetilde{\mathbf{H}} = (\mathbf{DP})^{-1}\mathbf{H}$ which shows that the QNMF is essentially unique by Definition 1. $\square$

Proposition 2 shows that known sufficient conditions for NMF directly extend to the QNMF case. For instance, if $\operatorname{Re} \mathbf{W}$

and $\mathbf{H}$ satisfy the separability related assumptions of [42], [43] or the more relaxed sufficiently scattered condition [44], then the QNMF $\mathbf{X} = \mathbf{WH}$ is essentially unique. While this result is reassuring – QNMF encompasses known NMF uniqueness conditions – it does not take advantage of polarization information contained in $\text{Im}\,\mathbf{W}$. The two remaining sections thus explore how this supplementary information together with the polarization constraint ($\mathcal{S}$) impact QNMF model identifiability.

B. Two sources case: range of admissible solutions and sufficient uniqueness condition

We first consider the simpler case of two sources ($P = 2$), which is particularly instrumental in understanding the role played by the polarization constraint ($\mathcal{S}$). The approach follows closely the one presented in [41] for the NMF case.

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

$$(\mathbf{W})_{mp} = w_{mp} \in \mathbb{H}_{\mathcal{S}} \text{ and } (\mathbf{H})_{pn} = h_{pn} \geq 0 \quad (16)$$

where $p = 1, 2$ denotes the source index. We use the convenient quaternion form (3) for w_{mp}, i.e.

$$w_{mp} = I_{mp} + I_{mp}\Phi_{mp}\boldsymbol{\mu}_{mp} . \quad (17)$$

To model QNMF indeterminacies, we parameterize the matrix $\mathbf{T}$ in Eqs. (11) – (12) as $\mathbf{T} = \mathbf{T}(\alpha, \beta)$, $\alpha, \beta \in \mathbb{R}$ such that

$$\mathbf{T}(\alpha, \beta) = \begin{bmatrix} 1 - \alpha & \beta \\ \alpha & 1 - \beta \end{bmatrix} . \quad (18)$$

This form for $\mathbf{T}(\alpha, \beta)$, together with the assumption $\alpha + \beta < 1$ ensures that $\mathbf{T}(\alpha, \beta)$ models only non-trivial ambiguities of QNMF, i.e. other than those given in Proposition 1. It follows that $\mathbf{T}$ is invertible, with inverse

$$\mathbf{T}^{-1}(\alpha, \beta) = \frac{1}{1 - \alpha - \beta} \begin{bmatrix} 1 - \beta & -\beta \\ -\alpha & 1 - \alpha \end{bmatrix} . \quad (19)$$

Now consider the linear transformed factors $\widetilde{\mathbf{W}} = \mathbf{WT}(\alpha, \beta)$ and $\widetilde{\mathbf{H}} = \mathbf{T}^{-1}(\alpha, \beta)\mathbf{H}$. Using (18), the mixing of sources explicitly reads for $1 \leq m \leq M$

$$\widetilde{w}_{m1} = (1 - \alpha)w_{m1} + \alpha w_{m2} \quad (20)$$
$$\widetilde{w}_{m2} = \beta w_{m1} + (1 - \beta)w_{m2} . \quad (21)$$

With (19) one gets the new activations for $1 \leq n \leq N$

$$\widetilde{h}_{1n} = \frac{(1 - \beta)h_{1n} - \beta h_{2n}}{1 - \alpha - \beta} \quad (22)$$
$$\widetilde{h}_{2n} = \frac{-\alpha h_{1n} + (1 - \alpha)h_{2n}}{1 - \alpha - \beta} . \quad (23)$$

Non-negativity of $\widetilde{\mathbf{H}}$ imposes that $\widetilde{h}_{1n} \geq 0$ and $\widetilde{h}_{2n} \geq 0$ for every n. The polarization constraints ($\mathcal{S}$) impose that $\widetilde{w}_{m1}, \widetilde{w}_{m2} \in \mathbb{H}_{\mathcal{S}}$. In other words, for every m, one has

$$(\text{NMF}) \begin{cases} \text{Re}\,\widetilde{w}_{m1} \geq 0 \\ \text{Re}\,\widetilde{w}_{m2} \geq 0 \end{cases} , \quad (24)$$

$$(\text{polarization}) \begin{cases} |\text{Im}\,\widetilde{w}_{m1}|^2 \leq (\text{Re}\,\widetilde{w}_{m1})^2 \\ |\text{Im}\,\widetilde{w}_{m2}|^2 \leq (\text{Re}\,\widetilde{w}_{m2})^2 \end{cases} . \quad (25)$$

Eq. (24) corresponds also to the usual NMF case [41]. However, condition (25) reveals the specificy of QNMF.

In order to find the range of admissible solutions for QNMF, we search for intervals $\mathcal{I}_\alpha^{\text{QNMF}}$ and $\mathcal{I}_\beta^{\text{QNMF}}$ such that, for every $(\alpha, \beta) \in \mathcal{I}_\alpha^{\text{QNMF}} \times \mathcal{I}_\beta^{\text{QNMF}}$ the linear transform $\mathbf{T}(\alpha, \beta)$ yields valid $\widetilde{\mathbf{W}}$ and $\widetilde{\mathbf{H}}$ factors. Remark that, since Eqs. (22) – (24) are identical to the NMF case, we can express the intervals $\mathcal{I}_\alpha^{\text{QNMF}}$ and $\mathcal{I}_\beta^{\text{QNMF}}$ as

$$\mathcal{I}_\alpha^{\text{QNMF}} = \mathcal{I}_\alpha^{\text{NMF}} \bigcap_m \mathcal{I}_{\alpha,m}^{\text{pol}} \quad (26)$$

$$\mathcal{I}_\beta^{\text{QNMF}} = \mathcal{I}_\beta^{\text{NMF}} \bigcap_m \mathcal{I}_{\beta,m}^{\text{pol}} \quad (27)$$

where $\mathcal{I}_\alpha^{\text{NMF}}$ (resp. $\mathcal{I}_\beta^{\text{NMF}}$) is the interval obtained for α (resp. β) using non-negativity constraints only, as explained below. For every m, $\mathcal{I}_{\alpha,m}^{\text{pol}}$ and $\mathcal{I}_{\beta,m}^{\text{pol}}$ indicate intervals for α and β such that the polarization constraint (25) holds. Thanks to this rewriting, Eqs. (26)–(27) 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 QNMF to improve NMF identifiability.

Computation of $\mathcal{I}_\alpha^{\text{NMF}}$ and $\mathcal{I}_\beta^{\text{NMF}}$. These intervals encode 2 non-negativity conditions: the real part of the sources and the activation coefficients. Their expression follow directly from NMF results [41], that is:

$$\mathcal{I}_\alpha^{\text{NMF}} = (\alpha_{\min}^{\text{NMF}}, \alpha_{\max}^{\text{NMF}}] \text{ and } \mathcal{I}_\beta^{\text{NMF}} = (\beta_{\min}^{\text{NMF}}, \beta_{\max}^{\text{NMF}}] . \quad (28)$$

Lower bounds depend on the ratio of the real parts of the sources like

$$\alpha_{\min}^{\text{NMF}} = -\min_{m \in \mathbb{M}_1} \frac{I_{m1}}{I_{m2} - I_{m1}}, \; \beta_{\min}^{\text{NMF}} = -\min_{m \in \mathbb{M}_2} \frac{I_{m2}}{I_{m1} - I_{m2}} \quad (29)$$

where for convenience $\text{Re}\,w_{mp} = I_{mp}$ as in (17) and $\mathbb{M}_1 = \{m | I_{m2} > I_{m1}\}$, $\mathbb{M}_2 = \{m | I_{m1} > I_{m2}\}$ are supposed to be non-empty sets. If $\mathbb{M}_1 = \emptyset$ (resp. $\mathbb{M}_2 = \emptyset$), one has $\alpha_{\min}^{\text{NMF}} = -\infty$ (resp. $\beta_{\min}^{\text{NMF}} = -\infty$). Activations coefficients control the upper bounds of $\mathcal{I}_\alpha^{\text{NMF}}$ and $\mathcal{I}_\beta^{\text{NMF}}$:

$$\alpha_{\max}^{\text{NMF}} = \min_n \frac{h_{2n}}{h_{1n} + h_{2n}}, \; \beta_{\max}^{\text{NMF}} = \min_n \frac{h_{1n}}{h_{1n} + h_{2n}} . \quad (30)$$

Computation of $\mathcal{I}_\alpha^{\text{pol}}$ and $\mathcal{I}_\beta^{\text{pol}}$. Let us now exploit the polarization condition (25), on which relies QNMF. Fix m ($1 \leq m \leq M$) and start with parameter α. Plugging (17) into (20) ones gets:

$$\text{Re}\,\widetilde{w}_{m1} = (1 - \alpha)I_{m1} + \alpha I_{m2} , \quad (31)$$
$$\text{Im}\,\widetilde{w}_{m1} = (1 - \alpha)I_{m1}\Phi_{m1}\boldsymbol{\mu}_{m1} + \alpha I_{m2}\Phi_{m2}\boldsymbol{\mu}_{m2} . \quad (32)$$

By developing all terms appearing in the polarization constraint (25) and reorganizing the expression as a second-order polynomial in α, we get the following inequality

$$\alpha^2 \left[I_{m1}^2(1 - \Phi_{m1}^2) + I_{m2}^2(1 - \Phi_{m2}^2) \right.$$
$$\left. -2I_{m1}I_{m2}(1 - \Phi_{m1}\Phi_{m2}\langle\boldsymbol{\mu}_{m1}, \boldsymbol{\mu}_{m2}\rangle) \right]$$
$$+2\alpha \left[I_{m1}I_{m2}(1 - \Phi_{m1}\Phi_{m2}\langle\boldsymbol{\mu}_{m1}, \boldsymbol{\mu}_{m2}\rangle) - I_{m1}^2(1 - \Phi_{m1}^2) \right]$$
$$+I_{m1}^2(1 - \Phi_{m1}^2) \geq 0 \quad (33)$$

where $\langle \boldsymbol{\mu}_1, \boldsymbol{\mu}_2 \rangle = -\mathrm{Re}\, \boldsymbol{\mu}_1 \boldsymbol{\mu}_2$ can be identified with the usual inner product of $\mathbb{R}^3$. Starting from (21), the same approach is used for β, leading to

$$\beta^2 \left[I_{m2}^2 (1 - \Phi_{m2}^2) + I_{m1}^2 (1 - \Phi_{m1}^2) \right.$$
$$\left. - 2 I_{m1} I_{m2} (1 - \Phi_{m1} \Phi_{m2} \langle \boldsymbol{\mu}_{m1}, \boldsymbol{\mu}_{m2} \rangle) \right]$$
$$+ 2\beta \left[I_{m1} I_{m2} (1 - \Phi_{m1} \Phi_{m2} \langle \boldsymbol{\mu}_{m1}, \boldsymbol{\mu}_{m2} \rangle) - I_{m2}^2 (1 - \Phi_{m2}^2) \right]$$
$$+ I_{m2}^2 (1 - \Phi_{m2}^2) \geq 0 \,. \tag{34}$$

For a given m, intervals $\mathcal{I}_{\alpha,m}^{\mathrm{pol}}$ and $\mathcal{I}_{\beta,m}^{\mathrm{pol}}$ are obtained by solving the corresponding polynomial inequalities (33) and (34). It involves finding the roots of the associated second-order polynomials in α and β, respectively. An easy but lengthy computation of polynomial discriminants shows that (33) and (34) always admit two real-valued solutions (possibly degenerate). Unfortunately, the amount of parameters involved (at least 4: $\Phi_{m1}, \Phi_{m2}, \langle \boldsymbol{\mu}_{m1}, \boldsymbol{\mu}_{m2} \rangle$ and the ratio I_{m1}/I_{m2}) prevents from performing a general theoretical study of the roots behaviour. Such a study would require to account for 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 when values of sources parameters are given, Eqs. (33) and (34) can be solved very efficiently numerically to yield desired intervals $\mathcal{I}_{\alpha,m}^{\mathrm{pol}}$ and $\mathcal{I}_{\beta,m}^{\mathrm{pol}}$. See Section VI for numerical illustrations on synthetic and real-world data.

Sufficient uniqueness condition. Despite the apparent complexity of the polarization conditions (33) and (34), a simple and interpretable sufficient condition for the essential uniqueness of QNMF with $P = 2$ can be formulated.

Proposition 3. *If the following conditions are satisfied:*
- $\exists\, m_1, m_2 \in \{1, 2, \ldots M\}$ *s.t.*

$$\begin{cases} \Phi_{m_1 1} = 1, \ \Phi_{m_1 2} \boldsymbol{\mu}_{m_1 2} \neq \boldsymbol{\mu}_{m_1 1} \\ I_{m_1 1} \geq \dfrac{1}{2} \dfrac{1 - \Phi_{m_1 2}^2}{1 - \Phi_{m_1 2} \langle \boldsymbol{\mu}_{m_1 1}, \boldsymbol{\mu}_{m_2 2} \rangle} I_{m_1 2} \\ \Phi_2(k_2) = 1, \ \Phi_{m_2 1} \boldsymbol{\mu}_{m_2 1} \neq \boldsymbol{\mu}_{m_2 2} \\ I_{m_2 2} \geq \dfrac{1}{2} \dfrac{1 - \Phi_{m_2 1}^2}{1 - \Phi_{m_2 1} \langle \boldsymbol{\mu}_{m_2 2}, \boldsymbol{\mu}_{m_2 1} \rangle} I_{m_2 1} \end{cases} \tag{C1}$$

- $\exists\, n_1, n_2 \in \{1, 2, \ldots N\}, n_1 \neq n_2$ *s.t.*

$$\begin{cases} h_{1 n_1} > 0 \ and \ h_{2 n_1} = 0 \\ h_{2 n_2} > 0 \ and \ h_{1 n_2} = 0 \end{cases} \tag{C2}$$

then the QNMF $\mathbf{X} = \mathbf{WH}$ is essentially unique.

Proof. See Appendix A-A. $\square$

On the one hand, condition (C2) is identical to the one found for the activation factor in the standard NMF case [41], [43], [44]. 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 [41], [43], [44], 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 condition (CS1')

$$\begin{cases} \Phi_{m1} = \Phi_{m2} = 1, \ \boldsymbol{\mu}_{m1} \neq \boldsymbol{\mu}_{m2} \\ I_{m1} > 0, \ I_{m2} > 0 \end{cases} . \tag{CS1'}$$

In other words, if there exist m such that both sources exhibit some arbitrary energy and are fully polarized with different polarization axes and if (C2) is satisfied, then the QNMF $\mathbf{X} = \mathbf{WH}$ is essentially unique.

The sufficient condition given in Proposition 3 for the uniqueness of QNMF in the two source case appears remarkably broad compared to the NMF case [41]. By incorporating polarization information and its associated constraints, QNMF permits to achieve uniqueness even when the sources never vanish – a typical case where NMF is known to be not unique [41], [44], [48].

C. General case ($P \geq 2$)

The study of the uniqueness of QNMF $\mathbf{X} = \mathbf{WH}$ for an arbitrary number of P sources is much more cumbersome than the $P = 2$ sources case, as the previous explicit modeling of indeterminacies is no longer amenable. However, the relationship between QNMF and NMF (see Section III-B) allows to formulate a necessary condition for uniqueness where standard NMF would fail.

Uniqueness of the NMF $\mathrm{Re}\, \mathbf{X} = [\mathrm{Re}\, \mathbf{W}] \mathbf{H}$ requires each column of $\mathrm{Re}\, \mathbf{W}$ to contain at least one entry equal to zero[2] [41], [43], [44]. Unfortunately, this condition is often violated, notably in hyperspectral imaging applications, and thus much effort has been devoted to the design of suitable optimization criteria [45], [46], [48], [49] to recover uniqueness of the (algorithmic) solution.

In contrast, QNMF relaxes the NMF zero-entry necessary condition on $\mathbf{W}$, stated by Proposition 4 below.

Proposition 4 (Necessary condition for uniqueness with non–vanishing sources). *Suppose that the QNMF $\mathbf{X} = \mathbf{WH}$ is essentially unique such that $\mathrm{Re}\, w_{mp} > 0$ for every m, p. Then the following conditions are satisfied:*
- $\forall (p, q), \ p \neq q,$

$$\exists m \ s.t. \ \Phi_{mp} = 1, \ \Phi_{mq} \boldsymbol{\mu}_{mq} \neq \boldsymbol{\mu}_{mp} \tag{A1}$$

- $\forall (p, q), \ p \neq q,$

$$\exists n \ s.t. \ h_{pn} = 0 \ and \ h_{qn} > 0 \tag{A2}$$

Proof. See Appendix A-B. $\square$

Proposition 4 shows that a necessary condition for uniqueness of the QNMF 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

---

[2]Note that requiring some entries of $\mathrm{Re}\, \mathbf{W}$ to be zero is equivalent to require the corresponding entries of $\mathbf{W}$ to be zero due to the nature of $(\mathcal{S})$.

be fully 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 [41], [43], [44].

In our opinion, the result stated by Proposition 4 is particularly interesting: it suggests that, in the general case of $P \geq 2$ sources, there may exist sufficient uniqueness conditions for QNMF that do not require at all the source factor $\mathbf{W}$ to contains zero-entries, unlike NMF. This conjecture is motivated by Proposition 3, which provides such a sufficient condition for the $P = 2$ sources case. It is also supported by our numerical experiments in Section VI-B, where QNMF uniqueness is observed on a synthetic, fully polarized spectro-polarimetric dataset with $P = 3$ sources. QNMF might offer a convenient framework to deal with situations where NMF alone does not guarantee uniqueness, such as with non-vanishing sources in hyperspectral imaging. This motivates us to search for general sufficient uniqueness conditions in future work.

V. AN ALGORITHM FOR QNMF

This section deals with the practical resolution of the QNMF problem. For a given value P, its resolution can be seen as an optimization procedure,

$$\min_{\substack{\mathbf{W} \in \mathbb{H}_{\mathcal{S}}^{M \times P} \\ \mathbf{H} \in \mathbb{R}_{+}^{P \times N}}} D(\mathbf{X}, \mathbf{W}\mathbf{H}) \quad (35)$$

where $D : \mathbb{H}^{M \times N} \times \mathbb{H}^{M \times N} \to \mathbb{R}_{+}$ can be an arbitary cost function for quaternion matrices. For simplicity, we choose here the Frobenius distance between quaternion matrices, referred to as the Euclidean cost in the sequel:

$$\begin{aligned} D(\mathbf{X}, \mathbf{W}\mathbf{H}) &= \|\mathbf{X} - \mathbf{W}\mathbf{H}\|_F^2 \\ &= \sum_{m,n} |X_{mn} - (\mathbf{W}\mathbf{H})_{mn}|^2 . \end{aligned} \quad (36)$$

The generic QNMF optimization problem (35) becomes

$$\min_{\substack{\mathbf{W} \in \mathbb{H}_{\mathcal{S}}^{M \times P} \\ \mathbf{H} \in \mathbb{R}_{+}^{P \times N}}} \|\mathbf{X} - \mathbf{W}\mathbf{H}\|_F^2 . \quad (37)$$

The formulation of the resolution of the QNMF problem (35) appears much alike the standard NMF with Euclidean cost. However, two fundamental questions need to be answered: *(i)* is the constraint $(\mathcal{S})$ easy to implement? and *(ii)* can we optimize w.r.t. $\mathbf{W}$ directly in the quaternion domain?

Fortunately, the answers to these two questions are affirmative. For *(i)*, the answer relies on the key link between $(\mathcal{S})$ 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 [50]–[53] – the so-called *generalized $\mathbb{H}\mathbb{R}$ calculus*.

Section V-A below explains how constraints on $\mathbf{W}$ and $\mathbf{H}$ factors are implemented. Section V-B describes the proposed alternating least squares strategy to solve (37). For completeness, Appendix B-A provides a quick introduction to quaternion optimization. Detailed computation of QNMF updates are found in Appendices B-C and B-B.

A. Projections onto constraints

Projections onto the constraint sets of $\mathbf{W}$ and $\mathbf{H}$ are a cornerstone of any method attempting to solve numerically the QNMF factorization problem (35) or (37). Projection operators onto constraints $\mathbb{R}_{+}$ and $(\mathcal{S})$ are denoted by $\Pi_{\mathbb{R}_{+}}$ and $\Pi_{\mathbb{H}_{\mathcal{S}}}$, respectively. The projection $\Pi_{\mathbb{R}_{+}}$ is classical [54], [55]:

$$[\Pi_{\mathbb{R}+}(\mathbf{M})]_{pn} = \max(0, (\mathbf{M})_{pn}) \quad (38)$$

where $\mathbf{M} \in \mathbb{R}^{P \times N}$, $1 \leq p \leq P$, $1 \leq n \leq N$.

To compute $\Pi_{\mathbb{H}_{\mathcal{S}}}$, we use the fact that the constraint $(\mathcal{S})$ can be derived from non-negativity (*i.e.* semi-definiteness) of 2-by-2 complex Hermitian matrices – as described in Section II-D. Consider an arbitrary matrix $\mathbf{M} \in \mathbb{H}^{M \times P}$ with (m, p) coefficient $(\mathbf{M})_{mp} = M_{mp} = \text{Re}\,M_{mp} + \boldsymbol{i}\text{Im}_{\boldsymbol{i}}\,M_{mp} + \boldsymbol{j}\text{Im}_{\boldsymbol{j}}\,M_{mp} + \boldsymbol{k}\text{Im}_{\boldsymbol{k}}\,M_{mp}$. Then define the following bijective mapping $f : \mathbb{H} \to \mathbb{C}^{2 \times 2}$ such that

$$\begin{aligned} &f(M_{mp}) \\ &\triangleq \frac{1}{2} \begin{bmatrix} \text{Re}\,M_{mp} + \text{Im}_{\boldsymbol{j}}\,M_{mp} & \text{Im}_{\boldsymbol{k}}\,M_{mp} + \boldsymbol{i}\text{Im}_{\boldsymbol{i}}\,M_{mp} \\ \text{Im}_{\boldsymbol{k}}\,M_{mp} - \boldsymbol{i}\text{Im}_{\boldsymbol{i}}\,M_{mp} & \text{Re}\,M_{mp} - \text{Im}_{\boldsymbol{j}}\,M_{mp} \end{bmatrix} . \end{aligned} \quad (39)$$

By construction, $f(M_{mp})$ is an Hermitian matrix. Its projection onto the set of non-negative Hermitian matrices $\mathcal{K}_{+}^2$ is well-known [56, p. 399], *i.e.*

$$\Pi_{\mathcal{K}_{+}^2}[f(M_{mp})] = \sum_{i=1}^{2} \max(0, \eta_i)\mathbf{v}_i\mathbf{v}_i^{\dagger} \quad (40)$$

where η_i and $\mathbf{v}_i$ are the i-th eigenvalue and eigenvector of the matrix $f(M_{mp})$, respectively. Since $\Pi_{\mathcal{K}_{+}^2}[f(M_{mp})]$ is a 2-by-2 non-negative Hermitian matrix, it can be expressed as

$$\Pi_{\mathcal{K}_{+}^2}[f(M_{mp})] = \begin{bmatrix} a & c \\ \bar{c} & b \end{bmatrix} \quad (41)$$

where $a \geq 0$, $b \geq 0$ and $c \in \mathbb{C}$ such that $ab - |c|^2 \geq 0$ on account of positive semi-definiteness of $\Pi_{\mathcal{K}_{+}^2}[f(M_{mp})]$. Remark that (41) is of the form (39): since the mapping f is bijective, the projection $[\Pi_{\mathbb{H}_{\mathcal{S}}}(\mathbf{M})]_{mp}$ is uniquely given by

$$[\Pi_{\mathbb{H}_{\mathcal{S}}}(\mathbf{M})]_{mp} = a + b + 2\boldsymbol{i}\text{Im}\,c + \boldsymbol{j}(a - b) + 2\boldsymbol{k}\text{Re}\,c , \quad (42)$$

where a, b and c as defined above.

B. Proposed algorithm: quaternion alternating least squares

The proposed algorithm adopts a popular strategy for solving the QNMF problem, based on the alternating constrained minimization of (37) w.r.t. $\mathbf{H}$ and $\mathbf{W}$. This choice is motivated by the fact that, whereas the Euclidean cost (37) is not convex in both $\mathbf{W}$ and $\mathbf{H}$, it is convex in each variable separately. After initialization of the factors, the iteration $r > 0$ reads

$$\mathbf{H}_{r+1} \leftarrow \operatorname*{arg\,min}_{\mathbf{H} \in \mathbb{R}_{+}^{P \times N}} \|\mathbf{X} - \mathbf{W}_r\mathbf{H}\|_F^2 \quad (43)$$

$$\mathbf{W}_{r+1} \leftarrow \operatorname*{arg\,min}_{\mathbf{W} \in \mathbb{H}_{\mathcal{S}}^{M \times P}} \|\mathbf{X} - \mathbf{W}\mathbf{H}_{r+1}\|_F^2 . \quad (44)$$

Equations (43)-(44) describe a two-block coordinate descent (CD) scheme in the quaternion domain. Two-block real-domain CD schemes are known to converge to a stationary

point under mild conditions [57]–[59]. These convergence results can be readily transposed to the quaternion case by identifying $\mathbb{H}$ with $\mathbb{R}^4$. Formally, since $\mathbb{R}^+$ and $\mathbb{H}_{\mathcal{S}}$ are closed convex sets, every limit point of the sequence $(\mathbf{W}_r, \mathbf{H}_r)$ generated by the two-block CD framework (43)-(44) is a stationary point of (37).

Convergence guarantees of the two-block CD scheme require to solve exactly both subproblems (43) and (44). However, the associated constraints make it difficult to obtain these updates directly in a closed form. The quaternion nature of the data $\mathbf{X}$ and source factor $\mathbf{W}$ adds a supplementary degree of difficulty due to non-commutativity of quaternion multiplication. Solving exactly each subproblem (43)-(44) requires to properly extend constrained optimization algorithms over the set of quaternions, which is beyond the scope of this paper.

Instead, we propose a simpler, closed-form algorithm that approximately solves each subproblem at each iteration, at the price of losing convergence guarantees. The main idea is the following: at a given iteration r, for each factor, one solves the unconstrained least-squares problem and projects the obtained solution onto the corresponding constraint:

$$\mathbf{H}_{r+1} \leftarrow \Pi_{\mathbb{R}_+} \left[\arg\min_{\mathbf{H}} \|\mathbf{X} - \mathbf{W}_r \mathbf{H}\|_F^2 \right] \tag{45}$$

$$\mathbf{W}_{r+1} \leftarrow \Pi_{\mathbb{H}_{\mathcal{S}}} \left[\arg\min_{\mathbf{W}} \|\mathbf{X} - \mathbf{W}\mathbf{H}_{r+1}\|_F^2 \right] . \tag{46}$$

Due to its resemblance with the usual alternating least squares (ALS) algorithm [60] for NMF, we call this strategy *quaternion alternating least squares* (QALS). Derivation of explicit updates for unconstrained least-squares problems requires special care because of the quaternion nature of $\mathbf{X}$ and $\mathbf{W}$. 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:

$$\mathbf{H}_{r+1} \leftarrow \Pi_{\mathbb{R}_+} \left[\left(\mathrm{Re} \left[\mathbf{W}_r^\top \overline{\mathbf{W}}_r \right] \right)^{-1} \mathrm{Re} \left[\mathbf{W}_r^\top \overline{\mathbf{X}} \right] \right] \tag{47}$$

$$\mathbf{W}_{r+1} \leftarrow \Pi_{\mathbb{H}_{\mathcal{S}}} \left[\mathbf{X}\mathbf{H}_{r+1}^\top \left(\mathbf{H}_{r+1}\mathbf{H}_{r+1}^\top \right)^{-1} \right] . \tag{48}$$

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 the standard ALS algorithm for NMF [61]. 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 [50]–[53]. Despite the lack of theoretical guarantees on its convergence, the quaternion ALS algorithm provides a good baseline for the resolution of QNMF with reasonably good results in most situations. Thus, it paves the way to further developments of more sophisticated algorithms for the QNMF problem.

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 QALS algorithm to solve the QNMF problem.

A. Polarization information and identifiability

This section illustrates the theoretical study provided in Section IV-B regarding the identifiability of QNMF in the $P = 2$ sources case. To emphasize how polarization impacts the range of admissible solutions, we choose a setting where the QNMF $\mathbf{X} = \mathbf{W}\mathbf{H}$ is not unique – and thus, neither is the associated NMF $\mathrm{Re}\,\mathbf{X} = \mathrm{Re}\,[\mathbf{W}]\,\mathbf{H}$, see (10).

Both sources and associated activations are non-vanishing, *i.e.* for all m, n, one has $w_{m1}, w_{m2}, h_{1n}, h_{2n} \neq 0$. Intensity responses $\mathrm{Re}\,w_{m1}, \mathrm{Re}\,w_{m2}$ encode realistic S_0 parameters, obtained for instance in hyperspectral imaging of wood components [62]. To simplify, both sources are chosen to exhibit constant polarization properties:

$$\Phi_{m1} = 0.7, \quad \boldsymbol{\mu}_{m1} = 0.87\boldsymbol{i} - 0.25\boldsymbol{j} - 0.43\boldsymbol{k} \tag{49}$$

$$\Phi_{m2} = 0.5, \quad \boldsymbol{\mu}_{m2} = -0.71\boldsymbol{i} + 0.44\boldsymbol{j} + 0.55\boldsymbol{k} . \tag{50}$$

Each source being partially polarized, it precludes the QNMF to be unique, since the necessary conditions stated by Proposition 4 are not fulfilled.

Numerical computations yield the range of admissible solutions for the QNMF and its associated NMF problem. Following Section IV-B, the range of admissible parameters α and β defining the transformation matrix $\mathbf{T}(\alpha, \beta)$ are found for the QNMF

$$\begin{aligned} \mathcal{I}_\alpha^{\mathrm{QNMF}} &= [-1.494 \cdot 10^{-1}, 7.692 \cdot 10^{-2}] \\ \mathcal{I}_\beta^{\mathrm{QNMF}} &= [-3.719 \cdot 10^{-1}, 1.538 \cdot 10^{-1}] \end{aligned} \tag{51}$$

and for the associated NMF problem:

$$\begin{aligned} \mathcal{I}_\alpha^{\mathrm{NMF}} &= [-1.799, 7.692 \cdot 10^{-2}] \\ \mathcal{I}_\beta^{\mathrm{NMF}} &= [-1.303 \cdot 10^1, 1.538 \cdot 10^{-1}] . \end{aligned} \tag{52}$$

Note that the QNMF and NMF intervals share the same upper bounds on α and β, which arise from the activation factor ratios (30).

Fig. 1 represents the set of admissible solutions corresponding to these NMF and QNMF intervals, respectively. Thick lines indicate the original sources and activations factors. Comparing respective ranges of solutions for S_0, it appears that QNMF significantly improves identifiability over standard NMF, by taking advantage of polarization information. Improvements are found on both S_0 and activations factors, and importantly, QNMF permits to reconstruct Stokes parameters S_1, S_2, S_3 with limited uncertainty. These results illustrate how QNMF takes advantage of the strong discriminative power of polarization in order to improve model identifiability.
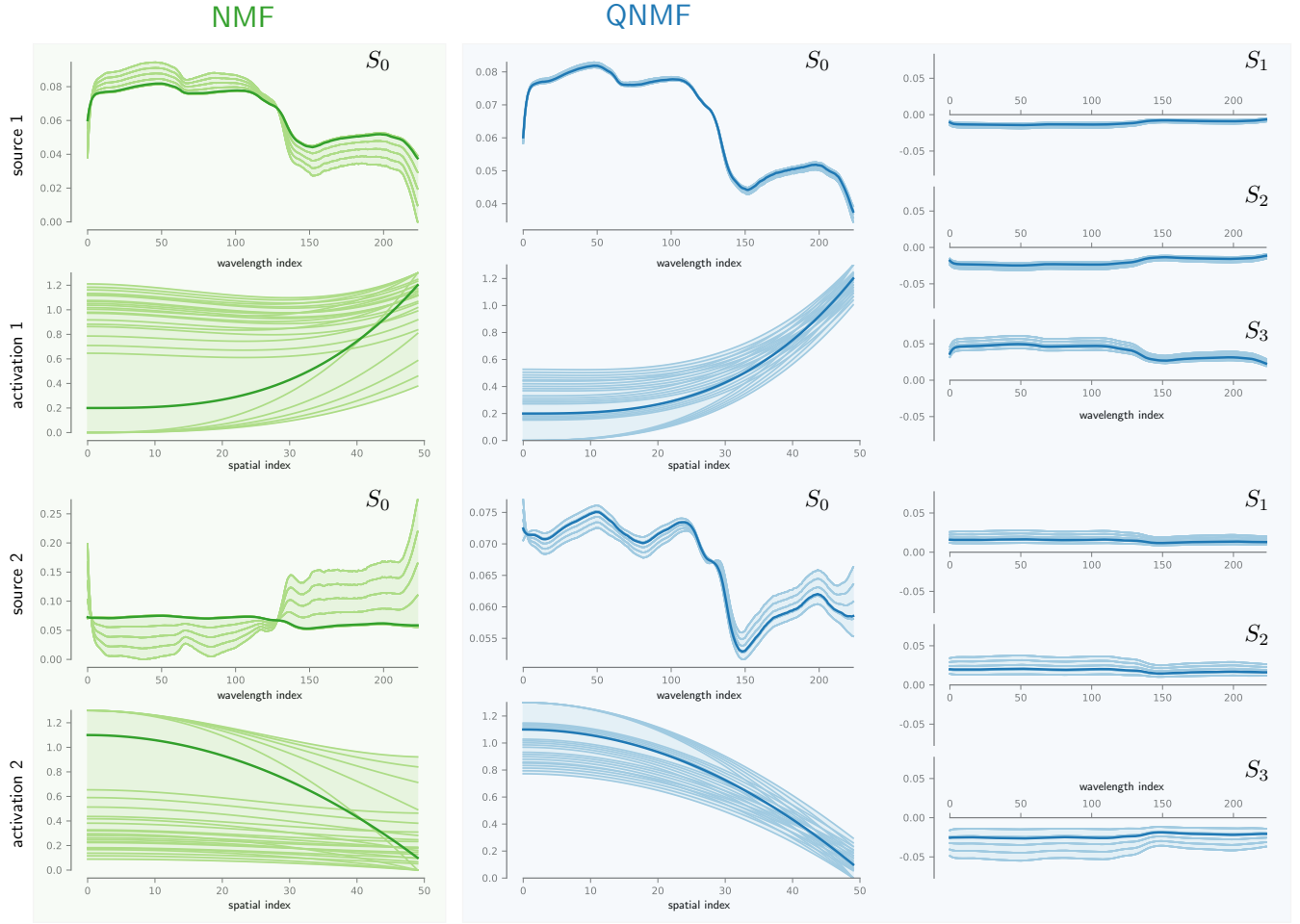
Fig. 1. Range of admissible solutions for the QNMF $\mathbf{X} = \mathbf{WH}$ and the associated NMF $\text{Re}\,\mathbf{X} = [\text{Re}\,\mathbf{W}]\,\mathbf{H}$. Taking advantage of polarization information with QNMF significantly improves model identifiability by shrinking the range of admissible solutions, compared to the sole use of intensity (S_0) information in the standard NMF approach.

B. Spectro-polarimetric blind unmixing using QALS

This section provides a first numerical validation of the QALS algorithm on realistic simulated spectro-polarimetric data. It consists in a linear mixture of $P = 3$ sources imaged at $M = 128$ wavelengths and $N = 512 \times 512$ spatial locations. Each source exhibits broad-band behavior, featuring non-vanishing intensity S_0 spectral profile together with spectrally varying polarization $S_1/S_0, S_2/S_0, S_3/S_0$ parameters. For simplicity, we assume that each source is fully polarized. Two-dimensional activations maps of each source are designed to represent a typical spectro-polarimetric imaging setting, *e.g.* solar spectro-polarimetry [7], [63], and such that the resulting activation matrix $\mathbf{H}$ satisfies the separability condition [44]. As a result, the designed $\mathbf{W}$ and $\mathbf{H}$ matrices satisfy the necessary conditions for QNMF identifiability stated in Proposition 4. Fig. 2A shows simulated spectro-polarimetric data $\mathbf{X} = \mathbf{WH}$ for two distinct wavelengths indices $m = 38$ and $m = 76$.

We show that QALS permits to effectively recover the QNMF factors. We adopt a random initialization strategy. The initial activation $\mathbf{H}_0$ is chosen as a matrix of i.i.d. entries drawn from $\mathcal{U}([0,1])$, the uniform distribution on $[0,1]$. The

initial source factor $\mathbf{W}_0$ has i.i.d. entries drawn from the quaternion circular unit Gaussian distribution [64] projected onto the constraint $(\mathcal{S})$. We then alternate Eqs. (47)–(48) until convergence. The latter was assessed by monitoring the relative error $\varepsilon_r = \|\mathbf{X} - \mathbf{W}_r\mathbf{H}_r\|_F^2/\|\mathbf{X}\|_F^2$ and stopping whenever the improvement was below a given threshold, *i.e.* $|\varepsilon_r - \varepsilon_{r-1}| \le 10^{-5}$ in our case. Repeating this procedure for $K = 100$ independent initializations of the QALS algorithm, it took on average 29 iterations for the algorithm to converge. Out of 100 initializations, we observed that the QALS algorithm recovered each time identical sources and activation factors, strongly indicating that the QNMF of the data $\mathbf{X}$ may be essentially unique.

Fig. 2B depicts the 3 sources and corresponding activations factors recovered by the QALS algorithm for one arbitrary initialization. Excellent match between reconstructed and ground truth factors is indicated by very small relative errors, *i.e.* $\varepsilon_{\mathbf{H}} = \|\hat{\mathbf{H}} - \mathbf{H}\|_F^2/\|\mathbf{H}\|_F^2 = 4.91 \cdot 10^{-5}$ and $\varepsilon_{\mathbf{W}} = \|\hat{\mathbf{W}} - \mathbf{W}\|_F^2/\|\mathbf{W}\|_F^2 = 3.92 \cdot 10^{-5}$. Looking at sources intensities S_0, we observe that they never vanish, *i.e.* $S_0 > 0$, plus that the first and second sources exhibit very similar
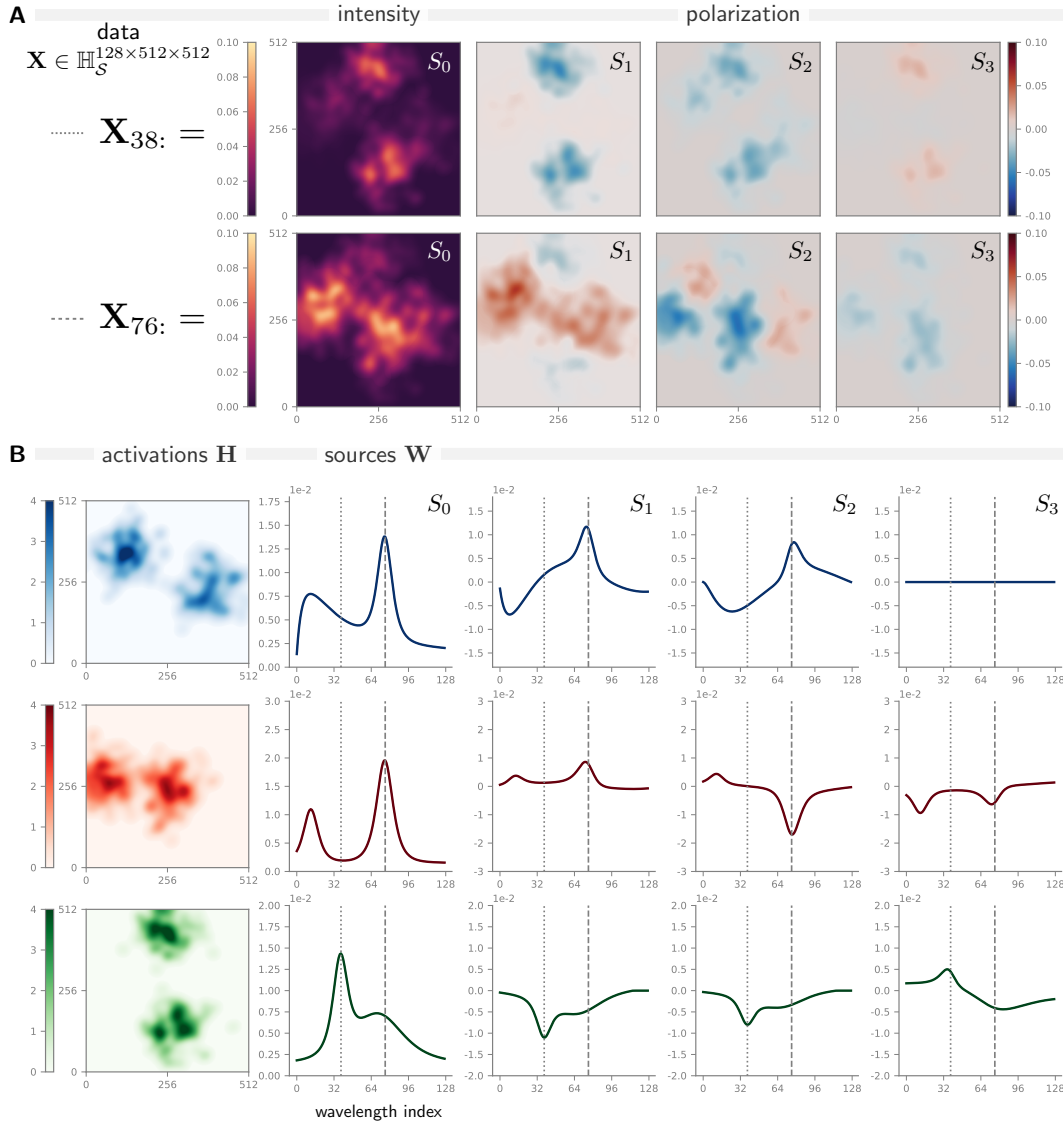
Fig. 2. Blind source separation of spectro-polarimetric data using QALS. **A** 2D intensity (S_0) and polarization (S_1, S_2, S_3) maps for two wavelength indices $m = 38$ and $m = 76$, corresponding to local intensity maxima of sources. **B** reconstructed 2D activations maps and sources Stokes spectral profiles for each of the three factors. Dotted and dashed lines correspond to wavelength indices $m = 38$ and $m = 76$, respectively.

peaks. This is a very challenging situation for NMF-based methods, which would require additional regularizations such as minimum volume [45]–[47] to ensure uniqueness of the solution. It is thus remarkable that the QNMF model with the QALS algorithm is able to solve this blind source problem in a simple manner by leveraging polarization information. One of the advantages of polarization in source separation problems can be directly seen by inspecting $\mathbf{X}_{76:}$ in Fig. 2A: while it is not possible to distinguish 2 sources using S_0 only, the distinctive polarization pattern on S_2 makes it easy to identify source 1 and source 2.

These simulation results demonstrate the ability of the proposed QALS algorithm to effectively solve the QNMF problem. Despite no associated provable convergence guarantees, QALS offers a simple, computationally efficient algorithm and provides a baseline for the development of more advanced algorithms for QNMF.

VII. CONCLUSION

This paper has introduced a new powerful tool called quaternion non-negative matrix factorization (QNMF), 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. We unveiled the relation between NMF and QNMF, showing that NMF sufficient conditions are also sufficient uniqueness conditions for QNMF. Remarkably, we have shown that QNMF improves model identifiability over NMF by leveraging the complete information available in the four Stokes parameters. This key role played by polarization information is stated by Proposition 3 for the two sources case and by Proposition 4 for the general case. Furthermore,

taking advantage of recent results in quaternion optimization, we have proposed a simple yet efficient algorithm for solving the QNMF problem in practice. These first results appear very promising and pave the way to future work on both theoretical and methodological aspects of QNMF. In particular, it is an open question whether one can generalize to the QNMF case the notion of non-negative rank [65] and its related properties. This fundamental question is intimately related to the definition of a polarized source; as such, it would greatly improve our understanding of the role played by polarization diversity in blind source separation problems.

APPENDIX A
UNIQUENESS CONDITIONS FOR QNMF

A. Proof of Proposition 3

Suppose that $P = 2$ and that the QNMF $\mathbf{X} = \mathbf{WH}$ exists. We suppose that conditions (C1) and (C2) are satisfied. To show that the QNMF is identifiable, we prove that intervals $\mathcal{I}_\alpha^{\text{QNMF}}$ and $\mathcal{I}_\beta^{\text{QNMF}}$ defined in (26)-(27) are restricted to $\{0\}$.

Condition (C1) permits to simplify (33) and (34) for m_1, m_2 like

$$
\begin{aligned}
&\alpha^2 \left[I_{m_1 2}^2 (1 - \Phi_{m_1 2}^2) - 2 I_{m_1 1} I_{m_1 2} (1 - \Phi_{m_1 2} \langle \boldsymbol{\mu}_{m_1 1}, \boldsymbol{\mu}_{m_1 2} \rangle) \right] \\
&+ 2\alpha \left[I_{m_1 1} I_{m_1 2} (1 - \Phi_{m_1 2} \langle \boldsymbol{\mu}_{m_1 1}, \boldsymbol{\mu}_{m_1 2} \rangle) \right] \geq 0 \quad (53)
\end{aligned}
$$

$$
\begin{aligned}
&\beta^2 \left[I_{m_2 1}^2 (1 - \Phi_{m_2 1}^2) - 2 I_{m_2 1} I_{m_2 2} (1 - \Phi_{m_2 1} \langle \boldsymbol{\mu}_{m_2 1}, \boldsymbol{\mu}_{m_2 2} \rangle) \right] \\
&+ 2\beta \left[I_{m_2 1} I_{m_2 2} (1 - \Phi_{m_2 1} \Phi_{m_2 2} \langle \boldsymbol{\mu}_{m_2 1}, \boldsymbol{\mu}_{m_2 2} \rangle) \right] \geq 0 . \quad (54)
\end{aligned}
$$

Using inequalities linking I_{m1}, I_{m2} for $m = m_1, m_2$ in (C1) yields the following domains of solutions to (53) and (54):

$$
\mathcal{I}_{\alpha, m_1}^{\text{pol}} = [0, \alpha_0] \quad \mathcal{I}_{\beta, m_2}^{\text{pol}} = [0, \beta_0], \quad (55)
$$

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

$$
\mathcal{I}_\alpha^{\text{NMF}} = [\alpha_0', 0] \quad \mathcal{I}_\beta^{\text{NMF}} = [\beta_0', 0], \quad (56)
$$

where $\alpha_0', \beta_0' \leq 0$. By intersection of intervals, one gets that $\mathcal{I}_\alpha^{\text{QNMF}}$ and $\mathcal{I}_\beta^{\text{QNMF}}$ defined in (26)-(27) are restricted to $\{0\}$, so that the QNMF is unique.

B. Proof of Proposition 4

Suppose that the QNMF $\mathbf{X} = \mathbf{WH}$ exists and that sources never vanish, i.e. $\operatorname{Re} w_{mp} > 0$ 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 $\mathbf{T}$ leading to different factors $\widetilde{\mathbf{W}} = \mathbf{WT}$ and $\widetilde{\mathbf{H}} = \mathbf{T}^{-1}\mathbf{H}$.

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

$$
\forall m, \ (\Phi_{m p_0} \in [0, 1) \text{ or } \Phi_{m q_0} \boldsymbol{\mu}_{m q_0} = \boldsymbol{\mu}_{m p_0}) , \quad (57)
$$

or equivalently,

$$
\forall m, \ [\Phi_{m p_0} \in [0, 1) \text{ or } (\Phi_{m q_0} = 1 \text{ and } \boldsymbol{\mu}_{m q_0} = \boldsymbol{\mu}_{m p_0})] , \quad (58)
$$

Now consider the transformation $\mathbf{T}_{p_0 q_0}^\alpha \in \mathbb{R}^{P \times P}$ defined by

$$
\forall k, \ell \in \{1, 2, \ldots, P\}, \begin{cases} \left(\mathbf{T}_{p_0 q_0}^\alpha \right)_{kk} & = 1 \\ \left(\mathbf{T}_{p_0 q_0}^\alpha \right)_{p_0 q_0} & = -\alpha \\ \left(\mathbf{T}_{p_0 q_0}^\alpha \right)_{k\ell} & = 0 \text{ otherwise} \end{cases} . \quad (59)
$$

By construction $\mathbf{T}_{p_0 q_0}^\alpha$ does not correspond to a trivial ambiguity of the QNMF when $\alpha \neq 0$. Note that $(\mathbf{T}_{p_0 q_0}^\alpha)^{-1} = \mathbf{T}_{p_0 q_0}^{-\alpha}$. Consider the new factors $\widetilde{\mathbf{W}} = \mathbf{WT}_{p_0 q_0}^\alpha$ and $\widetilde{\mathbf{H}} = \mathbf{T}_{p_0 q_0}^{-\alpha} \mathbf{H}$ such that $\mathbf{X} = \widetilde{\mathbf{W}}\widetilde{\mathbf{H}}$. From (59), the source indexed by p_0 is the only one affected by $\mathbf{T}_{p_0 q_0}^\alpha$ like

$$
\forall m, \ \widetilde{w}_{m p_0} = w_{m p_0} - \alpha w_{m q_0} . \quad (60)
$$

The corresponding activation coefficients read

$$
\forall n, \ \widetilde{h}_{p_0 n} = h_{p_0 n} + \alpha h_{q_0 n} . \quad (61)
$$

Supposing $\alpha > 0$, then $\widetilde{h}_{p_0 n} \geq 0$ for every n by non-negativity of the matrix $\mathbf{H}$. It remains to find at least one $\alpha > 0$ such that $\forall m, \ \widetilde{w}_{m p_0} = w_{m p_0} - \alpha w_{m q_0} \in \mathbb{H}_\mathcal{S}$. Imposing $\widetilde{w}_{m p_0} \in \mathbb{H}_\mathcal{S}$ for every m, yields

$$
\forall m, \begin{cases} I_{m p_0} - \alpha I_{m q_0} & \geq 0 \\ |I_{m p_0} \Phi_{m p_0} \boldsymbol{\mu}_{m p_0} - \alpha I_{m q_0} \Phi_{m q_0} \boldsymbol{\mu}_{m q_0}|^2 & \leq (I_{m p_0} - \alpha I_{m q_0})^2 \end{cases} \quad (62)
$$

where we used notation (17) for convenience. Since by assumption $I_{m p_0}, I_{q_0} > 0$, the first condition implies that $\alpha \leq \min_m (I_{m p_0} / I_{m q_0})$. The second condition related to polarization can be rewritten as, for every m

$$
\begin{aligned}
&\alpha^2 I_{m q_0}^2 (1 - \Phi_{m q_0}^2) \\
&- 2\alpha I_{m p_0} I_{m q_0} (1 - \Phi_{m p_0} \Phi_{m q_0} \langle \boldsymbol{\mu}_{m p_0}, \boldsymbol{\mu}_{m q_0} \rangle) \quad (63) \\
&+ I_{m p_0}^2 (1 - \Phi_{m p_0}^2) \geq 0 .
\end{aligned}
$$

According to (58) two cases may occur, not exclusively from each other. Fix m and suppose that $\Phi_{m p_0} \in [0, 1)$. Two cases are possible: either $\Phi_{m q_0} = 1$ or $\Phi_{m q_0} \in [0, 1)$. Assume that $\Phi_{m q_0} = 1$, then (63) is true for $\alpha \leq I_{m p_0} (1 - \Phi_{m p_0}^2) I_{m q_0}^{-1} ((1 - \Phi_{m p_0} \langle \boldsymbol{\mu}_{m p_0}, \boldsymbol{\mu}_{m q_0} \rangle)^{-1}$. For $\Phi_{m q_0} \in [0, 1)$, since $I_{m q_0}^2 (1 - \Phi_{m q_0}^2) > 0$ in virtue of assumptions, the discriminant Δ reads

$$
\begin{aligned}
\Delta = 4 I_{m p_0}^2 I_{m q_0}^2 &\left[(1 - \Phi_{m p_0} \Phi_{m q_0} \langle \boldsymbol{\mu}_{m p_0}, \boldsymbol{\mu}_{m q_0} \rangle)^2 \right. \\
&\left. - (1 - \Phi_{m p_0}^2)(1 - \Phi_{m q_0}^2) \right] .
\end{aligned} \quad (64)
$$

Ranges of parameters imply that $\Delta \geq 0$. As a result, (63) is true for $\alpha \in (-\infty, \alpha_-^m] \cup [\alpha_+^m, +\infty)$ where $\alpha_\pm^m$ are the roots of the polynomial. Sign of polynomial coefficients in (63) together with $0 \leq \Phi_{m p_0} < 1$ imply that $\alpha_\pm^m > 0$. Suppose now that $\Phi_{m q_0} = 1$ and $\boldsymbol{\mu}_{m q_0} = \boldsymbol{\mu}_{m p_0}$. Taking $\Phi_{m p_0} = 1$ Eq. (63) becomes trivial. The case $\Phi_{m q_0} \in [0, 1)$ is included in the previous discussion.

Summarizing all cases, for every m such that $\Phi_{m p_0} \in [0, 1)$ or $\Phi_{m q_0} \boldsymbol{\mu}_{m q_0} = \boldsymbol{\mu}_{m p_0}$, there always exists $\alpha > 0$ such that conditions (62) are satisfied, meaning that the QNMF is not unique. This leads to the first necessary condition (A1). Repeating the same approach with $\alpha < 0$ yields condition (A2).

APPENDIX B
DERIVATION OF QUATERNION ALS UPDATES

A. Quaternion derivatives using generalized $\mathbb{HR}$-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 [66], [67] 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 [66], so that other strategies need to be deployed. First, a pseudo-derivative approach can be used by treating $f(q)$ as a function $f(q_0, q_1, q_2, q_3)$ of the four real components of q. As pointed out in [50], such approach requires lengthy and cumbersome computations, thus limiting its applicatibility. In order to compute derivatives directly in the quaternion domain, crucial steps have been made recently with the development of the $\mathbb{HR}$-calculus [53], and subsequently the generalized $\mathbb{HR}$-calculus [52]. The latter provides a complete framework including chain rule and product rules, extending naturally the $\mathbb{CR}$-calculus [68] of complex-valued optimization to the case of quaternion functions.

For the quadratic-type of quaternion functions encountered in this paper, only important results related to quaternion optimization are necessary. These are given in Proposition 5 and Table I below. We refer to [50], [52] for a complete description of the generalized $\mathbb{HR}$-calculus.

Proposition 5 characterizes stationary points of cost functions of quaternion matrices.

Proposition 5 (Theorem 4.1 [50]). *Let $f : \mathbb{H}^{M \times N} \to \mathbb{R}$ and denote by $\nabla_{\mathbf{Q}} f$ its gradient w.r.t. $\mathbf{Q}$. Then,*

$$
\begin{aligned}
&\mathbf{Q}_0 \text{ is a stationary point} \\
\Leftrightarrow\ &\nabla_{\mathbf{Q}} f(\mathbf{Q}_0) = 0 \Leftrightarrow \nabla_{\overline{\mathbf{Q}}} f(\mathbf{Q}_0) = 0
\end{aligned} \tag{65}
$$

For practical computations, Table I reprints from [50] some useful functions of the quaternion matrix variable $\mathbf{Q}$ and their corresponding gradients.

B. Updates for $\mathbf{H}$

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

$$
\|\mathbf{X} - \mathbf{WH}\|_F^2 = \sum_{m,n} \left| x_{mn} - \sum_k w_{mk} h_{kn} \right|^2 \tag{66}
$$

$$
= \sum_{m,n} \left(x_{mn} - \sum_k w_{mk} h_{kn} \right) \left(\overline{x_{mn}} \sum_k \overline{w_{mk}} h_{kn} \right), \tag{67}
$$

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 (67)

TABLE I
MATRIX FUNCTIONS $f : \mathbb{H}^{M \times N} \to \mathbb{H}$ AND CORRESPONDING GRADIENTS. MATRICES $\mathbf{A}_1, \mathbf{A}_2$ ARE QUATERNION-VALUED WITH APPROPRIATE DIMENSIONS.

$f(\mathbf{Q})$	$\nabla_{\overline{\mathbf{Q}}} f$
$\operatorname{tr} \mathbf{A}_1 \mathbf{Q} \mathbf{A}_2$	$-\frac{1}{2} \mathbf{A}_1^\top \mathbf{A}_2^\dagger$
$\operatorname{tr} \mathbf{A}_1 \mathbf{Q}^\dagger \mathbf{A}_2$	$\operatorname{Re}[\mathbf{A}_2] \mathbf{A}_1$
$\operatorname{tr} \mathbf{A}_1 \mathbf{Q}^\dagger \mathbf{Q} \mathbf{A}_2$	$\operatorname{Re}[\mathbf{Q} \mathbf{A}_2] \mathbf{A}_1 - \frac{1}{2} (\mathbf{A}_1 \mathbf{Q}^\dagger)^\top \mathbf{A}_2^\dagger$

w.r.t. h_{ij} is

$$
\frac{\partial \|\mathbf{X} - \mathbf{WH}\|_F^2}{\partial h_{ij}}
$$

$$
\begin{aligned}
= &-\sum_{m,n} \frac{\partial \sum_k w_{mk} h_{kn}}{\partial h_{ij}} \left(\overline{x_{mn}} - \sum_k \overline{w_{mk}} h_{kn} \right) \\
&- \sum_{m,n} \left(w_{mn} - \sum_k w_{mk} h_{kn} \right) \frac{\partial \sum_k \overline{w_{mk}} h_{kn}}{\partial h_{ij}}
\end{aligned} \tag{68}
$$

$$
\begin{aligned}
= &-\sum_m w_{mi} \left(\overline{x_{mj}} - \sum_k \overline{w_{mk}} h_{kj} \right) \\
&- \sum_m \left(x_{mj} - \sum_k w_{mk} h_{kj} \right) \overline{w_{mi}}
\end{aligned} \tag{69}
$$

$$
\begin{aligned}
= &-\sum_m w_{mi} \left(\overline{x_{mj}} - \overline{(\mathbf{WH})_{mj}} \right) \\
&- \sum_m (x_{mj} - (\mathbf{WH})_{mj}) \overline{w_{mi}}
\end{aligned} \tag{70}
$$

$$
= -2\operatorname{Re} \left[\sum_m w_{mi} \left(\overline{x_{mj}} - \overline{(\mathbf{WH})_{mj}} \right) \right] \tag{71}
$$

$$
= -2\operatorname{Re} \left[\mathbf{W}^\top (\overline{\mathbf{X} - \mathbf{WH}}) \right]_{ij} . \tag{72}
$$

As a result, one gets

$$
\nabla_{\mathbf{H}} \|\mathbf{X} - \mathbf{WH}\|_F^2 = -2\operatorname{Re} \left[\mathbf{W}^\top (\overline{\mathbf{X} - \mathbf{WH}}) \right] . \tag{73}
$$

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

$$
-2\operatorname{Re} \left[\mathbf{W}^\top \overline{\mathbf{X}} \right] + 2\operatorname{Re} \left[\mathbf{W}^\top \overline{\mathbf{W}} \mathbf{H}_\star \right] = 0 , \tag{74}
$$

where again we used that $\overline{\mathbf{WH}_\star} = \overline{\mathbf{W}} \mathbf{H}_\star$ since $\mathbf{H}_\star$ is a real matrix. It also implies that $\operatorname{Re} \left[\mathbf{W}^\top \overline{\mathbf{W}} \mathbf{H}_\star \right] = \operatorname{Re} \left[\mathbf{W}^\top \overline{\mathbf{W}} \right] \mathbf{H}_\star$, leading to the following expression for $\mathbf{H}_\star$:

$$
\mathbf{H}_\star = \arg \min_{\mathbf{H}} \|\mathbf{X} - \mathbf{WH}\|_F^2 \tag{75}
$$

$$
= \left(\operatorname{Re} \left[\mathbf{W}^\top \overline{\mathbf{W}} \right] \right)^{-1} \operatorname{Re} \left[\mathbf{W}^\top \overline{\mathbf{X}} \right] . \tag{76}
$$

C. Updates for $\mathbf{W}$

Let us rewrite the Euclidean distance (37) in terms of a trace operator

$$
\|\mathbf{X} - \mathbf{WH}\|_F^2 = \operatorname{tr} \left[(\mathbf{X} - \mathbf{WH})^\dagger (\mathbf{X} - \mathbf{WH}) \right], \tag{77}
$$

where $\mathbf{Q}^{\dagger} = \overline{\mathbf{Q}}^{\top}$ is the (quaternion) conjugate-transpose of $\mathbf{Q}$. Developping[3] (77) one gets

$$\|\mathbf{X} - \mathbf{W}\mathbf{H}\|_F^2 = \operatorname{tr}\mathbf{X}^{\dagger}\mathbf{X} - \operatorname{tr}\mathbf{X}^{\dagger}\mathbf{W}\mathbf{H} \\ - \operatorname{tr}\mathbf{H}^{\dagger}\mathbf{W}^{\dagger}\mathbf{X} + \operatorname{tr}\mathbf{H}^{\dagger}\mathbf{W}^{\dagger}\mathbf{W}\mathbf{H} \,. \tag{78}$$

By Proposition 5, stationary points are defined by cancelling values of the gradient of (78) w.r.t. to $\overline{\mathbf{W}}$. Thanks to Table I one gets

$$\nabla_{\overline{\mathbf{W}}}\|\mathbf{X} - \mathbf{W}\mathbf{H}\|_F^2 = \frac{1}{2}\left(\mathbf{X}^{\dagger}\right)^{\top}\mathbf{H}^{\top} - \operatorname{Re}[\mathbf{X}]\mathbf{H}^{\top} \\ + \operatorname{Re}[\mathbf{W}\mathbf{H}]\mathbf{H}^{\top} - \frac{1}{2}\left(\mathbf{H}^{\top}\mathbf{W}^{\dagger}\right)^{\top}\mathbf{H}^{\top} \tag{79}$$

$$= -\left(\operatorname{Re}[\mathbf{X}] - \frac{1}{2}\overline{\mathbf{X}} + \frac{1}{2}\overline{\mathbf{W}\mathbf{H}} - \operatorname{Re}[\mathbf{W}\mathbf{H}]\right)\mathbf{H}^{\top} \tag{80}$$

$$= -\frac{1}{2}\left(\mathbf{X} - \mathbf{W}\mathbf{H}\right)\mathbf{H}^{\top} \,. \tag{81}$$

The Euclidean cost function is convex in $\mathbf{W}$. Thus, cancelling out the gradient (81) yields the global minimizer $\mathbf{W}_{\star}$ of this cost such that

$$-\mathbf{X}\mathbf{H}^{\top} + \mathbf{W}_{\star}\mathbf{H}\mathbf{H}^{\top} = 0 \tag{82}$$

so that

$$\mathbf{W}_{\star} = \arg\min_{\mathbf{W}}\|\mathbf{X} - \mathbf{W}\mathbf{H}\|_F^2 = \mathbf{X}\mathbf{H}^{\top}\left(\mathbf{H}\mathbf{H}^{\top}\right)^{-1} \,. \tag{83}$$

REFERENCES

[1] J. Flamant, S. Miron, and D. Brie, "Quaternion non-negative matrix factorization: a new tool for spectropolarimetric imaging," in *2019 IEEE 8th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, Guadeloupe, France, December 15-18 2019, pp. 1–5.

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

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

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

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

[6] 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.

[7] D. Orozco Suárez, L. Bellot Rubio, J. del Toro Iniesta, S. Tsuneta, B. Lites, K. Ichimoto, Y. Katsukawa, S. Nagata, T. Shimizu, R. Shine *et al.*, "Quiet-Sun internetwork magnetic fields from the inversion of Hinode measurements," *The Astrophysical Journal*, vol. 670, pp. L61–L64, 2007.

[8] N. Ghosh and A. I. Vitkin, "Tissue polarimetry: concepts, challenges, applications, and outlook," *Journal of biomedical optics*, vol. 16, no. 11, p. 110801, 2011.

[9] 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.

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

[11] 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.

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

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

[14] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM review*, vol. 51, no. 3, pp. 455–500, 2009.

[15] S. Javidi, C. C. Took, and D. P. Mandic, "Fast independent component analysis algorithm for quaternion valued signals," *IEEE transactions on neural networks*, vol. 22, no. 12, pp. 1967–1978, 2011.

[16] J. Vía, D. P. Palomar, L. Vielva, and I. Santamaría, "Quaternion ica from second-order statistics," *IEEE Transactions on Signal Processing*, vol. 59, no. 4, pp. 1586–1600, 2010.

[17] 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.

[18] J. Flamant, P. Chainais, and N. Le Bihan, "A complete framework for linear filtering of bivariate signals," *IEEE Transactions on Signal Processing*, vol. 66, no. 17, pp. 4541–4552, Sept 2018.

[19] J. Flamant, "A general approach for the analysis and filtering of bivariate signals," 2018. [Online]. Available: https://tel.archives-ouvertes.fr/tel-01926941/

[20] J. Flamant, N. Le Bihan, and P. Chainais, "Time-frequency analysis of bivariate signals," *Applied and Computational Harmonic Analysis*, vol. 46, no. 2, pp. 351 – 383, 2019.

[21] J. Flamant, N. Le Bihan, and P. Chainais, "Spectral analysis of stationary random bivariate signals," *IEEE Transactions on Signal Processing*, vol. 65, no. 23, pp. 6135–6145, 2017.

[22] J. Gonella, "A rotary-component method for analysing meteorological and oceanographic vector time series," in *Deep Sea Research and Oceanographic Abstracts*, vol. 19, no. 12. Elsevier, 1972, pp. 833–846.

[23] J. Flamant, P. Chainais, E. Chassande-Mottin, F. Feng, and N. Le Bihan, "Non-parametric characterization of gravitational-wave polarizations," in *26th European Signal Processing Conference (EUSIPCO)*, 2018.

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

[25] J. L. Brenner, "Matrices of quaternions." *Pacific Journal of Mathematics*, vol. 1, no. 3, pp. 329–335, 1951.

[26] 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. 253–260.

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

[28] N. Le Bihan and J. Mars, "Singular value decomposition of quaternion matrices: a new tool for vector-sensor signal processing," *Signal Processing*, vol. 84, no. 7, pp. 1177–1199, 2004.

[29] L. Rodman, *Topics in quaternion linear algebra*. Princeton University Press, 2014, vol. 45.

[30] C. Brosseau, *Fundamentals of polarized light: a statistical optics approach*. Wiley-Interscience, 1998.

[31] B. Schaefer, E. Collett, R. Smyth, D. Barrett, and B. Fraher, "Measuring the Stokes polarization parameters," *American Journal of Physics*, vol. 75, no. 2, pp. 163–168, 2007.

[32] H. G. Berry, G. Gabrielse, and A. Livingston, "Measurement of the Stokes parameters of light," *Applied optics*, vol. 16, no. 12, pp. 3200–3205, 1977.

[33] T. Tudor, "Vectorial Pauli algebraic approach in polarization optics. I. Device and state operators," *Optik - International Journal for Light and Electron Optics*, vol. 121, no. 13, pp. 1226–1235, 2010.

[34] ——, "Vectorial Pauli algebraic approach in polarization optics . II . Interaction of light with the canonical polarization devices," *Optik - International Journal for Light and Electron Optics*, vol. 121, no. 23, pp. 2149–2158, 2010.

[35] C. Whitney, "Pauli-Algebraic Operators in Polarization Optics," *Journal of the Optical Society of America*, vol. 61, no. 9, pp. 1207–1213, 1971.

[36] P. Pellat-Finet, "Représentation des états et des Opérateurs de Polarisation de la Lumière Par des Quaternions," *Optica Acta : International Journal of Optics*, vol. 31, pp. 415–434, 1984.

[37] M. Richartz and H.-Y. Hsü, "Analysis of Elliptical Polarization," *Journal of the Optical Society of America*, vol. 39, no. 2, p. 136, 1949.

[38] 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 journal of selected topics in applied earth observations and remote sensing*, vol. 5, no. 2, pp. 354–379, 2012.

[39] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE signal processing magazine*, vol. 19, no. 1, pp. 44–57, 2002.

---

[3]We use that, for any quaternion matrices $\mathbf{A}$ and $\mathbf{B}$, $(\mathbf{A}\mathbf{B})^{\dagger} = \mathbf{B}^{\dagger}\mathbf{A}^{\dagger}$. See e.g. [27] for more complete properties of quaternion matrices.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSP.2020.2974651, IEEE Transactions on Signal Processing

14

[40] 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.

[41] 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 International Conference on Acoustics, Speech, and Signal Processing*, vol. 5. IEEE, 2005, pp. 289–292.

[42] D. Donoho and V. Stodden, "When does non-negative matrix factorization give a correct decomposition into parts?" in *Advances in Neural Information Processing Systems*, 2004, pp. 1141–1148.

[43] H. Laurberg, M. G. Christensen, M. D. Plumbley, L. K. Hansen, and S. H. Jensen, "Theorems on positive data: On the uniqueness of NMF," *Computational intelligence and neuroscience*, vol. 2008, 2008.

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

[45] X. Fu, K. Huang, and N. D. Sidiropoulos, "On identifiability of non-negative matrix factorization," *IEEE Signal Processing Letters*, vol. 25, no. 3, pp. 328–332, 2018.

[46] C.-H. Lin, W.-K. Ma, W.-C. Li, C.-Y. Chi, and A. Ambikapathi, "Identifiability of the simplex volume minimization criterion for blind hyperspectral unmixing: The no-pure-pixel case," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 10, pp. 5530–5546, 2015.

[47] X. Fu, W.-K. Ma, K. Huang, and N. D. Sidiropoulos, "Blind separation of quasi-stationary sources: Exploiting convex geometry in covariance domain," *IEEE Transactions on Signal Processing*, vol. 63, no. 9, pp. 2306–2320, 2015.

[48] 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.

[49] M. D. Craig, "Minimum-volume transforms for remotely sensed data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 32, no. 3, pp. 542–552, 1994.

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

[51] 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.

[52] 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.

[53] 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.

[54] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, p. 788, 1999.

[55] ——, "Algorithms for non-negative matrix factorization," in *Advances in Neural Information Processing Systems*, 2001, pp. 556–562.

[56] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[57] D. P. Bertsekas, *Nonlinear programming*, 2nd ed. Athena Scientific, 1999.

[58] L. Grippo and M. Sciandrone, "On the convergence of the block nonlinear gauss–seidel method under convex constraints," *Operations research letters*, vol. 26, no. 3, pp. 127–136, 2000.

[59] J. Kim, Y. He, and H. Park, "Algorithms for nonnegative matrix and tensor factorizations: A unified view based on block coordinate descent framework," *Journal of Global Optimization*, vol. 58, no. 2, pp. 285–319, 2014.

[60] 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.

[61] 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.

[62] 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.

[63] P. I. McCauley, I. H. Cairns, S. M. White, S. Mondal, E. Lenc, J. Morgan, and D. Oberoi, "The low-frequency solar corona in circular polarization," *Solar Physics*, vol. 294, no. 8, p. 106, 2019.

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

[65] J. E. Cohen and U. G. Rothblum, "Nonnegative ranks, decompositions, and factorizations of nonnegative matrices," *Linear Algebra and its Applications*, vol. 190, pp. 149–168, 1993.

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

[67] R. Watson, "The generalized Cauchy-Riemann-Fueter equation and handedness," *Complex Variables*, vol. 48, no. 7, pp. 555–568, 2003.

[68] K. Kreutz-Delgado, "The complex gradient operator and the CR-calculus," *arXiv preprint arXiv:0906.4835*, 2009.