



**HAL**  
open science

## Multilinear analysis of quaternion arrays: theory and computation

Julien Flamant, Xavier Luciani, Sebastian Miron, Yassine Zniyed

### ► To cite this version:

Julien Flamant, Xavier Luciani, Sebastian Miron, Yassine Zniyed. Multilinear analysis of quaternion arrays: theory and computation. *SIAM Journal on Matrix Analysis and Applications*, 2026, 47 (1), pp.158-184. <10.1137/24M1715532>. <hal-04830078v2>

**HAL Id: hal-04830078**

**<https://hal.science/hal-04830078v2>**

Submitted on 16 Oct 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

# MULTILINEAR ANALYSIS OF QUATERNION ARRAYS: THEORY AND COMPUTATION\*

JULIEN FLAMANT<sup>†</sup>, XAVIER LUCIANI<sup>‡</sup>, SEBASTIAN MIRON<sup>†</sup>, AND YASSINE ZNIYED<sup>‡</sup>

**Abstract.** Multidimensional quaternion arrays (often referred to as “quaternion tensors”) and their decompositions have recently gained increasing attention in various fields such as color and polarimetric imaging or video processing. Despite this growing interest, the theoretical development of quaternion tensors remains limited. This paper introduces a novel multilinear framework for quaternion arrays, which extends the classical tensor analysis to multidimensional quaternion data in a rigorous manner. Specifically, we propose a new definition of quaternion tensors as  $\mathbb{H}\mathbb{R}$ -multilinear forms, addressing the challenges posed by the non-commutativity of quaternion multiplication. Within this framework, we establish the Tucker decomposition for quaternion tensors and develop a quaternion Canonical Polyadic Decomposition (Q-CPD). We thoroughly investigate the properties of the Q-CPD, including trivial ambiguities, complex equivalent models, and sufficient conditions for uniqueness. Additionally, we present two algorithms for computing the Q-CPD and demonstrate their effectiveness through numerical experiments. Our results provide a solid theoretical foundation for further research on quaternion tensor decompositions and offer new computational tools for practitioners working with quaternion multiway data.

**Key words.** quaternions, multilinear algebra, low-rank tensor decomposition, quaternion CPD.

**MSC codes.** 15A69, 15A33.

**1. Introduction.** Quaternions are famous in computer science by their ability to efficiently encode three-dimensional (3D) rotations (which is crucial for e.g., computer graphics [30]). In addition, recent years have seen an increased interest in quaternion arrays, where quaternions are used to encode as a single scalar a 3D or 4D real vector information. Important examples include RGB color images [1, 4, 39, 40], polarimetric images [11, 25], vector array processing [3, 14] or joint wind and temperature forecasting [31], to cite only a few. For a recent review on the usefulness of quaternion representations in data science, signal and image processing, see [24] and the recent special issue [33, 34].

Quaternion multiway arrays arise naturally when 3D or 4D vector data is collected across two or more diversities, such as time, space, wavelength, database index, etc. Over the last decades, an important research effort has been devoted to the study of quaternion vectors and matrices, see e.g., [38] for an introduction to quaternion linear algebra. On the other hand, the study of quaternion higher-order multiway arrays (i.e., with 3 dimensions or more) remains in its infancy. In particular, the notion of *quaternion tensor* – as a multilinear mathematical object – remains to be properly established, despite being commonly employed as a synonym for multidimensional arrays of quaternions. Several tensor-like tools for quaternion multiway arrays have been recently proposed [27, 22, 26] and used in color imaging applications [37, 23]. These works essentially build on generalization of the tensor singular value decomposition (t-SVD) [18, 17] to the quaternion domain. In fact, while in the real and complex cases

---

\*Submitted to the editors on October 10, 2025.

**Funding:** this work was supported by CNRS Informatics through the TenQ project (2022-2023) and in part by the ANR grant RICOCHET ANR-21-CE48-0013. For the purpose of Open Access, a CC-BY public copyright licence has been applied by the authors to the present document and will be applied to all subsequent versions up to the Author Accepted Manuscript arising from this submission. All article content, except where otherwise noted, is licensed under a [Creative Commons Attribution \(CC-BY\) license](#).

<sup>†</sup>Université de Lorraine, CNRS, CRAN, F-54000 Nancy, France

<sup>‡</sup>Université de Toulon, Aix-Marseille Université, CNRS, LIS UMR 7020, France

Notation	Description
$\mathbb{R}, \mathbb{C}, \mathbb{H}$	real, complex fields and quaternion skew-field
$\mathbb{C}_\mu = \mathbb{R} \oplus \mu\mathbb{R}$	complex subfield of $\mathbb{H}$ , with $\mu^2 = -1$
$a, \mathbf{a}, \mathbf{A}, \mathcal{A}$	scalar, vector, matrix and tensor
$\bar{a}, \bar{\mathbf{a}}, \bar{\mathbf{A}}$	conjugation of a scalar / vector / matrix
$\mathbf{A}^\top, \mathbf{A}^\mathbb{H}$	transpose and conjugate-transpose
$q = q_a + iq_b + jq_c + kq_d$	$q \in \mathbb{H}$ , $q_a, q_b, q_c, q_d$ real components
$q =  q  \exp(\mu_q \varphi_q)$	polar form of $q \in \mathbb{H}$
$\mathbf{A} \cdot_\triangleright \mathbf{B}, \mathbf{A} \cdot_\triangleleft \mathbf{B}$	direct and reverse quaternion matrix product
$\chi_\triangleright(\mathbf{A}), \chi_\triangleleft(\mathbf{A})$	direct and reverse complex adjoint
$\chi_\triangleright^\pi(\mathbf{A}), \chi_\triangleleft^\pi(\mathbf{A})$	direct and reverse column-wise complex adjoint
$\mathbf{A} \boxtimes \mathbf{B}, \mathbf{A} \boxtimes_\triangleright \mathbf{B}, \mathbf{A} \boxtimes_\triangleleft \mathbf{B}$	Kronecker product, direct and reverse Kronecker product
$\mathbf{A} \odot \mathbf{B}, \mathbf{A} \odot_\triangleright \mathbf{B}, \mathbf{A} \odot_\triangleleft \mathbf{B}$	Khatri-Rao product, direct and reverse Khatri-Rao product
$\mathbf{A} \boxtimes \mathbf{B}$	Hadamard product
$\mathbf{T}_{i::}, \mathbf{T}_{:j:}, \mathbf{T}_{::k}$	$i$ -th horizontal slice, $j$ -th lateral slice, $k$ -th frontal slice of tensor $\mathcal{T}$
$\mathbf{T}_{(n)}$	$n$ -mode unfolding of tensor $\mathcal{T}$
$\mathcal{T} \times_n^\triangleright \mathbf{A}, \mathcal{T} \times_n \mathbf{A}, \mathcal{T} \times_n^\triangleleft \mathbf{A}$	$n$ -mode products with possible direct or reverse quaternion matrix multiplication

TABLE 0.1

Notations used in this article. The matrices  $\mathbf{A}$  and  $\mathbf{B}$  have appropriate dimensions. When at least one is real-valued, the  $\triangleright$  and  $\triangleleft$  indices are irrelevant and they are dropped in notations.

it is known [13] that such decompositions can be interpreted as a specific multilinear tensor decomposition (such as the Canonical Polyadic decomposition), the lack of a multilinear definition of quaternion tensors makes it an interesting open question.

This paper deals with the study of quaternion multidimensional arrays from a multilinear point of view to meaningfully define the notion of *quaternion tensors*. The main difficulty lies in handling the non-commutativity of quaternion multiplication, which complicates the analysis and prevents straightforward extensions of the usual definitions from the real and complex cases. It builds on our preliminary results in [15], where only partial results were presented. The contributions of this paper are threefold: (i) we propose a coordinate-free definition of quaternion tensors as  $\mathbb{H}\mathbb{R}$ -multilinear forms and identify a quaternion multiway array with the representation of a quaternion tensor in some fixed product basis; (ii) we introduce the notion of quaternion Tucker format,  $n$ -mode products, change-of-basis and characterize their properties; (iii) we develop a complete theory for quaternion Canonical Polyadic Decomposition (Q-CPD), including characterization of trivial ambiguities, equivalent complex-valued models and establish several sufficient uniqueness conditions. Finally, two practical algorithms for computing the Q-CPD are presented and their performance is validated through numerical experiments. We believe the results of this paper can be of interest to practitioners dealing with quaternion multiway arrays, by providing a general framework for the development of new quaternion tensor decompositions and relevant models for applications that require uniqueness guarantees. In addition, the theoretical study presented in this paper relies extensively on the necessary distinction between *left* and *right* linear independence of quaternion vectors – due to quaternion

non-commutativity – leading us to introduce several new results on left and right ranks and Kruskal ranks of quaternion matrices, which may be of independent interest for researchers in quaternion linear algebra.

This paper is organized as follows. [Section 2](#) recalls basics about the quaternion algebra and introduces our notations. We then extend some key concepts of real and complex multilinear algebra to the quaternion case. In particular we establish several new results on the Kruskal rank of quaternion matrices that are crucial for our theoretical analysis of quaternion tensor decompositions. Our main contributions are exposed in subsequent sections. [Section 3](#) establishes a formal definition of a quaternion tensor as a multilinear form with a specific structure. This enables a natural derivation of the Tucker model for quaternion tensors. [Section 4](#) focuses on the Q-CPD and the study of its properties, equivalent complex representations and provides several sufficient uniqueness conditions. [Section 5](#) develops two different algorithms for computing the Q-CPD, which are further evaluated through numerical experiments in [Section 6](#). [Section 7](#) concludes the paper. [Appendix A](#) gathers technical details and proofs.

**2. Preliminaries.** This section starts by reviewing elementary properties of the set of quaternions  $\mathbb{H}$ . We proceed by introducing the basic notions of quaternion linear algebra, including quaternion matrix operations, complex adjoint(s), and quaternion vector spaces. The theory in this paper requires us to distinguish between *direct* and *reverse* quaternion matrix products; together with the notion of *left* and *right* linear independence of quaternion vectors, these notions lead to several new results that may be of independent interest. These are presented in [Subsection 2.2.4](#). In particular, [Proposition 2.7](#) links the right and left ranks of a quaternion matrix to those of its direct and reverse complex adjoints; [Definition 2.8](#) defines the notion of left and right Kruskal rank ( $k$ -rank) for quaternion matrices, and [Lemma 2.10](#) establishes a crucial link between these  $k$ -ranks of a quaternion matrix and the so-called  $k'$ -rank of its direct and reverse complex adjoint. For reference, [Table 0.1](#) summarizes the notations used in this article.

**2.1. Quaternions.** Let  $\mathbb{R}$  and  $\mathbb{C}$  denote the usual real and complex field, respectively. The set of quaternions  $\mathbb{H}$  defines a 4-dimensional normed division algebra over the real numbers  $\mathbb{R}$ . It has canonical basis  $\{1, \mathbf{i}, \mathbf{j}, \mathbf{k}\}$ , where  $\mathbf{i}, \mathbf{j}, \mathbf{k}$  are imaginary units such that

$$(2.1) \quad \mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -1, \quad \mathbf{ij} = -\mathbf{ji}, \mathbf{ij} = \mathbf{k}.$$

Any quaternion  $q \in \mathbb{H}$  can be written as

$$(2.2) \quad q = q_a + \mathbf{i}q_b + \mathbf{j}q_c + \mathbf{k}q_d,$$

where  $q_a, q_b, q_c, q_d \in \mathbb{R}$  are called the components of  $q$ . The real part of  $q$  is  $\text{Re } q = q_a$  whereas its imaginary part is  $\text{Im } q = \mathbf{i}q_b + \mathbf{j}q_c + \mathbf{k}q_d$ . A quaternion  $q$  is said to be purely imaginary (or simply, pure) if  $\text{Re } q = 0$ . The quaternion conjugate of  $q$  is denoted by  $\bar{q} = \text{Re } q - \text{Im } q$ . The modulus of  $q$  is  $|q| = \sqrt{q\bar{q}} = \sqrt{\bar{q}q} = \sqrt{q_a^2 + q_b^2 + q_c^2 + q_d^2}$ . Any non-zero quaternion  $q$  has an inverse  $q^{-1} = \bar{q}/|q|^2$ . Just like complex numbers, a quaternion can be written in polar form as  $q = |q| \exp(\boldsymbol{\mu}_q \varphi_q)$ , where its axis  $\boldsymbol{\mu}_q$  is a pure unit quaternion (i.e., such that  $\boldsymbol{\mu}_q^2 = -1$ ) and where  $\varphi_q \in \mathbb{R}$  is the angle of  $q$ .

One of the specific features of quaternions is the non-commutativity of the quaternion product. Indeed, for  $p, q \in \mathbb{H}$ , one has  $pq \neq qp$  in general since imaginary units do not commute. Nonetheless, standard operations such as conjugation and

inversion distribute well over products of two quaternions: one has  $\overline{(pq)} = \bar{q}\bar{p}$  and  $(pq)^{-1} = q^{-1}p^{-1}$ .

Given a pure unit quaternion  $\boldsymbol{\mu}$  (i.e., such that  $\text{Re } \boldsymbol{\mu} = 0$  and  $|\boldsymbol{\mu}| = 1$ ), one can define the set  $\mathbb{C}_{\boldsymbol{\mu}} = \mathbb{R} \oplus \boldsymbol{\mu}\mathbb{R}$  which is a *complex subfield* of  $\mathbb{H}$ . In particular, any quaternion  $q$  can be written as a combination of two complex numbers in  $\mathbb{C}_{\boldsymbol{\mu}}$ , a procedure known as the *Cayley-Dickson decomposition* of  $q$ . For  $\boldsymbol{\mu} = i$ , it reads

$$(2.3) \quad q = q_1 + q_2j, \quad q_1, q_2 \in \mathbb{C}_i.$$

The decomposition (2.3) permits to establish a useful complex matrix representation of quaternion arrays [38], known as the complex matrix adjoint. This will be investigated in detail in [Subsection 2.2.2](#).

**2.2. Quaternion linear algebra.** This section reviews important notions of linear algebra in the quaternion domain. For a comprehensive treatment of the subject, see e.g., [38] or [28]. In addition, we prove some new results that will be useful to study the uniqueness of quaternion tensor decompositions in Section 4.

**2.2.1. Quaternion matrix operations.** Quaternion vectors and matrices can be defined as 1-D and 2-D arrays with quaternion entries. The  $i$ -th entry of quaternion vector  $\mathbf{q} \in \mathbb{H}^N$  is given by  $(\mathbf{q})_n = q_n \in \mathbb{H}$ , while for a quaternion matrix  $\mathbf{A} \in \mathbb{H}^{M \times N}$ , its  $(m, n)$ -th entry reads  $(\mathbf{A})_{mn} = a_{mn} \in \mathbb{H}$ . The columns of  $\mathbf{A} \in \mathbb{H}^{M \times N}$  are denoted by  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_N \in \mathbb{H}^M$ . The transpose of  $\mathbf{A} \in \mathbb{H}^{M \times N}$  is denoted by  $\mathbf{A}^\top \in \mathbb{H}^{N \times M}$  and its conjugate-transpose is  $\mathbf{A}^H := \overline{(\mathbf{A}^\top)} = \overline{(\mathbf{A})}^\top \in \mathbb{H}^{N \times M}$ . Finally, we denote by  $\mathbf{I}_M$  the identity matrix in  $\mathbb{H}^{M \times M}$ , which is the same as in  $\mathbb{C}^{M \times M}$  and  $\mathbb{R}^{M \times M}$ .

Standard operations of linear algebra translate to the quaternion domain provided that operations take into account the non-commutativity of the quaternion product. Consider  $\mathbf{A} \in \mathbb{H}^{M \times N}$  and  $\mathbf{B} \in \mathbb{H}^{N \times P}$ . Then it is possible to define the direct and reverse quaternion matrix products as

$$(2.4) \quad \text{direct quaternion matrix product} \quad (\mathbf{A} \cdot_{\triangleright} \mathbf{B})_{mp} := \sum_{n=1}^N a_{mn} b_{np},$$

$$(2.5) \quad \text{reverse quaternion matrix product} \quad (\mathbf{A} \cdot_{\triangleleft} \mathbf{B})_{mp} := \sum_{n=1}^N b_{np} a_{mn}.$$

The direct or reverse denomination simply refers to the ordering of the two factors in the expression of the matrix product entries: direct refers to the left-to-right direction, while reverse means right-to-left direction. Note that, in many publications, the direct matrix product is often implicitly used [38]; however this distinction will appear crucial to our construction of quaternion tensors (see [Section 3](#)).

Due to quaternion non-commutativity, for two matrices with compatible sizes,  $\mathbf{A} \cdot_{\triangleright} \mathbf{B} \neq \mathbf{A} \cdot_{\triangleleft} \mathbf{B}$  in general; however, if one of the matrices has real-valued entries (say e.g.,  $\mathbf{B} \in \mathbb{R}^{N \times P}$ ), then  $\mathbf{A} \cdot_{\triangleright} \mathbf{B} = \mathbf{A} \cdot_{\triangleleft} \mathbf{B}$  and we can drop the subscript to write  $\mathbf{AB}$  if there is no risk of confusion from the context.

Following [29]<sup>1</sup>, the direct and reverse quaternion matrix products satisfy the

<sup>1</sup>In [29], the authors refer to direct and reverse quaternion matrix products as “left” and “right”, respectively. To avoid confusion with left and right quaternion vector spaces introduced in the next section, we chose this alternate denomination of quaternion matrix products.

following properties:

$$\begin{array}{ll}
 \text{transposition} & (\mathbf{A} \cdot_{\triangleright} \mathbf{B})^{\top} = \mathbf{B}^{\top} \cdot_{\triangleleft} \mathbf{A}^{\top} \quad \text{and} \quad (\mathbf{A} \cdot_{\triangleleft} \mathbf{B})^{\top} = \mathbf{B}^{\top} \cdot_{\triangleright} \mathbf{A}^{\top}, \\
 \text{conjugation} & \overline{(\mathbf{A} \cdot_{\triangleright} \mathbf{B})} = \overline{\mathbf{A}} \cdot_{\triangleleft} \overline{\mathbf{B}} \quad \text{and} \quad \overline{(\mathbf{A} \cdot_{\triangleleft} \mathbf{B})} = \overline{\mathbf{A}} \cdot_{\triangleright} \overline{\mathbf{B}}, \\
 \text{conjugate-transposition} & (\mathbf{A} \cdot_{\triangleright} \mathbf{B})^{\text{H}} = \mathbf{B}^{\text{H}} \cdot_{\triangleright} \mathbf{A}^{\text{H}} \quad \text{and} \quad (\mathbf{A} \cdot_{\triangleleft} \mathbf{B})^{\text{H}} = \mathbf{B}^{\text{H}} \cdot_{\triangleleft} \mathbf{A}^{\text{H}}.
 \end{array}$$

Proofs of these properties are readily obtained from the direct and reverse quaternion matrix product definitions.

The same need to distinguish between direct and reverse product applies to Kronecker and Khatri-Rao products, which are two standard matrix products encountered in tensor algebra. Given two quaternion matrices  $\mathbf{A} \in \mathbb{H}^{M \times N}$  and  $\mathbf{B} \in \mathbb{H}^{P \times Q}$ , their direct and reverse Kronecker product are defined as:

$$\begin{aligned}
 \mathbf{A} \boxtimes_{\triangleright} \mathbf{B} &:= \begin{bmatrix} a_{11} \mathbf{B} & \dots & a_{1N} \mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{M1} \mathbf{B} & \dots & a_{MN} \mathbf{B} \end{bmatrix} \in \mathbb{H}^{(MP) \times (NQ)}, \\
 \mathbf{A} \boxtimes_{\triangleleft} \mathbf{B} &:= \begin{bmatrix} \mathbf{B} a_{11} & \dots & \mathbf{B} a_{1N} \\ \vdots & \ddots & \vdots \\ \mathbf{B} a_{M1} & \dots & \mathbf{B} a_{MN} \end{bmatrix} \in \mathbb{H}^{(MP) \times (NQ)}.
 \end{aligned}$$

Similarly, given two matrices  $\mathbf{A} \in \mathbb{H}^{M \times P}$  and  $\mathbf{B} \in \mathbb{H}^{N \times P}$ , their direct and reverse Khatri-Rao product (column-wise Kronecker product) are defined as

$$\begin{aligned}
 \mathbf{A} \odot_{\triangleright} \mathbf{B} &:= [\mathbf{a}_1 \boxtimes_{\triangleright} \mathbf{b}_1 \quad \mathbf{a}_2 \boxtimes_{\triangleright} \mathbf{b}_2 \quad \dots \quad \mathbf{a}_P \boxtimes_{\triangleright} \mathbf{b}_P] \in \mathbb{H}^{(MN) \times P}, \\
 \mathbf{A} \odot_{\triangleleft} \mathbf{B} &:= [\mathbf{a}_1 \boxtimes_{\triangleleft} \mathbf{b}_1 \quad \mathbf{a}_2 \boxtimes_{\triangleleft} \mathbf{b}_2 \quad \dots \quad \mathbf{a}_P \boxtimes_{\triangleleft} \mathbf{b}_P] \in \mathbb{H}^{(MN) \times P}.
 \end{aligned}$$

Finally, the Hadamard product between two quaternion matrices  $\mathbf{A}, \mathbf{B} \in \mathbb{H}^{M \times N}$  is defined as

$$(2.6) \quad (\mathbf{A} \boxdot \mathbf{B})_{mn} = a_{mn} b_{mn} .$$

which is naturally non-commutative, unlike its real and complex counterparts. Therefore, we do not distinguish between direct and reverse product since it simply corresponds to exchanging the positions of the two matrices  $\mathbf{A}$  and  $\mathbf{B}$ .

Finally, note that similarly to the quaternion matrix product case, the distinction between direct and reverse operations becomes irrelevant for Kronecker and Khatri-Rao products as long as one of the matrices is real-valued. In this case, we usually drop the subscript if there is no risk of confusion within the context. More properties on direct and reverse matrix operations can be found in [29].

**2.2.2. Complex adjoints for quaternion matrix products.** The notion of complex adjoint of a quaternion matrix is a key tool in quaternion linear algebra [38]. In particular, it permits to establish the properties of quaternion matrices (rank, inversion, eigen-decompositions, etc.) from their complex adjoint counterpart. The definition of the complex adjoint is implicitly related to the choice of quaternion matrix multiplication, i.e., direct (2.4) or reverse (2.5). Therefore in full generality one needs to distinguish between *direct* and *reverse* adjoints. Let  $\mathbf{A} \in \mathbb{H}^{M \times N}$  and consider its Cayley-Dickson decomposition  $\mathbf{A} = \mathbf{A}_1 + \mathbf{A}_2 \mathbf{j}$ , where  $\mathbf{A}_1, \mathbf{A}_2$  are matrices in  $\mathbb{C}_i^{2M \times 2N}$ . The *direct* complex adjoint  $\chi_{\triangleright}(\mathbf{A})$  and *reverse* complex adjoint  $\chi_{\triangleleft}(\mathbf{A})$

enable an isomorphism between quaternion matrices (equipped with the corresponding matrix product) and complex matrices with block structure, such that

$$(2.7) \quad \chi_{\triangleright}(\mathbf{A}) = \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 \\ -\overline{\mathbf{A}_2} & \overline{\mathbf{A}_1} \end{bmatrix} \in \mathbb{C}_i^{2M \times 2N},$$

$$(2.8) \quad \chi_{\triangleleft}(\mathbf{A}) = \begin{bmatrix} \mathbf{A}_1 & -\overline{\mathbf{A}_2} \\ \mathbf{A}_2 & \overline{\mathbf{A}_1} \end{bmatrix} \in \mathbb{C}_i^{2M \times 2N}.$$

The only difference – yet crucial – between (2.7) and (2.8) resides in the placement of anti-diagonal blocks. [Proposition 2.1](#) below summarizes important properties of these complex adjoints.

**PROPOSITION 2.1** (Properties of the complex adjoint). *Let  $\mathbf{A}$  and  $\mathbf{B}$  be two quaternion matrices with arbitrary dimensions. The direct (2.7) and reverse (2.8) complex adjoint satisfy the following properties:*

- $\chi_{\triangleright}(\mathbf{I}_M) = \chi_{\triangleleft}(\mathbf{I}_M) = \mathbf{I}_{2M}$ ;
- $\chi_{\triangleright}(\mathbf{A} + \mathbf{B}) = \chi_{\triangleright}(\mathbf{A}) + \chi_{\triangleright}(\mathbf{B})$  and  $\chi_{\triangleleft}(\mathbf{A} + \mathbf{B}) = \chi_{\triangleleft}(\mathbf{A}) + \chi_{\triangleleft}(\mathbf{B})$  if  $\mathbf{A}, \mathbf{B}$  have the same size;
- $\chi_{\triangleright}(\mathbf{A} \cdot_{\triangleright} \mathbf{B}) = \chi_{\triangleright}(\mathbf{A})\chi_{\triangleright}(\mathbf{B})$  and  $\chi_{\triangleleft}(\mathbf{A} \cdot_{\triangleleft} \mathbf{B}) = \chi_{\triangleleft}(\mathbf{A})\chi_{\triangleleft}(\mathbf{B})$  if one can form the matrix product between  $\mathbf{A}$  and  $\mathbf{B}$ ;
- $\chi_{\triangleright}(\mathbf{A}^H) = \chi_{\triangleright}^H(\mathbf{A})$  and  $\chi_{\triangleleft}(\mathbf{A}^H) = \chi_{\triangleleft}^H(\mathbf{A})$ ;
- $\chi_{\triangleright}(\mathbf{A}) = \chi_{\triangleleft}^T(\mathbf{A}^T)$ ;

The proof of these results follows by direct calculations. For additional properties, we refer the reader to [38], where the direct complex adjoint is considered - similar results hold for the reverse adjoint by simple adaption of the proofs therein.

*Remark 2.2.* The standard definitions (2.7)–(2.8) of complex adjoints of a quaternion matrix can be related to the horizontal stack of the adjoint of each of its columns. Let  $\mathbf{A} = [\mathbf{a}_1 \mathbf{a}_2 \dots \mathbf{a}_N] \in \mathbb{H}^{M \times N}$ , and introduce the direct column-wise complex adjoint

$$(2.9) \quad \chi_{\triangleright}^{\pi}(\mathbf{A}) = [\chi_{\triangleright}(\mathbf{a}_1) \mid \chi_{\triangleright}(\mathbf{a}_2) \mid \dots \mid \chi_{\triangleright}(\mathbf{a}_N)] \in \mathbb{C}_i^{2M \times 2N}.$$

From the definition of the direct complex adjoint for a quaternion vector, it is easy to see that there exists a permutation matrix  $\mathbf{\Pi} \in \mathbb{R}^{2N \times 2N}$  such that  $\chi_{\triangleright}(\mathbf{A}) = \chi_{\triangleright}^{\pi}(\mathbf{A})\mathbf{\Pi}$ . Equation (2.9) also defines a uniform partition of  $\chi_{\triangleright}^{\pi}(\mathbf{A})$  in  $N$  submatrices  $\chi_{\triangleright}(\mathbf{a}_n) \in \mathbb{C}_i^{2M \times 2}$ . The same definition and properties holds for the columnwise reverse complex adjoint, which we denote by  $\chi_{\triangleleft}^{\pi}(\mathbf{A})$ . This will be useful later on for the study of tensor decompositions of quaternion multiway arrays.

**2.2.3. Quaternion vector spaces.** The set of quaternions  $\mathbb{H}$  is non-commutative and therefore it is not a field. As a result, since vector spaces are defined over a field (e.g.,  $\mathbb{R}$  or  $\mathbb{C}$ ), it is not possible to define the notion of vector space over  $\mathbb{H}$ , *per se*. Strictly speaking, one has to consider the notion of left- or right-module, which can be viewed as an extension of vector spaces where scalar multiplication is performed over a non-commutative ring. However, in the case of quaternions, it is common practice [28] to call left (resp. right)  $\mathbb{H}$ -modules left (resp. right) quaternion vector spaces, with little abuse. In particular, since  $\mathbb{H}$  is a division algebra, quaternion vector spaces exhibit the usual properties of vector spaces (dimension, basis, linear maps, etc.) up to the necessary distinction between left and right linearity.

For completeness, we give the definition of a left quaternion vector space below. The definition for a right vector space follows by direct adaption.

DEFINITION 2.3 (Left quaternion vector space [16]). A left vector space  $\mathcal{H}$  over  $\mathbb{H}$  is an additive abelian group in which the operation of scalar multiplication by elements of  $\mathbb{H}$  is defined. Scalar multiplication is assumed to obey the following laws for all  $\mathbf{x}, \mathbf{y} \in \mathcal{H}$  and  $q, p \in \mathbb{H}$ :

- $q(\mathbf{x} + \mathbf{y}) = q\mathbf{x} + q\mathbf{y}$ ;
- $(q + p)\mathbf{x} = q\mathbf{x} + p\mathbf{x}$ ;
- $(qp)\mathbf{x} = q(p\mathbf{x})$ ;
- $1 \cdot \mathbf{x} = \mathbf{x}$ , where  $1 \in \mathbb{H}$ .

A subscript  $_L$  or  $_R$  permits to distinguish between left and right quaternion vector spaces. Two important examples are:  $\mathbb{H}_L^M$ , the left quaternion vector space of vectors in  $\mathbb{H}^M$  and  $\mathbb{H}_R^M$ , the right quaternion vector space of vectors in  $\mathbb{H}^M$ .

**2.2.4. The rank and Kruskal rank of quaternion matrices.** In this section we review the *rank* of a quaternion matrix and extend the definition of the Kruskal rank, or simply k-rank, to quaternion matrices. In the real and complex cases, the rank of a matrix refers to the maximum number of *linearly independent* columns (or rows) in the matrix. When extending this notion to quaternion matrices, non-commutativity of quaternion multiplication calls once again for a distinction between left and right linear independence. The rank of a quaternion matrix  $\mathbf{A}$  is usually defined to be the maximum number of *right* linearly independent columns [38]. However, as it will appear clearly in the sequel, the theory presented in this paper calls for a distinction between the *left* and *right* ranks (and thus, k-ranks) of a quaternion matrix.

DEFINITION 2.4 (Left and right columns spans). Let  $\mathbf{A} \in \mathbb{H}^{M \times N}$  and denote by  $\mathbf{a}_1, \dots, \mathbf{a}_N \in \mathbb{H}^M$  its columns. Then, the span of left-linear combination of columns of  $\mathbf{A}$  is denoted as

$$(2.10) \quad \text{span}_L(\mathbf{A}) := \left\{ \sum_{n=0}^N x_n \mathbf{a}_n \mid x_1, \dots, x_n \in \mathbb{H} \right\},$$

and the span of right-linear combinations of columns of  $\mathbf{A}$  is given by

$$(2.11) \quad \text{span}_R(\mathbf{A}) := \left\{ \sum_{n=0}^N \mathbf{a}_n x_n \mid x_1, \dots, x_n \in \mathbb{H} \right\}.$$

From these definitions, one observes that  $\text{span}_L(\mathbf{A})$  defines a left vector space over the columns of  $\mathbf{A}$ , while  $\text{span}_R(\mathbf{A})$  defines a right vector space over the same columns.

The following definition of left and right ranks of a quaternion matrix follows naturally.

DEFINITION 2.5 (Left and right ranks of a quaternion matrix). Let  $\mathbf{A} \in \mathbb{H}^{M \times N}$ . The left rank of  $\mathbf{A}$  is defined as  $\text{rank}_L \mathbf{A} = \dim \text{span}_L(\mathbf{A})$  while the right rank is given by  $\text{rank}_R \mathbf{A} = \dim \text{span}_R(\mathbf{A})$ .

*Remark 2.6.* For a real or complex matrix  $\mathbf{A}$ , we drop the subscripts and keep the unambiguous notations  $\text{span}(\mathbf{A})$  to denote the span of its columns and its rank by  $\text{rank } \mathbf{A}$ .

As in the usual real and complex cases, the dimension of the left (resp. right) column span encodes the maximal number of left (resp. right) linearly independent vectors. For quaternion matrices, the left and right ranks are in general different; see e.g., [38, Example 7.3]. Hence the distinction is critical. Nonetheless, the left and right ranks satisfy several important properties similar to their real and complex counterparts. In

particular, for a quaternion matrix  $\mathbf{A} \in \mathbb{H}^{M \times N}$ , the rank cannot exceed its dimensions, i.e.,  $\text{rank}_L \mathbf{A} \leq \min(M, N)$  and  $\text{rank}_R \mathbf{A} \leq \min(M, N)$ . The matrix  $\mathbf{A}$  is said to be full left rank if  $\text{rank}_L \mathbf{A} = \min(M, N)$  and full right rank if  $\text{rank}_R \mathbf{A} = \min(M, N)$ .

The following proposition relates the left and right ranks of a quaternion matrix to those of its direct and reverse complex adjoints.

**PROPOSITION 2.7** (Left and right ranks from complex adjoints). *Let  $\mathbf{A} \in \mathbb{H}^{M \times N}$  be a quaternion matrix. Then the left and right ranks of  $\mathbf{A}$  can be obtained from direct and reverse complex adjoints as*

1.  $\text{rank } \chi_{\triangleright}(\mathbf{A}) = 2 \text{rank}_R \mathbf{A}$ ;
2.  $\text{rank } \chi_{\triangleleft}(\mathbf{A}) = 2 \text{rank}_L \mathbf{A}$ .

*Proof.* We only prove the first statement; the second follows from direct adaptation of the proof. Let  $\mathbf{A} \in \mathbb{H}^{M \times N}$  and suppose that  $\text{rank}_R \mathbf{A} = r \leq \min(M, N)$ .

Let us first prove that  $\text{rank } \chi_{\triangleright}(\mathbf{A}) \geq 2r$ . Since  $\text{rank}_R \mathbf{A} = r$ , this implies that there exist  $r$  columns  $\mathbf{a}_1, \dots, \mathbf{a}_r$  of  $\mathbf{A}$  which are right linearly independent. Denote by  $\mathbf{A}_r = [\mathbf{a}_1 \cdots \mathbf{a}_r] \in \mathbb{H}^{M \times r}$  this submatrix of  $\mathbf{A}$ . Now let  $\boldsymbol{\beta} \in \mathbb{C}_i^{2r}$  such that  $\chi_{\triangleright}(\mathbf{A}_r)\boldsymbol{\beta} = 0$ . Partition  $\boldsymbol{\beta} = [\boldsymbol{\beta}_1^\top \quad -\overline{\boldsymbol{\beta}_2}^\top]^\top$  with  $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2 \in \mathbb{C}_i^r$ . The complex adjoint structure of  $\chi_{\triangleright}(\mathbf{A}_r)$  implies that  $\chi_{\triangleright}(\mathbf{A}_r)\boldsymbol{\beta}^\perp = 0$  as well, where  $\boldsymbol{\beta}^\perp := [\boldsymbol{\beta}_2^\top \quad \overline{\boldsymbol{\beta}_1}^\top]^\top$ . Therefore, posing  $\boldsymbol{\alpha} = \boldsymbol{\beta}_1 + \boldsymbol{\beta}_2 \mathbf{j}$  and observing that  $[\boldsymbol{\beta} \boldsymbol{\beta}^\perp] = \chi_{\triangleright}(\boldsymbol{\alpha})$ , one has  $\chi_{\triangleright}(\mathbf{A}_r)\chi_{\triangleright}(\boldsymbol{\alpha}) = 0$ . Using the properties of the direct adjoint we thus have  $\chi_{\triangleright}(\mathbf{A}_r \cdot_{\triangleright} \boldsymbol{\alpha}) = 0$  and then  $\mathbf{A}_r \cdot_{\triangleright} \boldsymbol{\alpha} = 0$ . Since we have supposed that the columns of  $\mathbf{A}_r$  were linearly independent, we necessarily have  $\boldsymbol{\alpha} = 0$  hence  $\boldsymbol{\beta} = 0$ . Therefore the  $2r$  columns of  $\chi_{\triangleright}(\mathbf{A}_r)$  are linearly independent over  $\mathbb{C}_i$ . Since  $\chi_{\triangleright}(\mathbf{A}_r)$  is a submatrix of  $\chi_{\triangleright}(\mathbf{A})$  we have  $\text{rank } \chi_{\triangleright}(\mathbf{A}) \geq 2r$ . We now prove that  $\text{rank } \chi_{\triangleright}(\mathbf{A}) \leq 2r$ . Since  $\text{rank}_R \mathbf{A} = r$ , upon proper permutations of columns, one can factorize  $\mathbf{A} = \mathbf{A}_r \cdot_{\triangleright} [\mathbf{I}_r \mathbf{B}]$ , where  $\mathbf{A}_r = [\mathbf{a}_1 \cdots \mathbf{a}_r] \in \mathbb{H}^{M \times r}$  has linearly independent columns and where  $\mathbf{B} \in \mathbb{H}^{r \times (N-r)}$  encodes the right-linear combinations that yields the remaining  $N - r$  columns of  $\mathbf{A}$ . Applying the direct complex adjoint to this factorization, one gets  $\chi_{\triangleright}(\mathbf{A}) = \chi_{\triangleright}(\mathbf{A}_r)\chi_{\triangleright}([\mathbf{I}_r \mathbf{B}])$ . This shows that  $\chi_{\triangleright}(\mathbf{A})$  can always be factorized as a product of a  $2M \times 2r$  matrix by a  $2r \times 2N$  matrix, hence  $\text{rank } \chi_{\triangleright}(\mathbf{A}) \leq 2r$ . Bringing the two parts together yields  $\text{rank } \chi_{\triangleright}(\mathbf{A}) = 2r = 2 \text{rank}_R \mathbf{A}$ .  $\square$

The definition of the left and right Kruskal ranks follows naturally from that of the left and right ranks of a quaternion matrix.

**DEFINITION 2.8** (Left and right Kruskal ranks of a quaternion matrix). *Let  $\mathbf{A} \in \mathbb{H}^{M \times N}$  be a quaternion matrix. Then:*

- *the left Kruskal rank of  $\mathbf{A}$  is denoted by  $k_L(\mathbf{A})$ . It is the maximal number  $r$  such that any set of  $r$  columns of  $\mathbf{A}$  are left linearly independent.*
- *the right Kruskal rank of  $\mathbf{A}$  is denoted by  $k_R(\mathbf{A})$ . It is the maximal number  $r$  such that any set of  $r$  columns of  $\mathbf{A}$  are right linearly independent.*

For a real or complex matrix  $\mathbf{A}$ , we simply write  $k(\mathbf{A})$  to denote its usual Kruskal rank. For quaternion matrices, we use the shorthand  $k_L$ -rank and  $k_R$ -rank to mean the left and right Kruskal ranks, respectively.

It is possible to relate the  $k_L$ -rank and  $k_R$ -rank of a quaternion matrix to the  $k'$ -rank of its direct and reverse column-wise complex adjoints (see definitions in [Remark 2.2](#)). Let us first recall the definition of the  $k'$ -rank of a partitioned matrix.

**DEFINITION 2.9** ( $k'$ -rank of a partitioned matrix [8]). *Let  $\mathbf{X} \in \mathbb{F}^{M \times N}$  be a real or complex partitioned matrix (not necessarily uniformly). Then the  $k'$ -rank of  $\mathbf{X}$ , denoted by  $k'(\mathbf{X})$ , is the maximal number  $r$  such that any set of  $r$  submatrices of  $\mathbf{X}$*

yields a set of linearly independent columns.

The following lemma establishes the relation between the  $k_L$ -rank and  $k_R$ -rank of a quaternion matrix and the  $k$ '-rank of its column-wise complex adjoints.

**LEMMA 2.10** (*k*'-rank of column-wise adjoint matrices). *Let  $\mathbf{A} \in \mathbb{H}^{M \times N}$  and let  $\chi_{\triangleright}^{\pi}(\mathbf{A}) \in \mathbb{C}_i^{2M \times 2N}$  and  $\chi_{\triangleleft}^{\pi}(\mathbf{A}) \in \mathbb{C}_i^{2M \times 2N}$  be its column-wise direct complex adjoint and reverse complex adjoint, respectively, with the uniform partition introduced in [Remark 2.2](#). Then  $k_R(\mathbf{A}) = k'(\chi_{\triangleright}^{\pi}(\mathbf{A}))$  and  $k_L(\mathbf{A}) = k'(\chi_{\triangleleft}^{\pi}(\mathbf{A}))$ .*

*Proof.* We start by proving the result for the  $k_R$ -rank of  $\mathbf{A}$ . Let us assume that  $k_R(\mathbf{A}) = c$ . According to the definition of the  $k_R$ -rank, any subset of  $c$  columns of  $\mathbf{A}$  forms a submatrix  $\mathbf{A}_c$  of  $\mathbf{A}$  such that  $\text{rank}_R \mathbf{A}_c = c$ . Using [Proposition 2.7](#), we obtain  $\text{rank} \chi_{\triangleright}(\mathbf{A}_c) = 2c$ . This is equivalent to saying that any set of direct adjoint matrices associated with the  $c$  columns of  $\mathbf{A}_c$  yields a set of linearly independent columns of  $\chi_{\triangleright}^{\pi}(\mathbf{A})$ , which, by [Definition 2.9](#) implies  $k'(\chi_{\triangleright}^{\pi}(\mathbf{A})) \geq c = k_R(\mathbf{A})$ .

Using once again the definition of the  $k_R$ -rank, there exists at least a subset of  $c+1$  columns of  $\mathbf{A}$ , such that the matrix  $\mathbf{A}_{c+1}$  formed by these  $c+1$  columns is right rank deficient, i.e.,  $\text{rank}_R \mathbf{A}_{c+1} < c+1$ , implying  $\text{rank} \chi_{\triangleright}(\mathbf{A}_{c+1}) < 2(c+1)$ . Consequently, as  $\chi_{\triangleright}(\mathbf{A}) = \chi_{\triangleright}^{\pi}(\mathbf{A})\mathbf{I}$ , there exists a set of  $c+1$  submatrices of  $\chi_{\triangleright}^{\pi}(\mathbf{A})$  that yields a set of linearly dependent columns, thereby completing the proof. The second result can be proved in an analogous way, by replacing  $\chi_{\triangleright}(\mathbf{A})$  by  $\chi_{\triangleleft}(\mathbf{A})$  and right ranks by left ranks.  $\square$

This result will be essential to the study of uniqueness conditions for the quaternion Canonical Polyadic Decomposition, see further in [Section 4](#).

**3. Quaternion tensors: construction and properties.** In the standard real and complex cases, three alternative definitions of *tensors* are possible [\[21\]](#): tensors as objects obeying certain transformation rules, tensors as multilinear maps or tensors as elements of tensor product spaces. While the first definition has historical importance, it has now been superseded by the two other definitions [\[5\]](#). Over the past few years, there has been an increasing interest in studying multidimensional arrays of quaternions [\[27, 22, 26, 37, 23\]](#); while the term *quaternion tensor* is often used, a proper algebraic, coordinate-free definition remains to be established.

In this paper, we choose to construct quaternion tensors as quaternion-valued multilinear forms. Our definition of a quaternion tensor can be viewed as an instance of [\[21, Definition 3.3\]](#) specified to quaternion-valued forms satisfying sufficient multilinearity properties. More precisely, we identify a tensor of order  $D$  with a quaternion-valued multilinear form  $f : \mathcal{S}_1 \times \mathcal{S}_2 \times \dots \times \mathcal{S}_D \rightarrow \mathbb{H}$  where  $\{\mathcal{S}_d\}_{d=1}^D$  are some carefully chosen vector spaces. This definition enables a natural generalization of the notion of tensors to the quaternion domain, while revealing the crucial caveats associated to the non-commutativity of quaternion multiplication. In what follows, [Subsection 3.1](#) discusses the issue of quaternion multilinearity, [Subsection 3.2](#) establishes the definition of quaternion tensors as multilinear forms. [Subsection 3.3](#) derives elementary transformations for quaternion-valued tensors, which are direct consequences of [Definition 3.3](#) below. Finally, [Subsection 3.4](#) introduces the Tucker format of a quaternion tensor.

**3.1. The quaternion multilinearity issue.** The definition of quaternion multilinearity is cumbersome due to the non-commutativity of quaternion multiplication. To illustrate this, considering the three-dimensional case is sufficient. Let  $\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3$  be three left quaternion vector spaces and let  $f : \mathcal{S}_1 \times \mathcal{S}_2 \times \mathcal{S}_3 \rightarrow \mathbb{H}$  be a multilinear form.

By definition,  $f$  is *left-linear* in each one of its arguments, i.e., for any  $\alpha, \beta \in \mathbb{H}$  and for any  $\mathbf{x}_d, \mathbf{y}_d \in \mathcal{S}_d$  with  $d = 1, 2, 3$ , one has

$$(3.1) \quad f(\alpha \mathbf{x}_1 + \beta \mathbf{y}_1, \mathbf{x}_2, \mathbf{x}_3) = \alpha f(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) + \beta f(\mathbf{y}_1, \mathbf{x}_2, \mathbf{x}_3),$$

$$(3.2) \quad f(\mathbf{x}_1, \alpha \mathbf{x}_2 + \beta \mathbf{y}_2, \mathbf{x}_3) = \alpha f(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) + \beta f(\mathbf{x}_1, \mathbf{y}_2, \mathbf{x}_3),$$

$$(3.3) \quad f(\mathbf{x}_1, \mathbf{x}_2, \alpha \mathbf{x}_3 + \beta \mathbf{y}_3) = \alpha f(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) + \beta f(\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}_3).$$

However, when one has to evaluate quantities such as  $f(\alpha \mathbf{x}_1, \beta \mathbf{x}_2, \mathbf{x}_3)$ , one is faced with a disturbing issue: in which order should linearity properties (3.1) – (3.3) be applied? Starting with (3.1) and then (3.2) gives  $f(\alpha \mathbf{x}_1, \beta \mathbf{x}_2, \mathbf{x}_3) = \alpha \beta f(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ , while the reversed order yields  $f(\alpha \mathbf{x}_1, \beta \mathbf{x}_2, \mathbf{x}_3) = \beta \alpha f(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ . Since the product of two quaternions is non-commutative,  $\alpha \beta \neq \beta \alpha$  in general and thus  $\beta \alpha f(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \neq \alpha \beta f(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ . Any other choice of variables leads to similar contradictions. In other terms, without specifying a somewhat arbitrary ordering of the variables  $\mathbf{x}_1, \mathbf{x}_2$  and  $\mathbf{x}_3$  (or equivalently, of quaternion vector spaces  $\mathcal{S}_1, \mathcal{S}_2$  and  $\mathcal{S}_3$ ), the definition of quaternion multilinearity is unpractical as it is.

The issue is not limited to left quaternion vector spaces. Considering  $\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3$  to be right quaternion vector spaces leads to the same conclusions. Mixing left and right quaternion vector spaces does not solve the issue either. To see this, suppose without loss of generality that  $\mathcal{S}_1, \mathcal{S}_2$  are left quaternion vector spaces, while  $\mathcal{S}_3$  is a right quaternion vector space. Quaternion multilinearity between the first and last variables is unambiguous: indeed, for any  $\alpha, \beta \in \mathbb{H}$ , one has  $f(\alpha \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3 \beta) = \alpha f(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3 \beta) = f(\alpha \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \beta = \alpha f(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \beta$ . However, the first two modes suffer from the previous issue since for any  $\alpha, \beta \in \mathbb{H}$ ,  $f(\alpha \mathbf{x}_1, \beta \mathbf{x}_2, \mathbf{x}_3)$  can be equal to  $\alpha \beta f(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$  or  $\beta \alpha f(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ , depending on the order in which linearity properties are applied.

In the next section, we circumvent the issue of general quaternion multilinearity by introducing a special class of multilinear quaternion-valued forms. By carefully specifying the nature of the vector spaces  $\{\mathcal{S}_d\}$ , we can recover fundamental multilinearity properties, similar to those in the real and complex cases. This class forms the foundation for the definition of quaternion tensors in Section 3.2 and the establishment of their properties.

**3.2. Quaternion tensors as  $\mathbb{H}\mathbb{R}$ -multilinear forms.** In order to extend relevant multilinearity-like properties to the quaternion linear form  $f : \mathcal{S}_1 \times \cdots \times \mathcal{S}_D \rightarrow \mathbb{H}$ , it is necessary to restrict the nature of the different vector spaces  $\{\mathcal{S}_d\}$  defining the domain of  $f$ .

**DEFINITION 3.1** ( $\mathbb{H}\mathbb{R}$ -multilinear form). *Let  $D \geq 3$  be an integer and let the vector spaces  $\{\mathcal{S}_d\}$  be such that:  $\mathcal{S}_1$  is a left quaternion vector space;  $\mathcal{S}_d$ ,  $1 < d < D$  are real vector spaces; and such that  $\mathcal{S}_D$  is a right quaternion vector space. Then  $f : \mathcal{S}_1 \times \cdots \times \mathcal{S}_D \rightarrow \mathbb{H}$  is said to be  $\mathbb{H}\mathbb{R}$ -multilinear if*

- $f$  is left-quaternion linear in its first argument, i.e.,

$$(3.4) \quad \forall \alpha, \beta \in \mathbb{H}, \forall \mathbf{x}_1, \mathbf{y}_1 \in \mathcal{S}_1, f(\alpha \mathbf{x}_1 + \beta \mathbf{y}_1, \dots) = \alpha f(\mathbf{x}_1, \dots) + \beta f(\mathbf{y}_1, \dots);$$

- $f$  is real-linear in its  $d$ -th argument,  $1 < d < D$ , i.e.,

$$(3.5) \quad \begin{aligned} \forall \alpha, \beta \in \mathbb{R}, \forall \mathbf{x}_d, \mathbf{y}_d \in \mathcal{S}_d, \\ f(\dots, \alpha \mathbf{x}_d + \beta \mathbf{y}_d, \dots) &= f(\dots, \mathbf{x}_d \alpha + \mathbf{y}_d \beta, \dots) \\ &= \alpha f(\dots, \mathbf{x}_d, \dots) + \beta f(\dots, \mathbf{y}_d, \dots) \\ &= f(\dots, \mathbf{x}_d, \dots) \alpha + f(\dots, \mathbf{y}_d, \dots) \beta; \end{aligned}$$

- $f$  is right-quaternion linear in its last argument, i.e., :

$$(3.6) \quad \forall \alpha, \beta \in \mathbb{H}, \forall \mathbf{x}_D, \mathbf{y}_D \in \mathcal{S}_D, f(\dots, \mathbf{x}_D \alpha + \mathbf{y}_D \beta) = f(\dots, \mathbf{x}_D) \alpha + f(\dots, \mathbf{y}_D) \beta.$$

*Remark 3.2.* Several remarks are in order. The term  $\mathbb{H}\mathbb{R}$ -multilinearity refers to the fact that conditions (3.4)–(3.6) mix between real and quaternion linearity properties. In addition, note that if a function is  $\mathbb{H}\mathbb{R}$ -multilinear, by definition it is also  $\mathbb{R}$ -multilinear. Finally, it is worth noting that the ordering of vector spaces in Definition 3.1 is completely arbitrary, and can be permuted if necessary. The only restriction lies in considering exactly one quaternion left vector space, one quaternion right vector space, and  $D - 2$  real-vector spaces for an arbitrary number of dimensions  $D \geq 3$ .

In the sequel, we restrict our attention to finite dimensional vector spaces  $\{\mathcal{S}_d\}$  with respective dimension  $\dim \mathcal{S}_d = N_d$ . Therefore, without loss of generality, we can assume that  $\mathcal{S}_1 = \mathbb{H}_L^{N_1}$ ,  $\mathcal{S}_d = \mathbb{R}^{N_d}$  for  $1 < d < D$  and  $\mathcal{S}_D = \mathbb{H}_R^{N_D}$ . The next definition defines quaternion tensors as  $\mathbb{H}\mathbb{R}$ -multilinear forms.

**DEFINITION 3.3** (Quaternion tensor). *Let  $f : \mathbb{H}_L^{N_1} \times \mathbb{R}^{N_2} \times \dots \times \mathbb{R}^{N_{D-1}} \times \mathbb{H}_R^{N_D} \rightarrow \mathbb{H}$  be a  $\mathbb{H}\mathbb{R}$ -multilinear form. Let  $\{\mathbf{e}_{i_d}^{(d)}\}_{i_d=1:N_d}$  denote a basis of the vector space  $\mathcal{S}_d$  for  $1 \leq d \leq D$ . Then the quaternion tensor of order  $D$  associated to  $f$  is the  $D$ -dimensional quaternion array  $\mathcal{T} \in \mathbb{H}^{N_1 \times \dots \times N_D}$  defined by its entries as*

$$(3.7) \quad \mathcal{T}_{i_1 i_2 \dots i_D} = f\left(\mathbf{e}_{i_1}^{(1)}, \mathbf{e}_{i_2}^{(2)}, \dots, \mathbf{e}_{i_D}^{(D)}\right)$$

i.e.,  $\mathcal{T}$  is the representation of  $f$  in the Cartesian product basis  $\times_{d=1}^D \{\mathbf{e}_{i_d}^{(d)}\}_{i_d=1:N_d}$ .

*Remark 3.4.* Following standard practice in multilinear algebra, we identify a tensor with its multidimensional array coordinate representation in a given basis.

**Definition 3.3** suggests that different choices of bases for  $\{\mathcal{S}_d\}$  lead to different multidimensional quaternion arrays representations of the same  $\mathbb{H}\mathbb{R}$ -multilinear form  $f$ . Therefore a natural question is how these representations are related to one another. Let  $f : \mathbb{H}_L^{N_1} \times \mathbb{R}^{N_2} \times \dots \times \mathbb{R}^{N_{D-1}} \times \mathbb{H}_R^{N_D} \rightarrow \mathbb{H}$  be a  $\mathbb{H}\mathbb{R}$ -multilinear form and consider two quaternion multidimensional array representations:

$$(3.8) \quad \text{“new” representation } \mathcal{T} \in \mathbb{H}^{N_1 \times \dots \times N_D} \text{ in the basis } \times_{d=1}^D \{\mathbf{e}_{i_d}^{(d)}\}_{i_d=1:N_d},$$

$$(3.9) \quad \text{“old” representation } \tilde{\mathcal{T}} \in \mathbb{H}^{N_1 \times \dots \times N_D} \text{ in the basis } \times_{d=1}^D \{\tilde{\mathbf{e}}_{j_d}^{(d)}\}_{j_d=1:N_d}.$$

Expressing “new” basis vectors as a function of “old” ones one gets

$$(3.10) \quad \mathbf{e}_{i_1}^{(1)} = \sum_{j_1=1}^{N_1} a_{i_1 j_1}^{(1)} \tilde{\mathbf{e}}_{j_1}^{(1)}, \quad a_{i_1 j_1}^{(1)} \in \mathbb{H}, \quad i_1 = 1, \dots, N_1$$

$$(3.11) \quad \mathbf{e}_{i_d}^{(d)} = \sum_{j_d=1}^{N_d} a_{i_d j_d}^{(d)} \tilde{\mathbf{e}}_{j_d}^{(d)}, \quad a_{i_d j_d}^{(d)} \in \mathbb{R}, \quad i_d = 1, \dots, N_d \text{ and } 1 < d < D$$

$$(3.12) \quad \mathbf{e}_{i_D}^{(D)} = \sum_{j_D=1}^{N_D} \tilde{\mathbf{e}}_{j_D}^{(d)} a_{i_D j_D}^{(D)}, \quad a_{i_D j_D}^{(D)} \in \mathbb{H}, \quad i_D = 1, \dots, N_D,$$

on account of the nature of the vector spaces  $\{\mathcal{S}_d\}$ . Plugging expressions (3.10) – (3.12) into the definition (3.7) of  $\mathcal{T}$ , one obtains the fundamental change-of-basis relation for quaternion tensors relating the entries of the tensor  $\mathcal{T}$  to that of  $\tilde{\mathcal{T}}$ :

$$(3.13) \quad \mathcal{T}_{i_1 i_2 \dots i_D} = \sum_{j_1, j_2, \dots, j_D=1}^{N_1, N_2, \dots, N_D} a_{i_1 j_1}^{(1)} a_{i_2 j_2}^{(2)} \dots a_{i_{D-1} j_{D-1}}^{(D-1)} \tilde{\mathcal{T}}_{j_1 j_2 \dots j_D} a_{i_D j_D}^{(D)}.$$

On the one hand, remark the particular position of terms  $a_{i_1 j_1}^{(1)}$  and  $a_{i_D j_D}^{(D)}$ , to the left and to the right of  $\tilde{\mathcal{T}}_{j_1 j_2 \dots j_D}$ , respectively, due to the left (resp. right) linear nature of the associated quaternion vector spaces. On the other hand, the position of central terms  $a_{i_d j_d}^{(d)}$ ,  $1 < d < D$ , with respect to  $\tilde{\mathcal{T}}_{j_1 j_2 \dots j_D}$  has no importance since they are real-valued.

**3.3. Transformations rules for quaternion tensors.** The change-of-basis equation (3.13) is the first example of transformation rules for quaternion tensors. Such transformations rely on the  $\mathbb{H}\mathbb{R}$ -multilinear form representation of quaternion tensors (Definition 3.3) and the linearity properties of the underlying vector spaces. Extending the change-of-basis rule to arbitrary linear transformations in each variable leads to the definition of the  $n$ -mode product, which needs to be carefully specified for quaternion tensors due to  $\mathbb{H}\mathbb{R}$ -multilinearity. From now on, we fix a basis for the  $D$  vector spaces  $\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_D$  and identify tensors with  $D$ -dimensional array of numbers.

**DEFINITION 3.5** ( $n$ -mode product for quaternion tensors). *Let  $\mathcal{T} \in \mathbb{H}^{N_1 \times \dots \times N_D}$  be a quaternion tensor of order  $D$ . The  $n$ -mode-product operation corresponds to multiplying the  $n$ -mode fibers of  $\mathcal{T}$  by a matrix  $\mathbf{U}$  of appropriate size and domain, depending on the considered mode. Explicitly,*

- the 1-mode product is defined as the direct quaternion matrix product by  $\mathbf{U} \in \mathbb{H}^{J \times N_1}$  such that

$$(3.14) \quad (\mathcal{T} \times_1^{\triangleright} \mathbf{U})_{j i_2 \dots i_D} = \sum_{i_1=1}^{N_1} u_{j i_1} \mathcal{T}_{i_1 i_2 \dots i_D},$$

- for  $1 < d < D$ , the  $d$ -mode product is defined as the real matrix product by  $\mathbf{U} \in \mathbb{R}^{J \times N_d}$  such that

$$(3.15) \quad (\mathcal{T} \times_d \mathbf{U})_{i_1 \dots j \dots i_D} = \sum_{i_d=1}^{N_d} \mathcal{T}_{i_1 \dots i_d \dots i_D} u_{j i_d},$$

- the  $D$ -mode product is defined as the reverse quaternion matrix product by  $\mathbf{U} \in \mathbb{H}^{J \times N_D}$  such that

$$(3.16) \quad (\mathcal{T} \times_D^{\triangleleft} \mathbf{U})_{i_1 \dots i_{D-1} j} = \sum_{i_D=1}^{N_D} \mathcal{T}_{i_1 i_2 \dots i_D} u_{j i_D}.$$

Just like for quaternion matrix multiplication, symbols  $\triangleright$  and  $\triangleleft$  indicate the type of quaternion matrix multiplication involved in mode-products when relevant. In addition, one can express  $n$ -mode products in terms of unfoldings of quaternion tensors and quaternion matrix multiplication operations. Following standard practice, we denote by  $\mathbf{T}_{(d)} \in \mathbb{H}^{N_d \times \prod_{i \neq d} N_i}$  the  $d$ -th mode unfolding of a quaternion tensor  $\mathcal{T} \in \mathbb{H}^{N_1 \times \dots \times N_D}$  and use the indexing convention of [19]. The  $n$ -mode product operations

relate to matrix unfoldings as

$$(3.17) \quad \mathcal{T}' = \mathcal{T} \times_1^{\triangleright} \mathbf{U} \iff \mathbf{T}'_{(1)} = \mathbf{U} \cdot_{\triangleright} \mathbf{T}_{(1)},$$

$$(3.18) \quad \mathcal{T}' = \mathcal{T} \times_d \mathbf{U} \iff \mathbf{T}'_{(d)} = \mathbf{U} \mathbf{T}_{(d)},$$

$$(3.19) \quad \mathcal{T}' = \mathcal{T} \times_D^{\triangleleft} \mathbf{U} \iff \mathbf{T}'_{(D)} = \mathbf{U} \cdot_{\triangleleft} \mathbf{T}_{(D)},$$

i.e., the matrix unfoldings inherit the direct and reverse order of the  $n$ -mode product operation.

Proposition 3.6 and Proposition 3.7 show that the [Definition 3.5](#) of  $n$ -modes product permits to preserve the classical properties of successive  $n$ -mode products: (i) commutativity between distinct modes and (ii) matrix composition between identical modes. These desirable properties are directly inherited from the definition a quaternion tensor as a  $\mathbb{H}\mathbb{R}$ -multilinear form.

**PROPOSITION 3.6 (Commutativity).** *Let  $\mathcal{T} \in \mathbb{H}^{N_1 \times N_2 \times \dots \times N_D}$  be a quaternion tensor of order  $D$ . Let  $\mathbf{U}_1 \in \mathbb{H}^{J_1 \times N_1}$  and  $\mathbf{U}_D \in \mathbb{H}^{J_D \times N_D}$  be arbitrary quaternion matrices, and let  $\mathbf{U}_d \in \mathbb{R}^{J_d \times N_d}$  and  $\mathbf{U}_{d'} \in \mathbb{R}^{J_{d'} \times N_{d'}}$  be arbitrary real matrices with  $d \neq d', 1 < d, d' < D$ . The following properties are satisfied:*

$$(3.20) \quad (\mathcal{T} \times_1^{\triangleright} \mathbf{U}_1) \times_D^{\triangleleft} \mathbf{U}_D = (\mathcal{T} \times_D^{\triangleleft} \mathbf{U}_D) \times_1^{\triangleright} \mathbf{U}_1,$$

$$(3.21) \quad (\mathcal{T} \times_1^{\triangleright} \mathbf{U}_1) \times_d \mathbf{U}_d = (\mathcal{T} \times_d \mathbf{U}_d) \times_1^{\triangleright} \mathbf{U}_1,$$

$$(3.22) \quad (\mathcal{T} \times_d \mathbf{U}_d) \times_D^{\triangleleft} \mathbf{U}_D = (\mathcal{T} \times_D^{\triangleleft} \mathbf{U}_D) \times_d \mathbf{U}_d,$$

$$(3.23) \quad (\mathcal{T} \times_d \mathbf{U}_d) \times_{d'} \mathbf{U}_{d'} = (\mathcal{T} \times_{d'} \mathbf{U}_{d'}) \times_d \mathbf{U}_d.$$

**PROPOSITION 3.7 (Composition).** *Let  $\mathcal{T} \in \mathbb{H}^{N_1 \times N_2 \times \dots \times N_D}$  be a quaternion tensor of order  $D$ . Let  $\mathbf{U}$  and  $\mathbf{V}$  be matrices with adequate dimensions and values. The following properties are satisfied:*

$$(3.24) \quad \mathcal{T} \times_1^{\triangleright} \mathbf{U} \times_1^{\triangleright} \mathbf{V} = \mathcal{T} \times_1^{\triangleright} (\mathbf{V} \cdot_{\triangleright} \mathbf{U}),$$

$$(3.25) \quad \mathcal{T} \times_D^{\triangleleft} \mathbf{U} \times_D^{\triangleleft} \mathbf{V} = \mathcal{T} \times_D^{\triangleleft} (\mathbf{V} \cdot_{\triangleleft} \mathbf{U}),$$

$$(3.26) \quad \mathcal{T} \times_d \mathbf{U} \times_d \mathbf{V} = \mathcal{T} \times_d (\mathbf{V} \mathbf{U}), \quad 1 < d < D.$$

Proofs of Proposition 3.6 and Proposition 3.7 are found by direct calculations.

*Remark 3.8 (Change-of-basis).* The change of basis equation [\(3.13\)](#) can be rewritten as successive  $n$ -mode products. Let  $\mathbf{A}^{(1)} \in \mathbb{H}^{N_1 \times N_1}$ ,  $\mathbf{A}^{(2)} \in \mathbb{R}^{N_2 \times N_2}$ ,  $\dots$ ,  $\mathbf{A}^{(D-1)} \in \mathbb{R}^{N_{D-1} \times N_{D-1}}$ ,  $\mathbf{A}^{(D)} \in \mathbb{H}^{N_D \times N_D}$  the change-of-basis matrices. Then the change-of-basis equation can be rewritten as

$$(3.27) \quad \mathcal{T} = \tilde{\mathcal{T}} \times_1^{\triangleright} \mathbf{A}^{(1)} \times_2 \mathbf{A}^{(2)} \dots \times_{D-1} \mathbf{A}^{(D-1)} \times_D^{\triangleleft} \mathbf{A}^{(D)},$$

where  $n$ -mode products can be performed in any order according to [Proposition 3.6](#).

**3.4. Tucker format of quaternion tensors.** Without loss of generality, in the remainder of the paper we focus on the case of third order tensors. The next definition defines the Tucker format for quaternion tensors.

**DEFINITION 3.9 (Tucker format).** *Let  $\mathcal{T} \in \mathbb{H}^{N_1 \times N_2 \times N_3}$  be a third-order quaternion tensor. The tensor  $\mathcal{T}$  is said to be expressed in Tucker format if there exists a core tensor  $\mathcal{S} \in \mathbb{H}^{F_1 \times F_2 \times F_3}$  and matrices  $\mathbf{A} \in \mathbb{H}^{N_1 \times F_1}$ ,  $\mathbf{B} \in \mathbb{R}^{N_2 \times F_2}$  and  $\mathbf{C} \in \mathbb{H}^{N_3 \times F_3}$  such that*

$$(3.28) \quad \mathcal{T} = \mathcal{S} \times_1^{\triangleright} \mathbf{A} \times_2 \mathbf{B} \times_3^{\triangleleft} \mathbf{C}.$$

The Tucker format of  $\mathcal{T}$  is written in short-hand notation as  $\mathcal{T} = \llbracket \mathcal{S}; \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket$ .

Direct calculations show that the entries of a tensor  $\mathcal{T} = \llbracket \mathcal{S}; \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket$  can be written in terms of the entries of the core tensor  $\mathcal{S}$  and factor matrices  $\mathbf{A}, \mathbf{B}, \mathbf{C}$  as

$$(3.29) \quad \mathcal{T}_{i_1 i_2 i_3} = \sum_{f_1=1}^{F_1} \sum_{f_2=1}^{F_2} \sum_{f_3=1}^{F_3} A_{i_1 f_1} B_{i_2 f_2} S_{f_1 f_2 f_3} C_{i_3 f_3},$$

for  $i_1 = 1, \dots, N_1, i_2 = 1, \dots, N_2$  and  $i_3 = 1, \dots, N_3$ . Note that due to quaternion non-commutativity, elements in (3.29) do not commute except for the entries of  $\mathbf{B}$ . For tensors written in Tucker format it is also possible to derive expressions of their mode-unfoldings. Following the unfolding convention of Kolda and Bader [19], if  $\mathcal{T} = \llbracket \mathcal{S}; \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket$  then its matrix unfoldings along the first and last mode read

$$(3.30) \quad \mathbf{T}_{(1)} = \mathbf{A} \cdot_{\triangleright} \mathbf{S}_{(1)} \cdot_{\triangleright} (\mathbf{C} \boxtimes \mathbf{B})^\top,$$

$$(3.31) \quad \mathbf{T}_{(3)} = \mathbf{C} \cdot_{\triangleleft} \mathbf{S}_{(3)} \cdot_{\triangleleft} (\mathbf{B} \boxtimes \mathbf{A})^\top.$$

However, due to quaternion non-commutativity, the unfolding along the central mode can only be expressed in a simple way when the core tensor is real-valued, i.e.,

$$(3.32) \quad \mathbf{T}_{(2)} = \mathbf{B} \mathbf{S}_{(2)} \left( \mathbf{C} \boxtimes_{\triangleleft} \mathbf{A} \right)^\top, \quad \text{when } \mathcal{S} \text{ is real-valued.}$$

#### 4. Quaternion Canonical Polyadic Decomposition.

**4.1. Definition.** If in (3.28) the tensor  $\mathcal{S}$  is diagonal, it enables the derivation of the quaternion equivalent of a well-known tensor model, the Canonical Polyadic Decomposition (CPD). By absorbing the diagonal entries of  $\mathcal{S}$  in the factor matrices  $\mathbf{A}, \mathbf{B}, \mathbf{C}$ , we define the Quaternion Canonical Polyadic Decomposition (Q-CPD) as follows.

DEFINITION 4.1 (Quaternion Canonical Polyadic Decomposition (Q-CPD)). *The Q-CPD of a third-order quaternion tensor  $\mathcal{T} \in \mathbb{H}^{N_1 \times N_2 \times N_3}$  is given element-wise by :*

$$(4.1) \quad \mathcal{T}_{i_1 i_2 i_3} = \sum_{f=1}^F A_{i_1 f} B_{i_2 f} C_{i_3 f},$$

for  $i_1 = 1, \dots, N_1, i_2 = 1, \dots, N_2$  and  $i_3 = 1, \dots, N_3$ . The smallest value of  $F$  for which equality (4.1) holds exactly is called the rank of  $\mathcal{T}$ . By expressing the elements of the decomposition in matrix format, the Q-CPD of  $\mathcal{T}$  can be succinctly expressed as  $\mathcal{T} = \llbracket \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket$ , where  $\mathbf{A} \in \mathbb{H}^{N_1 \times F}$ ,  $\mathbf{B} \in \mathbb{R}^{N_2 \times F}$ ,  $\mathbf{C} \in \mathbb{H}^{N_3 \times F}$ .

*Remark 4.2.* Contrary to quaternion matrices, the rank of a quaternion tensor is unambiguous as defined and thus the left or right distinction does not apply.

Notation for the Q-CPD  $\mathcal{T} = \llbracket \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket$  refers to the fact that the Q-CPD can be interpreted as a Tucker model in which the core tensor can always be chosen to be the identity tensor. In particular, this enables a direct derivation of the unfoldings of the model; see Eqs. (4.2)-(4.4) below. Moreover, similarly to the Tucker model, elements in (4.1) do not commute except for the entries of  $\mathbf{B}$ . First, let us analyze the trivial ambiguities of the Q-CPD.

Let  $\mathcal{T} = \llbracket \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket$  follow a CPD of rank  $F$ . The first trivial ambiguity corresponds to an arbitrary joint permutation of the columns of the factor matrices, i.e.,  $\mathcal{T} = \llbracket \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket = \llbracket \mathbf{A}\mathbf{\Pi}, \mathbf{B}\mathbf{\Pi}, \mathbf{C}\mathbf{\Pi} \rrbracket$  with  $\mathbf{\Pi} \in \mathbb{R}^{F \times F}$  an arbitrary permutation matrix. The

second trivial ambiguity relates to scaling between factors: unlike the real and complex CPD, one has to distinguish between *real* and *quaternion* scaling ambiguities due to factor matrices taking values in different spaces. Define three diagonal scaling matrices  $\mathbf{\Lambda}_A, \mathbf{\Lambda}_C \in \mathbb{H}^{F \times F}$  and  $\mathbf{\Lambda}_B \in \mathbb{R}^{F \times F}$  (with non-zero diagonal entries). Direct calculations show that  $\mathcal{T} = \llbracket \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket = \llbracket \mathbf{A} \cdot_{\triangleright} \mathbf{\Lambda}_A, \mathbf{B} \mathbf{\Lambda}_B, \mathbf{C} \cdot_{\triangleleft} \mathbf{\Lambda}_C \rrbracket$  holds for arbitrary factor matrices  $\mathbf{A}, \mathbf{B}, \mathbf{C}$  if and only if  $(\mathbf{\Lambda}_A \cdot_{\triangleright} \mathbf{\Lambda}_C) \mathbf{\Lambda}_B = \mathbf{I}_F$ , where  $\mathbf{I}_F$  is the  $F \times F$  identity matrix. By exploiting the polar form of scaling coefficients further interpretation is possible.

**PROPOSITION 4.3** (Interpretation of scaling ambiguities). *Consider three diagonal matrices  $\mathbf{\Lambda}_A = \text{diag}(\alpha_1, \dots, \alpha_F) \in \mathbb{H}^{F \times F}$ ,  $\mathbf{\Lambda}_B = \text{diag}(\beta_1, \dots, \beta_F) \in \mathbb{R}^{F \times F}$  and  $\mathbf{\Lambda}_C = \text{diag}(\gamma_1, \dots, \gamma_F) \in \mathbb{H}^{F \times F}$ . Write quaternion entries in polar form such that  $\alpha_f = |\alpha_f| \exp(\boldsymbol{\mu}_{\alpha_f} \varphi_{\alpha_f})$  and  $\gamma_f = |\gamma_f| \exp(\boldsymbol{\mu}_{\gamma_f} \varphi_{\gamma_f})$  where  $|\cdot|$  denotes the modulus and  $(\boldsymbol{\mu}, \varphi)$  the axis and angle of the quaternion, respectively. Then, the scaling condition  $(\mathbf{\Lambda}_A \cdot_{\triangleright} \mathbf{\Lambda}_C) \mathbf{\Lambda}_B = \mathbf{I}_F$  can be equivalently interpreted as the combination of two ambiguities:*

- a modulus scaling ambiguity:  $|\alpha_f| |\beta_f| |\gamma_f| = 1$ ,  $f = 1, \dots, F$
- a angular scaling ambiguity: for  $f = 1, \dots, F$ , if  $\beta_f$  is positive, it implies there is a choice of  $\boldsymbol{\mu}_{\alpha_f}, \varphi_{\alpha_f}$  and  $\boldsymbol{\mu}_{\gamma_f}, \varphi_{\gamma_f}$ , such that  $\boldsymbol{\mu}_{\alpha_f} = \boldsymbol{\mu}_{\gamma_f}$  and  $\varphi_{\alpha_f} + \varphi_{\gamma_f} = 0$ . In the case  $\beta_f$  is negative, it implies there is a choice of  $\boldsymbol{\mu}_{\alpha_f}, \varphi_{\alpha_f}$  and  $\boldsymbol{\mu}_{\gamma_f}, \varphi_{\gamma_f}$ , such that  $\boldsymbol{\mu}_{\alpha_f} = \boldsymbol{\mu}_{\gamma_f}$  and  $\varphi_{\alpha_f} + \varphi_{\gamma_f} = \pi$ .

*i.e., moduli must multiply to 1 and arguments of quaternion factors compensate each other according to the sign of  $\beta_f$ .*

*Proof.* Starting from  $(\mathbf{\Lambda}_A \cdot_{\triangleright} \mathbf{\Lambda}_C) \mathbf{\Lambda}_B = \mathbf{I}_F$  we get  $\alpha_f \beta_f \gamma_f = 1$  for every  $f = 1, \dots, F$ . Using polar forms, the latter reads  $|\alpha_f| \exp(\boldsymbol{\mu}_{\alpha_f} \varphi_{\alpha_f}) \beta_f |\gamma_f| \exp(\boldsymbol{\mu}_{\gamma_f} \varphi_{\gamma_f}) = 1$ . Taking the modulus of this expression one gets  $|\alpha_f| |\beta_f| |\gamma_f| = 1$ . For the argument, we have  $\exp(\boldsymbol{\mu}_{\alpha_f} \varphi_{\alpha_f}) \exp(\boldsymbol{\mu}_{\gamma_f} \varphi_{\gamma_f}) = \frac{\beta_f}{|\beta_f|}$ . By considering separately the cases  $\beta_f > 0$  and  $\beta_f < 0$  one obtains the desired property.  $\square$

These ambiguities play a crucial role to define the uniqueness of the Q-CPD; see subsection 4.3 for more details. Finally, note that handling the Q-CPD trivial ambiguities does not pose any challenge as long as non-commutativity of quaternion multiplication is taken into account.

The matrix unfoldings of the quaternion CPD model follow directly from the unfoldings (3.30) – (3.32) of the Tucker model. As already mentioned, the Q-CPD can be viewed as a special case of the Tucker model with a core identity tensor; therefore the unfoldings of  $\mathcal{T} = \llbracket \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket$  read

$$(4.2) \quad \mathbf{T}_{(1)} = \mathbf{A} \cdot_{\triangleright} (\mathbf{C} \odot \mathbf{B})^{\top},$$

$$(4.3) \quad \mathbf{T}_{(2)} = \mathbf{B} \left( \mathbf{C} \odot_{\triangleleft} \mathbf{A} \right)^{\top},$$

$$(4.4) \quad \mathbf{T}_{(3)} = \mathbf{C} \cdot_{\triangleleft} (\mathbf{B} \odot \mathbf{A})^{\top}.$$

Once again, these matrix unfoldings are similar to those of the real and complex cases (see, e.g., [19]) yet matrix and Khatri-Rao products have direct or reverse orderings due to non-commutativity of the quaternion product. Moreover, these unfoldings are essential to the practical computation of the factor matrices  $\mathbf{A}, \mathbf{B}$  and  $\mathbf{C}$ ; see Section 5 for further details.

To conclude this section, note that it is also possible to derive mathematical expressions for the matrix slices of the Q-CPD (4.1). The frontal slices  $\mathbf{T}_{::i_3} \in \mathbb{H}^{N_1 \times N_2}$ , lateral slices  $\mathbf{T}_{:i_2:} \in \mathbb{H}^{N_1 \times N_3}$  and horizontal slices  $\mathbf{T}_{i_1:} \in \mathbb{H}^{N_2 \times N_3}$  of  $\mathcal{T} = \llbracket \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket$

can be written in terms of factor matrices as

$$(4.5) \quad \mathbf{T}_{:i_1:} = \mathbf{B} \left( \mathbf{D}_{\mathbf{A}}^{(i_1)} \cdot_{\triangleright} \mathbf{C}^{\top} \right), \quad \mathbf{D}_{\mathbf{A}}^{(i_1)} := \text{diag}(\mathbf{A}_{i_1:}) \in \mathbb{H}^{F \times F}, \quad i_1 = 1, \dots, N_1,$$

$$(4.6) \quad \mathbf{T}_{:i_2:} = \left( \mathbf{A} \mathbf{D}_{\mathbf{B}}^{(i_2)} \right) \cdot_{\triangleright} \mathbf{C}^{\top}, \quad \mathbf{D}_{\mathbf{B}}^{(i_2)} := \text{diag}(\mathbf{B}_{i_2:}) \in \mathbb{R}^{F \times F}, \quad i_2 = 1, \dots, N_2,$$

$$(4.7) \quad \mathbf{T}_{::i_3} = \left( \mathbf{A} \cdot_{\triangleright} \mathbf{D}_{\mathbf{C}}^{(i_3)} \right) \mathbf{B}^{\top}, \quad \mathbf{D}_{\mathbf{C}}^{(i_3)} := \text{diag}(\mathbf{C}_{i_3:}) \in \mathbb{H}^{F \times F}, \quad i_3 = 1, \dots, N_3.$$

These expressions can be deduced from the definition (4.1) of the Q-CPD. They will be useful in the next section to study equivalent models of the Q-CPD.

**4.2. Equivalent models.** In this subsection, we propose different ways to rewrite the Q-CPD as structured or constrained decompositions in the complex domain. Two main approaches are investigated. The first one models complex adjoints of lateral slices of  $\mathcal{T}$  while the second one models separately the two Cayley-Dickson parts of of  $\mathcal{T}$ . The former will allow us to establish uniqueness properties of the Q-CPD in [Subsection 4.3](#). The latter will be useful in [Section 5](#) to derive a simple complex-domain Q-CPD algorithm.

**4.2.1. Complex adjoint representation.** Consider a third-order tensor  $\mathcal{T} \in \mathbb{H}^{N_1 \times N_2 \times N_3}$  following a rank- $F$  quaternion CPD such that  $\mathcal{T} = \llbracket \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket$ . Recall that the lateral slices of  $\mathcal{T}$  are given by (4.6) and let us write  $\mathbf{A} = \mathbf{A}_1 + \mathbf{A}_2 \mathbf{j}$  and  $\mathbf{C} = \mathbf{C}_1 + \mathbf{C}_2 \mathbf{j}$ , where  $\mathbf{A}_1, \mathbf{A}_2 \in \mathbb{C}_i^{N_1 \times F}$  and  $\mathbf{C}_1, \mathbf{C}_2 \in \mathbb{C}_i^{N_3 \times F}$ . Then, the direct complex adjoint of the  $i_2$ -th lateral slice (4.6) is given by

$$(4.8) \quad \begin{aligned} \chi_{\triangleright}(\mathbf{T}_{:i_2:}) &= \chi_{\triangleright} \left( \left( \mathbf{A} \cdot_{\triangleright} \mathbf{D}_{\mathbf{B}}^{(i_2)} \right) \cdot_{\triangleright} \mathbf{C}^{\top} \right) \\ &= \chi_{\triangleright}(\mathbf{A}) \chi_{\triangleright}(\mathbf{D}_{\mathbf{B}}^{(i_2)}) \chi_{\triangleright}(\mathbf{C}^{\top}) \\ &= \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 \\ -\overline{\mathbf{A}_2} & \overline{\mathbf{A}_1} \end{bmatrix} \begin{bmatrix} \text{diag}(\mathbf{B}_{i_2:}) & 0 \\ 0 & \text{diag}(\mathbf{B}_{i_2:}) \end{bmatrix} \begin{bmatrix} \mathbf{C}_1^{\top} & \mathbf{C}_2^{\top} \\ -\mathbf{C}_2^{\text{H}} & \mathbf{C}_1^{\text{H}} \end{bmatrix}. \end{aligned}$$

Let us denote by  $\mathcal{T}_{\mathbb{C}}$  the complex-valued tensor such that its frontal slices<sup>2</sup> are  $\chi_{\triangleright}(\mathbf{T}_{:i_2:})$  for  $i_2 = 1, \dots, N_2$ . It is a structured tensor of dimensions  $2N_1 \times 2N_3 \times N_2$ . The expression (4.8) of its (now) frontal slices shows that  $\mathcal{T}_{\mathbb{C}}$  follows a standard complex CPD model with colinearities. Direct calculations yield, after reorganizing terms, the expression

$$(4.9) \quad \mathcal{T}_{\mathbb{C}} = \sum_{f=1}^F \chi_{\triangleright}(\mathbf{a}_f) \chi_{\triangleright}(\mathbf{c}_f^{\top}) \circ \mathbf{b}_f = \sum_{f=1}^F \chi_{\triangleright}(\mathbf{a}_f) \chi_{\triangleleft}^{\top}(\mathbf{c}_f) \circ \mathbf{b}_f,$$

where  $\circ$  is the usual outer product. Eq. (4.9) shows that  $\mathcal{T}_{\mathbb{C}}$  follows a decomposition into  $F$  rank-(2, 2, 1) terms [9]. This result is summarized in the following proposition.

**PROPOSITION 4.4.** *Let  $\mathcal{T} \in \mathbb{H}^{N_1 \times N_2 \times N_3}$  be a third-order tensor as defined in [Definition 3.3](#) and denote by  $\mathcal{T}_{\mathbb{C}} \in \mathbb{C}_i^{2N_1 \times 2N_3 \times N_2}$  the tensor having frontal slices given by complex adjoints of lateral slices of  $\mathcal{T}$ . Then  $\mathcal{T}$  follows a rank- $F$  quaternion CPD if and only if  $\mathcal{T}_{\mathbb{C}}$  follows a structured rank-(2, 2, 1) decomposition of the form (4.9).*

*Proof.* The result holds by isomorphism between the set of quaternions matrices equipped with the direct quaternion matrix product and the set of direct complex adjoints.  $\square$

<sup>2</sup>Note that the permutation of the second mode (real) and the third mode (quaternionic) is for ease of presentation only and does not affect the generality of the equivalence.

Proposition 4.4 is particularly useful to establish uniqueness results for the quaternion CPD; see Subsection 4.3 below.

It is worth noting that the ambiguities of the decomposition (4.9) are more restrictive than those of the standard unconstrained rank- $(L, L, 1)$  decomposition [9]. Indeed consider such an arbitrary decomposition of a complex tensor  $\mathcal{T} = \sum_{f=1}^F \mathbf{U}_f \mathbf{W}_f^\top \circ \mathbf{v}_f$ , where  $\mathbf{U}_f, \mathbf{W}_f$  have full column rank  $L$ , and where  $\mathbf{v}_f$  are vectors. Then it is well known that only the low-rank  $\mathbf{U}_f \mathbf{W}_f^\top$  subspaces can be identified since  $\mathbf{U}_f \mathbf{W}_f^\top = \mathbf{U}_f \mathbf{R}_U (\mathbf{V}_f \mathbf{R}_V)^\top$  where  $\mathbf{R}_U, \mathbf{R}_V \in \mathbb{C}^{L \times L}$  are such that  $\mathbf{R}_U \mathbf{R}_V^\top = \mathbf{I}_L$ . For the structured rank- $(2, 2, 1)$  decomposition (4.9), the situation is more restrictive with the help of the following lemma.

**LEMMA 4.5.** *Let  $\mathbf{p}, \tilde{\mathbf{p}} \in \mathbb{H}^M$  and  $\mathbf{q}, \tilde{\mathbf{q}} \in \mathbb{H}^N$ . Then  $\chi_{\triangleright}(\mathbf{p})\chi_{\triangleleft}^\top(\mathbf{q}) = \chi_{\triangleright}(\tilde{\mathbf{p}})\chi_{\triangleleft}^\top(\tilde{\mathbf{q}})$  if and only if there exists a non-zero quaternion  $r \in \mathbb{H}$  such that  $\tilde{\mathbf{p}} = \mathbf{p}r$  and  $\tilde{\mathbf{q}} = r^{-1}\mathbf{q}$ .*

*Proof.* Using complex adjoint properties, one has  $\chi_{\triangleright}(\mathbf{p})\chi_{\triangleleft}^\top(\mathbf{q}) = \chi_{\triangleright}(\mathbf{p})\chi_{\triangleright}(\mathbf{q}^\top) = \chi_{\triangleright}(\mathbf{p}\mathbf{q}^\top)$ . Thus  $\chi_{\triangleright}(\mathbf{p})\chi_{\triangleleft}^\top(\mathbf{q}) = \chi_{\triangleright}(\tilde{\mathbf{p}})\chi_{\triangleleft}^\top(\tilde{\mathbf{q}})$  if and only if  $\chi_{\triangleright}(\mathbf{p}\mathbf{q}^\top) = \chi_{\triangleright}(\tilde{\mathbf{p}}\tilde{\mathbf{q}}^\top)$ . By the properties of the complex adjoint, this is equivalent to saying that  $\mathbf{p}\mathbf{q}^\top = \tilde{\mathbf{p}}\tilde{\mathbf{q}}^\top$ . The equality is satisfied if and only if there exists a non-zero  $r \in \mathbb{H}$  such that  $\tilde{\mathbf{p}} = \mathbf{p}r$  and  $\tilde{\mathbf{q}}^\top = r^{-1}\mathbf{q}^\top$ . Since  $r$  is a scalar, this equivalently means that  $\tilde{\mathbf{q}} = r^{-1}\mathbf{q}$  and it concludes the proof.  $\square$

Applying Lemma 4.5 to the decomposition (4.9), one obtains that  $\chi_{\triangleright}(\mathbf{a}_f)\chi_{\triangleleft}^\top(\mathbf{c}_f) = \chi_{\triangleright}(\mathbf{a}_f)\mathbf{R}_A(\chi_{\triangleleft}(\mathbf{c}_f)\mathbf{R}_C)^\top$  if and only if there exists a non-zero scalar  $r_f \in \mathbb{H}$  such that  $\mathbf{R}_A = \chi_{\triangleright}(r_f)$  and  $\mathbf{R}_C^\top = \chi_{\triangleright}(r_f^{-1})$ , or equivalently,  $\mathbf{R}_C = \chi_{\triangleright}^\top(r_f^{-1}) = \chi_{\triangleleft}(r_f^{-1})$ . In other terms, the ambiguities of the rank- $(2, 2, 1)$  factors are directly inherited from those of the quaternion factors of the Q-CPD.

**4.2.2. Coupled CONFAC model.** Suppose again that  $\mathcal{T} = \llbracket \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket$  follows a rank- $F$  Q-CPD and consider its Cayley-Dickson decomposition as  $\mathcal{T} = \mathcal{T}_1 + \mathcal{T}_2 \mathbf{j}$ , where  $\mathcal{T}_1, \mathcal{T}_2 \in \mathbb{C}_i^{N_1 \times N_2 \times N_3}$ . Then, another approach to build a complex representation of the Q-CPD of  $\mathcal{T}$  consists in modeling separately  $\mathcal{T}_1$  and  $\mathcal{T}_2$  and interpreting the Q-CPD as a coupled decomposition of these two components. In the following we define as a constraint matrix any matrix that is full row rank and whose each column have one, and only one, non-null entry. We then say that a third order tensor  $\mathcal{X} \in \mathbb{C}_i^{N_1 \times N_2 \times N_3}$  admits a CONFAC decomposition [10] (or follows a CONFAC model) of rank  $F_c$  if there exist three matrices,  $\mathbf{U}, \mathbf{V}$  and  $\mathbf{W}$  of sizes  $(N_1, L_1), (N_2, L_2), (N_3, L_3)$  respectively and three constraint matrices,  $\Psi, \Phi, \Omega$  of sizes  $(L_1, F_c), (L_2, F_c), (L_3, F_c)$  such that

$$(4.10) \quad \mathcal{X} = \llbracket \mathbf{U}\Psi, \mathbf{V}\Phi, \mathbf{W}\Omega \rrbracket.$$

Provided that the constraint matrices  $\Psi, \Phi, \Omega$  are known, the factor matrices  $\mathbf{U}, \mathbf{V}$  and  $\mathbf{W}$  are easily identified by means of an ALS algorithm [6]. Note that the CONFAC model is closely related to the PARALIND model [2]. Theorem 4.6 states that the Q-CPD implies that each component follows a CONFAC decomposition model.

**THEOREM 4.6** (Link between the Q-CPD and the complex CONFAC decomposition). *Let  $\mathcal{T}$  be a third order tensor of quaternions of size  $(N_1, N_2, N_3)$ . Let  $\llbracket \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket$  be its rank- $F$  Q-CPD and  $\mathcal{T}_1$  and  $\mathcal{T}_2$  the two  $\mathbb{C}_i$ -valued tensors such that  $\mathcal{T} = \mathcal{T}_1 + \mathcal{T}_2 \mathbf{j}$ . Then,  $\mathcal{T}_1$  and  $\mathcal{T}_2$  each admits a CONFAC decomposition of rank  $2F$  given by*

$$(4.11a) \quad \mathcal{T}_1 = \llbracket \mathbf{U}\Psi, \mathbf{B}\Phi, \overline{\mathbf{W}}\mathbf{I}_{2F} \rrbracket,$$

$$(4.11b) \quad \mathcal{T}_2 = \llbracket \mathbf{U}\mathbf{I}_{2F}, \mathbf{B}\Phi, \mathbf{W}\mathbf{I}_{2F} \rrbracket,$$

where the factor matrices  $\mathbf{U}$  and  $\mathbf{W}$  read in terms of Cayley-Dickson factors of  $\mathbf{A}$  and  $\mathbf{C}$

$$(4.12) \quad \mathbf{U} = [\mathbf{A}_1 \quad \mathbf{A}_2], \quad \mathbf{W} = [\mathbf{C}_2 \quad \bar{\mathbf{C}}_1],$$

and where the constraint matrices  $\Psi$  and  $\Phi$  are given by

$$(4.13) \quad \Psi = \begin{bmatrix} 0 & \mathbf{I}_F \\ -\mathbf{I}_F & 0 \end{bmatrix}, \quad \Phi = [\mathbf{I}_F \quad \mathbf{I}_F].$$

*Proof.* Consider the Cayley-Dickson decomposition of quaternion factor matrices as  $\mathbf{A} = \mathbf{A}_1 + \mathbf{A}_2\mathbf{j}$  and  $\mathbf{C} = \mathbf{C}_1 + \mathbf{C}_2\mathbf{j}$ . Then, the Q-CPD of  $\mathcal{T} = \llbracket \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket$  rewrites as

$$\mathcal{T} = \sum_{f=1}^F (\mathbf{a}_{1f} + \mathbf{a}_{2f}\mathbf{j}) \circ \mathbf{b}_f \circ (\mathbf{c}_{1f} + \mathbf{c}_{2f}\mathbf{j})$$

where  $\mathbf{a}_{1f}$  is the column  $f$  of matrix  $\mathbf{A}_1$ . This leads to

$$(4.14) \quad \mathcal{T}_1 = \sum_{f=1}^F \mathbf{a}_{1f} \circ \mathbf{b}_f \circ \mathbf{c}_{1f} - \mathbf{a}_{2f} \circ \mathbf{b}_f \circ \bar{\mathbf{c}}_{2f},$$

$$(4.15) \quad \mathcal{T}_2 = \sum_{f=1}^F \mathbf{a}_{1f} \circ \mathbf{b}_f \circ \mathbf{c}_{2f} + \mathbf{a}_{2f} \circ \mathbf{b}_f \circ \bar{\mathbf{c}}_{1f}.$$

Thus, equation (4.14) decomposes  $\mathcal{T}_1$  as a sum of two CPDs and there is an unfolding of  $\mathcal{T}_1$  of size  $(N_2, N_3 N_1)$ , denoted  $\mathbf{T}_1^B$  and such that

$$(4.16) \quad \begin{aligned} \mathbf{T}_1^B &= \mathbf{B}((-\mathbf{A}_2 \odot \bar{\mathbf{C}}_2)^\top + (\mathbf{A}_1 \odot \mathbf{C}_1)^\top) \\ &= [\mathbf{B} \quad \mathbf{B}] [-\mathbf{A}_2 \odot \bar{\mathbf{C}}_2 \quad \mathbf{A}_1 \odot \mathbf{C}_1]^\top \\ &= [\mathbf{B} \quad \mathbf{B}] [[-\mathbf{A}_2 \quad \mathbf{A}_1] \odot [\bar{\mathbf{C}}_2 \quad \mathbf{C}_1]]^\top. \end{aligned}$$

Introducing  $\mathbf{U} = [\mathbf{A}_1 \quad \mathbf{A}_2]$  and  $\mathbf{W} = [\mathbf{C}_2 \quad \bar{\mathbf{C}}_1]$ , Equation (4.16) rewrites

$$(4.17) \quad \mathbf{T}_1^B = \mathbf{B}\Phi(\mathbf{U}\Psi \odot \bar{\mathbf{W}})^\top$$

which is equivalent to equation (4.11a). The same reasoning holds for  $\mathcal{T}_2$  and leads to

$$(4.18) \quad \mathbf{T}_2^B = \mathbf{B}\Phi(\mathbf{U} \odot \mathbf{W})^\top$$

which is equivalent to equation (4.11b). Eventually, one can immediately check that  $\Psi$ ,  $\Phi$ , and  $\mathbf{I}_{2F}$  have full row rank and their columns have one and only one non-null entry.  $\square$

The system of equations (4.11a) and (4.11b) (or (4.17) and (4.18)) defines a coupled CONFAC decomposition<sup>3</sup>. It is worth mentioning that this decomposition was introduced by [7, section IV.A] in a totally different context. The main advantage of this model with respect to the complex adjoint representation of equation (4.9) is that it only involves the  $2N_1N_2N_3$  complex equations that are directly induced by the

<sup>3</sup>One can easily check that a similar reasoning applies to the four real components of  $\mathcal{T}$  leading to a real coupled CONFAC decomposition of the four tensors.

Q-CPD model. It is thus more suitable for calculating the Q-CPD in the complex field. Moreover, it has been shown in [7] that it can be computed using a simple algorithm, from which we have drawn inspiration for the Q-CPD algorithm presented in Section 5.2.

In addition, we will show in Subsection 4.3 that under some assumptions, the real factor matrix can be identified from any of the two CONFAC models involved in the coupled CONFAC decomposition.

**4.3. Uniqueness of the Q-CPD.** As explained in Subsection 4.1, the Q-CPD exhibits trivial scaling and permutation ambiguities, leading to the following definition of uniqueness.

**DEFINITION 4.7** (Uniqueness of the Q-CPD). *Let  $\mathcal{T} = \llbracket \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket$  admit a rank- $F$  Q-CPD. We say that the Q-CPD is essentially unique (or simply, unique) if the only factor matrices  $\tilde{\mathbf{A}}, \tilde{\mathbf{B}}, \tilde{\mathbf{C}}$  such that  $\mathcal{T} = \llbracket \tilde{\mathbf{A}}, \tilde{\mathbf{B}}, \tilde{\mathbf{C}} \rrbracket$  are related through*

$$(4.19) \quad \tilde{\mathbf{A}} = \mathbf{A} \cdot_{\triangleright} (\Lambda_{\mathbf{A}} \Pi), \quad \tilde{\mathbf{B}} = \mathbf{B} \Lambda_{\mathbf{B}} \Pi, \quad \tilde{\mathbf{C}} = \mathbf{C} \cdot_{\triangleleft} (\Lambda_{\mathbf{C}} \Pi),$$

where  $\Pi \in \mathbb{R}^{F \times F}$  is a permutation matrix,  $\Lambda_{\mathbf{A}}, \Lambda_{\mathbf{C}}$  are quaternion-valued diagonal scaling matrices and where  $\Lambda_{\mathbf{B}}$  is a real diagonal matrix such that  $(\Lambda_{\mathbf{A}} \cdot_{\triangleright} \Lambda_{\mathbf{C}}) \Lambda_{\mathbf{B}} = \mathbf{I}_F$ .

Now, uniqueness results for the quaternion CPD can be readily obtained through the equivalence with the structured rank-(2, 2, 1) decomposition (4.9) (see also Proposition 4.4). This statement is made precise in the following proposition, which is a direct consequence of Proposition 4.4.

**PROPOSITION 4.8.** *Let  $\mathcal{T} \in \mathbb{H}^{N_1 \times N_2 \times N_3}$  be a third-order tensor and denote by  $\mathcal{T}_{\mathbb{C}} \in \mathbb{C}_i^{2N_1 \times N_2 \times 2N_3}$  the tensor constructed from direct complex adjoints of lateral slices of  $\mathcal{T}$ . Suppose that  $\mathcal{T}$  follows a rank- $F$  quaternion CPD (4.1). Then the quaternion CPD is unique if and only if the structured rank-(2,2,1) decomposition (4.9) of  $\mathcal{T}_{\mathbb{C}}$  is unique.*

The proof is straightforward by observing that the mapping between the factor matrices of  $\mathcal{T}$  and those of  $\mathcal{T}_{\mathbb{C}}$  is one-to-one. Building on this result, Theorem 4.9 below establishes a sufficient uniqueness condition for the Q-CPD.

**THEOREM 4.9.** *Let  $\mathbf{A} \in \mathbb{H}^{N_1 \times F}$ ,  $\mathbf{B} \in \mathbb{R}^{N_2 \times F}$  and  $\mathbf{C} \in \mathbb{H}^{N_3 \times F}$  be factor matrices. Suppose that (at least) one of the following conditions is satisfied:*

1.  $\text{rank}_{\mathbb{R}} \mathbf{A} = F$ ,  $\text{rank}_{\mathbb{L}} \mathbf{C} = F$  and  $\mathbf{B}$  does not have proportional columns;
2.  $k(\mathbf{B}) = F$  and  $k_{\mathbb{R}}(\mathbf{A}) + k_{\mathbb{L}}(\mathbf{C}) \geq F + 2$ ;
3.  $k_{\mathbb{R}}(\mathbf{A}) = F$  and  $k(\mathbf{B}) + k_{\mathbb{L}}(\mathbf{C}) \geq F + 2$ ;
4.  $k_{\mathbb{L}}(\mathbf{C}) = F$  and  $k_{\mathbb{R}}(\mathbf{A}) + k(\mathbf{B}) \geq F + 2$ ;
5.  $N_1 N_3 \geq F$  and  $k_{\mathbb{R}}(\mathbf{A}) + k(\mathbf{B}) + k_{\mathbb{L}}(\mathbf{C}) \geq 2F + 2$ .

*Then the Q-CPD  $\mathcal{T} = \llbracket \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket$  is unique in the sense of Definition 4.7.*

*Proof.* The sufficient uniqueness conditions for the Q-CPD derive directly from the usual uniqueness conditions for rank-( $L, L, 1$ ) decompositions [9], using the equivalence of Subsection 4.2.1 and our Lemma 2.10. Let  $\mathcal{T} = \llbracket \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket$  and construct the complex adjoint tensor  $\mathcal{T}_{\mathbb{C}}$  as in (4.9). From Proposition 4.8 the Q-CPD of  $\mathcal{T}$  is unique if and only if the structured rank-(2,2,1) decomposition of  $\mathcal{T}_{\mathbb{C}}$  is unique. Following the approach in [9], remark that the horizontal stack of rank-2 factors simply corresponds to the direct (resp. reverse) column-wise complex adjoint (2.9) of the matrix  $\mathbf{A}$  (resp.

$\mathbf{C}$ ), i.e.,

$$\begin{aligned}\chi_{\triangleright}^{\pi}(\mathbf{A}) &= [\chi_{\triangleright}(\mathbf{a}_1) \quad \cdots \quad \chi_{\triangleright}(\mathbf{a}_F)] \in \mathbb{H}^{2N_1 \times 2F}, \\ \chi_{\triangleleft}^{\pi}(\mathbf{C}) &= [\chi_{\triangleleft}(\mathbf{c}_1) \quad \cdots \quad \chi_{\triangleleft}(\mathbf{c}_F)] \in \mathbb{H}^{2N_3 \times 2F}.\end{aligned}$$

Recall from [Remark 2.2](#) that there exists a permutation matrix  $\mathbf{\Pi} \in \mathbb{R}^{2F \times 2F}$  such that  $\chi_{\triangleright}^{\pi}(\mathbf{A}) = \chi_{\triangleright}(\mathbf{A})\mathbf{\Pi}$  and  $\chi_{\triangleleft}^{\pi}(\mathbf{C}) = \chi_{\triangleleft}(\mathbf{C})\mathbf{\Pi}$ . Therefore uniqueness of the rank-(2, 2, 1) decomposition of  $\mathcal{T}_{\mathbb{C}}$  can directly be characterized in terms of matrices  $\chi_{\triangleright}(\mathbf{A}), \chi_{\triangleleft}(\mathbf{C})$  and  $\mathbf{B}$ . Moreover, from the properties of the complex adjoints we have that  $\text{rank } \chi_{\triangleright}(\mathbf{A}) = \text{rank } \chi_{\triangleright}^{\pi}(\mathbf{A}) = 2F$  if and only if  $\text{rank}_{\mathbb{R}} \mathbf{A} = F$  ([Proposition 2.7](#)) and  $k'(\chi_{\triangleright}^{\pi}(\mathbf{A})) = k_{\mathbb{R}}(\mathbf{A})$  (by [Lemma 2.10](#)). Similarly for  $\mathbf{C}$  one has  $\text{rank } \chi_{\triangleleft}(\mathbf{C}) = \text{rank } \chi_{\triangleleft}^{\pi}(\mathbf{C}) = 2F$  if and only if  $\text{rank}_{\mathbb{L}} \mathbf{C} = F$  and  $k'(\chi_{\triangleleft}^{\pi}(\mathbf{C})) = k_{\mathbb{L}}(\mathbf{C})$ . It now suffices to apply, in order, Theorems 4.1, 4.4, 4.5(a), 4.5(b) and 4.7 [9] to the rank-(2, 2, 1) decomposition of  $\mathcal{T}_{\mathbb{C}}$  to obtain conditions 1., 2., 3., 4., and 5.  $\square$

*Remark 4.10.* Condition 5 in [Theorem 4.9](#) resembles the usual Kruskal uniqueness condition for the CPD in the real and complex cases, up to (i) the necessary distinction between left and right Kruskal ranks of a quaternion matrix; and (ii) the (usually non-restrictive) condition  $N_1 N_3 \geq F$ , i.e., that the product of the dimensions of the two quaternion-valued factors is at least equal to the rank of the tensor. In particular, when  $F \leq \max(N_1, N_3)$ , the condition is automatically satisfied and Condition 5 boils down to the well-known Kruskal condition generalized to the quaternion case.

*Remark 4.11.* For a generic quaternion matrix, left and right k-ranks are both equal, i.e., if  $\mathbf{A} \in \mathbb{H}^{M \times F}$  is generic, then  $k_{\mathbb{L}}(\mathbf{A}) = k_{\mathbb{R}}(\mathbf{A}) = \min(M, F)$ . This results directly from the definition of complex matrix adjoints (2.7) – (2.8) and [Lemma 2.10](#).

Eventually, [Proposition 4.12](#) states that the CONFAC decomposition of only one component of the Cayley-Dickson decomposition of  $\mathcal{T}$  is likely to be sufficient to identify the real factor matrix of the Q-CPD but insufficient to identify the quaternion factor matrices.

**PROPOSITION 4.12 (Essential uniqueness of  $\mathbf{B}$ ).** *Let  $\mathcal{T}$  be a third order tensor of quaternions such that  $\mathcal{T} = \mathcal{T}_1 + \mathcal{T}_2 \mathbf{j}$ . Let  $\llbracket \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket$  be the rank- $F$  Q-CPD of  $\mathcal{T}$  and  $\llbracket \mathbf{U}\Psi, \mathbf{B}\Phi, \mathbf{W}\mathbf{I}_{2F} \rrbracket$  the CONFAC decomposition of  $\mathcal{T}_1$  given in [Theorem 4.6](#). If  $\mathbf{B}, \mathbf{U}$  and  $\mathbf{W}$  are full column rank then  $\mathbf{B}$  is essentially unique.*

A proof is given in the Appendix. This result also holds for  $\mathcal{T}_2$  and Equation (4.11b).

**5. Algorithms for computing the quaternion CPD.** Given an arbitrary tensor  $\mathcal{T} \in \mathbb{H}^{N_1 \times N_2 \times N_3}$ , the Q-CPD approximation problem of rank- $F$  can be formulated as the following optimization problem

$$(5.1) \quad (\hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}}) = \arg \min_{\mathbf{A}, \mathbf{B}, \mathbf{C}} \|\mathcal{T} - \llbracket \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket\|_F^2.$$

To solve (5.1), we introduce two algorithms that operate in the quaternion domain and the complex domain, respectively. Both algorithms rely on the well known Alternating Least Squares procedure (ALS). As we will show, they yield equivalent update rules; however subtle differences arise in their practical implementation, as discussed in [Section 6](#).

**5.1. Computing the Q-CPD in  $\mathbb{H}$ : the Q-ALS algorithm.** The Q-ALS algorithm extends the ALS algorithm to the Q-CPD. It resorts to the matrix unfoldings

of the Q-CPD given in equations (4.2)-(4.4). Thereby, at each iteration the factor matrices are updated as:

$$(5.2) \quad \mathbf{A} \leftarrow \arg \min_{\mathbf{A}} \|\mathbf{T}_{(1)} - \mathbf{A} \cdot_{\triangleright} (\mathbf{C} \odot \mathbf{B})^{\top}\|_F^2,$$

$$(5.3) \quad \mathbf{B} \leftarrow \arg \min_{\mathbf{B}} \|\mathbf{T}_{(2)} - \mathbf{B}(\mathbf{C} \odot_{\triangleleft} \mathbf{A})^{\top}\|_F^2 \quad \text{s.t. } \mathbf{B} \text{ has real-valued entries,}$$

$$(5.4) \quad \mathbf{C} \leftarrow \arg \min_{\mathbf{C}} \|\mathbf{T}_{(3)} - \mathbf{C} \cdot_{\triangleleft} (\mathbf{B} \odot \mathbf{A})^{\top}\|_F^2,$$

where  $\|\cdot\|_F$  denotes the Frobenius norm of a matrix. We thus have to solve three different kinds of quaternion matrix least squares subproblems:

$$(5.5) \quad \hat{\mathbf{X}}_{\triangleright} = \arg \min_{\mathbf{X}} \|\mathbf{M} - \mathbf{X} \cdot_{\triangleright} \mathbf{N}\|_F^2,$$

$$(5.6) \quad \hat{\mathbf{X}} = \arg \min_{\mathbf{X}} \|\mathbf{M} - \mathbf{X}\mathbf{N}\|_F^2 \quad \text{s.t. } \mathbf{X} \text{ has real-valued entries,}$$

$$(5.7) \quad \hat{\mathbf{X}}_{\triangleleft} = \arg \min_{\mathbf{X}} \|\mathbf{M} - \mathbf{X} \cdot_{\triangleleft} \mathbf{N}\|_F^2.$$

The solution to (5.5) can be found in [36] and is given by  $\hat{\mathbf{X}}_{\triangleright} = \mathbf{M} \cdot_{\triangleright} \mathbf{N}^{\mathbf{H}} \cdot_{\triangleright} (\mathbf{N} \cdot_{\triangleright} \mathbf{N}^{\mathbf{H}})^{-1}$ . (Note that in this section, quaternion matrix inverses are always defined with respect to the direct quaternion matrix product and can be computed through inversion of the corresponding complex adjoint, see [38, Theorem 4.2].) The second subproblem is a constrained optimization problem since  $\mathbf{X}$  is constrained to belong to the set of real-valued matrices. The optimal solution can be found in [11, Appendix B] and is given by  $\hat{\mathbf{X}} = \text{Re}\{\mathbf{M} \cdot_{\triangleright} \mathbf{N}^{\mathbf{H}}\} \text{Re}\{(\mathbf{N} \cdot_{\triangleright} \mathbf{N}^{\mathbf{H}})\}^{-1}$ . Eventually, in order to solve (5.7), we resort to the properties of direct and reverse quaternion matrix products:  $\|\mathbf{M} - \mathbf{X} \cdot_{\triangleleft} \mathbf{N}\|_F^2 = \|\mathbf{M}^{\top} - \mathbf{N}^{\top} \cdot_{\triangleright} \mathbf{X}^{\top}\|_F^2$  so that the solution is given by:  $\hat{\mathbf{X}}_{\triangleleft} = \mathbf{M} \cdot_{\triangleleft} \mathbf{N}^{\mathbf{H}} \cdot_{\triangleleft} \left[ (\overline{\mathbf{N}} \cdot_{\triangleright} \mathbf{N}^{\top})^{-1} \right]^{\top}$ .

As a consequence, the updates computed during one iteration of the Q-ALS algorithm are given by:

$$(5.8) \quad \mathbf{A} \leftarrow \mathbf{T}_{(1)} \cdot_{\triangleright} (\overline{\mathbf{C} \odot \mathbf{B}}) \cdot_{\triangleright} ((\mathbf{C} \odot \mathbf{B})^{\top} \cdot_{\triangleright} (\overline{\mathbf{C} \odot \mathbf{B}}))^{-1}$$

$$(5.9) \quad \mathbf{B} \leftarrow \text{Re} \left\{ \mathbf{T}_{(2)} \cdot_{\triangleright} (\overline{\mathbf{C} \odot_{\triangleleft} \mathbf{A}}) \right\} \text{Re} \left\{ (\mathbf{C} \odot_{\triangleleft} \mathbf{A})^{\top} \cdot_{\triangleright} (\overline{\mathbf{C} \odot_{\triangleleft} \mathbf{A}}) \right\}^{-1}$$

$$(5.10) \quad \mathbf{C} \leftarrow \mathbf{T}_{(3)} \cdot_{\triangleleft} (\overline{\mathbf{B} \odot \mathbf{A}}) \cdot_{\triangleleft} \left( (\mathbf{B} \odot \mathbf{A})^{\mathbf{H}} \cdot_{\triangleright} (\mathbf{B} \odot \mathbf{A})^{-1} \right)^{\top}.$$

The Q-ALS updates resemble those of the usual ALS updates for computing the real and complex CPD. Therefore we anticipate that existing convergence results for ALS [32] can be adapted to the quaternion setting, exploiting recent advances in quaternion optimization [12, 35, 36]. However, a precise characterization of such properties requires a dedicated study, which is left for future work.

**5.2. Computing the Q-CPD in  $\mathbb{C}$ : the C-ALS algorithm.** This second approach relies on the complex coupled CONFAC representation of the Q-CPD presented in Subsection 4.2.2. First, start from the Cayley-Dickson decomposition of  $\mathcal{T} = \mathcal{T}_1 + \mathcal{T}_2 \mathbf{j}$  and recall that, if  $\mathcal{T}$  admits a Q-CPD of rank- $F$ , then from Theorem 4.6 one has  $\mathcal{T}_1 = \llbracket \mathbf{U} \Psi, \mathbf{B} \Phi, \overline{\mathbf{W}} \rrbracket$  and  $\mathcal{T}_2 = \llbracket \mathbf{U}, \mathbf{B} \Phi, \mathbf{W} \rrbracket$ , where  $\mathbf{U} = [\mathbf{A}_1 \ \mathbf{A}_2]$  and  $\mathbf{W} = [\mathbf{C}_2 \ \overline{\mathbf{C}}_1]$  are defined from Cayley-Dickson parts of  $\mathbf{A}$  and  $\mathbf{C}$ , respectively,

and where the constraint matrices  $\Psi$  and  $\Phi$  are given in (4.13). Then one can rewrite the optimization problem (5.1) as

$$(5.11) \quad (\hat{\mathbf{U}}, \hat{\mathbf{B}}, \hat{\mathbf{W}}) = \arg \min_{\mathbf{U}, \mathbf{B}, \mathbf{W}} \|\mathcal{T}_1 - \llbracket \mathbf{U} \Psi, \mathbf{B} \Phi, \overline{\mathbf{W}} \rrbracket\|_F^2 + \|\mathcal{T}_2 - \llbracket \mathbf{U}, \mathbf{B} \Phi, \mathbf{W} \rrbracket\|_F^2.$$

since the Frobenius norm is preserved by the Cayley-Dickson decomposition. The structure of the cost function in (5.11) enables the derivation of ALS updates for  $\mathbf{U}$ ,  $\mathbf{B}$  and  $\mathbf{W}$ , leading to the C-ALS algorithm detailed below.

For simplicity, let us start with the estimation of the real matrix  $\mathbf{B}$  given complex matrices  $\mathbf{U}$  and  $\mathbf{W}$ . Denote by  $\mathbf{T}_1^B$  and  $\mathbf{T}_2^B$  the unfoldings of  $\mathcal{T}_1$  and  $\mathcal{T}_2$  along the second mode, respectively and define  $\mathbf{T}^B = [\mathbf{T}_2^B \quad \overline{\mathbf{T}_1^B}]$ . Then, exploiting unfoldings (4.17) and (4.18) of the coupled CONFAC model along the second mode, the  $\mathbf{B}$ -update optimization problem can be written as

$$(5.12) \quad \mathbf{B} \leftarrow \arg \min_{\mathbf{B}} \|\mathbf{T}^B - \mathbf{B} \Phi(\chi_{\triangleright}(\mathbf{A}) \odot \mathbf{W})^\top\|_F^2 \quad \text{s.t. } \mathbf{B} \text{ has real-valued entries,}$$

which arises from the definition of  $\mathbf{U}$  as a horizontal stack of Cayley-Dickson parts of  $\mathbf{A}$  and simple manipulations of the unfoldings (4.17) and (4.18). The explicit solution to (5.12) is given below in (5.16).

A similar approach is used to get the  $\mathbf{U}$  and  $\mathbf{W}$ -updates. For  $\mathbf{W}$ , define a matrix  $\mathbf{T}^C = [\mathbf{T}_2^C \quad \overline{\mathbf{T}_1^C}]$  where  $\mathbf{T}_1^C$  and  $\mathbf{T}_2^C$  are the unfoldings along the third mode of  $\mathcal{T}_1$  and  $\mathcal{T}_2$ , respectively. From unfoldings of the CONFAC model, the  $\mathbf{W}$ -update is found as the solution to the optimization problem

$$(5.13) \quad \mathbf{W} \leftarrow \arg \min_{\mathbf{W}} \|\mathbf{T}^C - \mathbf{W}(\chi_{\triangleright}(\mathbf{A}) \odot \mathbf{B} \Phi)^\top\|_F^2.$$

For  $\mathbf{U}$ , similar but slightly more tedious manipulations yield to

$$(5.14) \quad \mathbf{U} \leftarrow \arg \min_{\mathbf{U}} \|\mathbf{T}^A - \mathbf{U}(\chi_{\triangleleft}(\mathbf{C}) \odot \mathbf{B} \Phi)^\top\|_F^2,$$

where this time  $\mathbf{T}^A = [\mathbf{T}_1^A \quad \mathbf{T}_2^A]$  gathers the unfoldings of  $\mathcal{T}_1$  and  $\mathcal{T}_2$  along the first mode. Eventually, the C-ALS procedure solves the Q-CPD approximation problem of rank- $F$  (5.1) thanks to the coupled CONFAC interpretation (5.11) by iterating the following explicit updates

$$(5.15) \quad \mathbf{U} \leftarrow \mathbf{T}^A \overline{(\chi_{\triangleleft}(\mathbf{C}) \odot \mathbf{B} \Phi)} \left( (\chi_{\triangleleft}(\mathbf{C}) \odot \mathbf{B} \Phi)^\top \overline{(\chi_{\triangleleft}(\mathbf{C}) \odot \mathbf{B} \Phi)} \right)^{-1},$$

$$(5.16)$$

$$\mathbf{B} \leftarrow \text{Re} \left\{ \mathbf{T}^B \overline{(\chi_{\triangleright}(\mathbf{A}) \odot \mathbf{W}) \Phi^\top} \right\} \text{Re} \left\{ \Phi \left( (\chi_{\triangleright}(\mathbf{A}) \odot \mathbf{W})^\top \overline{(\chi_{\triangleright}(\mathbf{A}) \odot \mathbf{W})} \right) \Phi^\top \right\}^{-1},$$

$$(5.17) \quad \mathbf{W} \leftarrow \mathbf{T}^C \overline{(\chi_{\triangleright}(\mathbf{A}) \odot \mathbf{B} \Phi)} \left( (\chi_{\triangleright}(\mathbf{A}) \odot \mathbf{B} \Phi)^\top \overline{(\chi_{\triangleright}(\mathbf{A}) \odot \mathbf{B} \Phi)} \right)^{-1}.$$

At convergence, matrices  $\mathbf{A}$  and  $\mathbf{C}$  are directly deduced from  $\mathbf{U}$  and  $\mathbf{W}$ .

It is worth mentioning that, by definition of matrices  $\mathbf{U}$  and  $\mathbf{W}$  on the one hand, and by construction of the C-ALS cost function (5.11) on the other hand, the least squares problems in (5.12), (5.13) and (5.14) are equivalent to those in (5.3), (5.4) and (5.2), respectively. As a result, Q-ALS iterations (5.8)–(5.10) and C-ALS iterations (5.15)–(5.17) are one-to-one as well. In other words, the Q-ALS and C-ALS algorithms provide equivalent algorithms to solve the Q-CPD approximation problem (5.1) of

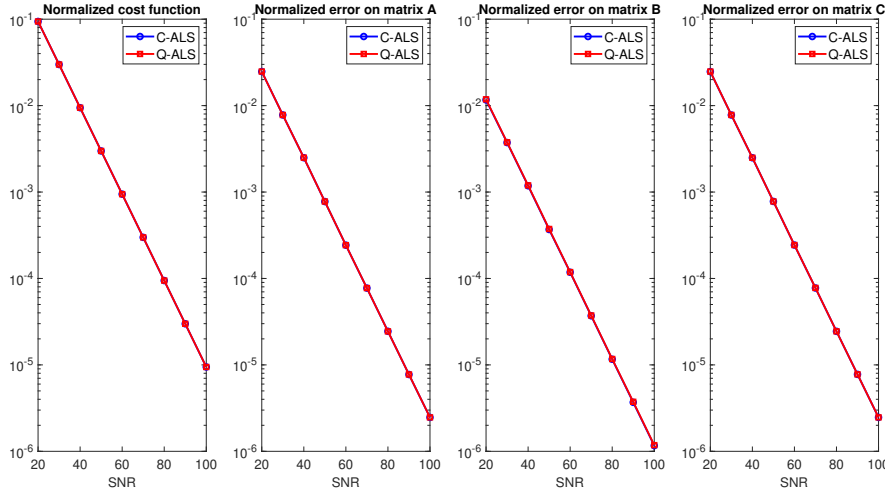


FIG. 6.1. Median value of the cost function and median NMSE on the factor matrices with respect to the SNR.

rank- $F$ , operating in the quaternion and complex domain, respectively. Q-ALS updates enjoy somewhat simpler explicit expressions by leveraging directly the Q-CPD model; however practical implementation requires the use of dedicated quaternion linear algebra libraries. In contrary, the C-ALS algorithm can directly take advantage of standard (complex) linear algebra libraries. These subtle implementation differences are exemplified in the next section.

**6. Numerical simulations.** The purpose of the following numerical simulations is twofold. We thus want to : *i*) verify that the Q-CPD can be computed effectively ; *ii*) assess the performances of the proposed algorithms in terms of convergence speed and accuracy in a simple but classical scenario. To this purpose, we consider here the Q-CPD decomposition of 500 third order tensors built according to the Q-CPD model (4.1), with  $F = 5$  and  $N_1 = N_2 = N_3 = 10$ . The entries of the three factor matrices were drawn randomly according to a normal distribution and are different for each tensor. Before computing the Q-CPD, an additive white Gaussian noise was applied to each tensor. The noise powers were set in order to obtain the desired Signal to Noise Ratio (SNR). More complex scenarios involving correlated factors, higher ranks, higher order tensors and further comparison between the algorithms deserve a specific study and are not in the scope of this paper. All simulations were performed using MATLAB environment. The Q-ALS algorithm has been programmed using the QTfM toolbox<sup>4</sup>. Both Q-ALS and C-ALS algorithms were initialized with the same (random) values and were stopped when the relative deviation between two consecutive values of the cost function fell under  $10^{-8}$  with a maximum of 500 iterations. At the end of each run, after removing the scaling and permutation ambiguities, the relative root mean-squared errors between the actual and estimated factor matrices were computed for both algorithms. Eventually, the median values of these errors and of the relative cost function were calculated over the 500 tensors realizations. This scenario was repeated for different SNR. Results are plotted on Figure 6.1.

Three main conclusions can be drawn. First, as expected, the curves corresponding to the two algorithms overlap perfectly. Second, the errors decrease linearly with the

<sup>4</sup><https://nl.mathworks.com/matlabcentral/fileexchange/76158-quaternion-toolbox-for-matlab>

noise power. This behavior is similar to the standard CPD algorithms used for real or complex tensors. Third, the error values are small enough to consider that Q-CPD decomposition can actually be computed, which validates the possibility of using this new tensor model in practical applications. In order to illustrate the practical interest of C-ALS, we also computed for each algorithm the average running time, running time per iteration and number of iterations. The average running times for C-ALS and Q-ALS were about 0.078s and 0.134s respectively. As both algorithms compute the same updates at each iteration (and therefore reach the same number of iterations), this difference is entirely explained by the average running time per iteration: about 0.0016s and 0.0027s for C-ALS and Q-ALS respectively and by the QTFM toolbox functions, which are not as optimized as native MATLAB functions. These observations hold for all the considered SNR. Eventually the same experiments were run for  $F = 10$ , leading to similar conclusions (plots not shown).

**7. Conclusion.** In this paper, we rigorously extended the notion of tensors to quaternion algebra. The key insight is that a quaternion tensor can be understood as a peculiar type of multilinear form, an  $\mathbb{H}\mathbb{R}$ -multilinear form, which preserves the fundamental multilinearity properties expected of such mathematical objects. Building on this foundation, we extended the Tucker and Canonical Polyadic Decompositions (CPD) to quaternion tensors. In particular, we showed the equivalence of the quaternion CPD (Q-CPD) to a specific complex-valued rank-(2, 2, 1) decomposition and established sufficient conditions for its uniqueness. These conditions rely on a careful handling of left and right linear independence in the quaternion domain, through the notion of left and right ranks, and left and right Kruskal ranks which are defined herein. Additionally, we derived two ALS-based algorithms for computing the Q-CPD, one operating in the quaternion domain and the other in the complex domain, demonstrating that the new decomposition can be computed efficiently. Future work will focus on further evaluation of the Q-CPD algorithms and on exploring their potential applications in data analysis and related fields.

**Acknowledgments.** The authors would like to thank the anonymous reviewer for their valuable comments and constructive suggestions, which have greatly improved the quality and clarity of this manuscript.

**Appendix A. Proof of Proposition 4.12.** Let us first define  $w(\mathbf{u})$  as the number of non zero entries of the complex vector  $\mathbf{u}$  and recall the Kruskal permutation lemma [20].

LEMMA A.1 (Kruskal permutation lemma). *Let  $\mathbf{B}$  and  $\tilde{\mathbf{B}}$  be two (real or complex) matrices of the same size and assume that  $k_{\mathbf{B}} \geq 2$  (no colinear columns). Let  $F$  denote the number of columns. If for any complex vector  $\mathbf{x}$  such that  $w(\tilde{\mathbf{B}}^H \mathbf{x}) \leq F - \text{rank } \tilde{\mathbf{B}} + 1$  we have  $w(\mathbf{B}^H \mathbf{x}) \leq w(\tilde{\mathbf{B}}^H \mathbf{x})$  then there exists a non singular diagonal matrix  $\mathbf{D}$  and a permutation matrix  $\mathbf{P}$  so that  $\tilde{\mathbf{B}} = \mathbf{BPD}$ .*

We can now start the proof of proposition 4.12.

*Proof.* Suppose that there exists an alternative solution of the CONFAC decomposition of  $\mathcal{T}_1$  given by  $(\tilde{\mathbf{U}}, \tilde{\mathbf{B}}, \tilde{\mathbf{W}})$ . Then, we have  $(\mathbf{T}_1^B)^T = (\tilde{\mathbf{U}}\Psi \odot \tilde{\mathbf{W}})\Phi^T \tilde{\mathbf{B}}^T = (\mathbf{U}\Psi \odot \bar{\mathbf{W}})\Phi^T \mathbf{B}^T$ . From the full column rank assumptions, we immediately verify that  $\text{rank}(\mathbf{U}\Psi \odot \bar{\mathbf{W}})\Phi^T \mathbf{B}^T = \text{rank } \mathbf{B} = F = \text{rank}(\tilde{\mathbf{U}}\Psi \odot \tilde{\mathbf{W}})\Phi^T \tilde{\mathbf{B}}^T$ . Therefore  $\text{rank } \tilde{\mathbf{B}} \geq F$ . Since  $\text{rank } \tilde{\mathbf{B}} \leq F$  by construction,  $\text{rank } \tilde{\mathbf{B}} = F$ . As a consequence, the condition in the lemma becomes simply: for any  $\mathbf{x}$  such that  $w(\tilde{\mathbf{B}}^T \mathbf{x}) \leq 1$  we have

$$w(\mathbf{B}^\top \mathbf{x}) \leq w(\tilde{\mathbf{B}}^\top \mathbf{x}).$$

If  $w(\tilde{\mathbf{B}}^\top \mathbf{x}) = 0$  then  $(\mathbf{U}\Psi \odot \overline{\mathbf{W}})\Phi^\top \mathbf{B}^\top \mathbf{x} = \mathbf{0}$ .  $\mathbf{U}\Psi$  and  $\mathbf{W}$  being full column rank,  $\mathbf{U}\Psi \odot \overline{\mathbf{W}}$  is also full column rank, thus we necessarily have  $\mathbf{B}^\top \mathbf{x} = \mathbf{0}$  so that  $w(\mathbf{B}^\top \mathbf{x}) = 0 = w(\tilde{\mathbf{B}}^\top \mathbf{x})$ . The condition in the lemma is verified.

Let us now suppose that  $w(\tilde{\mathbf{B}}^\top \mathbf{x}) = 1$ , there exists  $i$ ,  $1 \leq i \leq F$ , so that  $\tilde{\mathbf{B}}^\top \mathbf{x} = \mathbf{e}_i$  where  $\mathbf{e}_i$  is the  $i^{\text{th}}$  unit vector of size  $F$ . Denoting  $\mathbf{z} = \mathbf{B}^\top \mathbf{x}$  and  $\mathbf{M}^+$  the Moore-Penrose pseudo-inverse of  $\mathbf{M}$ , we have:

$$(A.1) \quad (\tilde{\mathbf{U}}\Psi \odot \overline{\tilde{\mathbf{W}}})\Phi^\top \mathbf{e}_i = (\mathbf{U}\Psi \odot \overline{\mathbf{W}})\Phi^\top \mathbf{z}$$

$$(A.2) \quad (\tilde{\mathbf{U}}\Psi \boxtimes \overline{\tilde{\mathbf{W}}})(\mathbf{I}_{2F} \odot \mathbf{I}_{2F})\Phi^\top \mathbf{e}_i = (\mathbf{U}\Psi \boxtimes \overline{\mathbf{W}})(\mathbf{I}_{2F} \odot \mathbf{I}_{2F})\Phi^\top \mathbf{z}$$

$$(A.3) \quad (\mathbf{U}\Psi \boxtimes \overline{\mathbf{W}})^+(\tilde{\mathbf{U}}\Psi \boxtimes \overline{\tilde{\mathbf{W}}})(\mathbf{I}_{2F} \odot \mathbf{I}_{2F})\Phi^\top \mathbf{e}_i = (\mathbf{I}_{2F} \odot \mathbf{I}_{2F})\Phi^\top \mathbf{z}$$

$$(A.4) \quad (\Psi^+ \mathbf{U}^+ \tilde{\mathbf{U}}\Psi \boxtimes \overline{\tilde{\mathbf{W}}^+ \tilde{\mathbf{W}}})(\mathbf{I}_{2F} \odot \mathbf{I}_{2F})\Phi^\top \mathbf{e}_i = (\mathbf{I}_{2F} \odot \mathbf{I}_{2F})\Phi^\top \mathbf{z}.$$

We then define the two non-singular matrices of size  $2F \times 2F$ :  $\mathbf{R} = \Psi^+ \mathbf{U}^+ \tilde{\mathbf{U}}\Psi$  and  $\mathbf{S} = \overline{\tilde{\mathbf{W}}^+ \tilde{\mathbf{W}}}$ . It yields

$$(A.5) \quad (\mathbf{R} \odot \mathbf{S})\Phi^\top \mathbf{e}_i = (\mathbf{I}_{2F} \odot \mathbf{I}_{2F})\Phi^\top \mathbf{z}.$$

On the one hand

$$(A.6) \quad (\mathbf{R} \odot \mathbf{S})\Phi^\top \mathbf{e}_i = \mathbf{r}_i \boxtimes \mathbf{s}_i + \mathbf{r}_{F+i} \boxtimes \mathbf{s}_{N+i}$$

$$(A.7) \quad = \text{vec}\{\mathbf{s}_i \mathbf{r}_i^\top + \mathbf{s}_{N+i} \mathbf{r}_{N+i}^\top\}$$

where  $\mathbf{r}_i$  and  $\mathbf{s}_i$  denote the  $i^{\text{th}}$  columns of  $\mathbf{R}$  and  $\mathbf{S}$  respectively.

On the other hand

$$(A.8) \quad (\mathbf{I}_{2F} \odot \mathbf{I}_{2F})\Phi^\top \mathbf{z} = \sum_{f=1}^F z_f \left( \begin{pmatrix} \mathbf{e}_f \\ \mathbf{0} \end{pmatrix} \boxtimes \begin{pmatrix} \mathbf{e}_f \\ \mathbf{0} \end{pmatrix} + \begin{pmatrix} \mathbf{0} \\ \mathbf{e}_f \end{pmatrix} \boxtimes \begin{pmatrix} \mathbf{0} \\ \mathbf{e}_f \end{pmatrix} \right)$$

$$(A.9) \quad = \text{vec}\left\{ \begin{pmatrix} \text{Diag}\{\mathbf{z}\} & \mathbf{0} \\ \mathbf{0} & \text{Diag}\{\mathbf{z}\} \end{pmatrix} \right\}.$$

So that equation (A.5) rewrites

$$(A.10) \quad \mathbf{s}_i \mathbf{r}_i^\top + \mathbf{s}_{F+i} \mathbf{r}_{F+i}^\top = \begin{pmatrix} \text{Diag}\{\mathbf{z}\} & \mathbf{0} \\ \mathbf{0} & \text{Diag}\{\mathbf{z}\} \end{pmatrix}.$$

Eventually, we just need to consider the two upper blocks. So that defining

$$(A.11) \quad \mathbf{r}_i^0 = (R_{1,i}, \dots, R_{F,i})^\top$$

$$(A.12) \quad \mathbf{r}_{F+i}^0 = (R_{1,F+i}, \dots, R_{F,F+i})^\top$$

$$(A.13) \quad \mathbf{r}_i^1 = (R_{F+1,i}, \dots, R_{2F,i})^\top$$

$$(A.14) \quad \mathbf{r}_{F+i}^1 = (R_{F+1,F+i}, \dots, R_{2F,F+i})^\top$$

$$(A.15) \quad \mathbf{s}_i^0 = (S_{1,i}, \dots, S_{F,i})^\top$$

$$(A.16) \quad \mathbf{s}_{F+i}^0 = (S_{1,F+i}, \dots, S_{F,F+i})^\top$$

yields

$$(A.17) \quad \mathbf{s}_i^0 (\mathbf{r}_i^1)^\top + \mathbf{s}_{F+i}^0 (\mathbf{r}_{F+i}^1)^\top = \mathbf{0}$$

$$(A.18) \quad R_{f,i} \mathbf{s}_i^0 + R_{f,F+i} \mathbf{s}_{F+i}^0 = z_f \mathbf{e}_f, \quad \forall f = 1, \dots, F.$$

Thus there exists  $\alpha \in \mathbb{C}$  such that

$$(A.19) \quad (R_{f,i} + \alpha)\mathbf{s}_i^0 = z_f \mathbf{e}_f, \quad \forall f = 1, \dots, F.$$

We immediately deduce that there is at most one  $f$  verifying  $z_f \neq 0$ . Thereby,  $w(\mathbf{B}^\top \mathbf{x}) \leq 1$ . Hence  $w(\mathbf{B}^\top \mathbf{x}) \leq w(\tilde{\mathbf{B}}^\top \mathbf{x})$  and, again, the condition in the permutation lemma is verified. We can thus apply it and conclude the proof.  $\square$

#### REFERENCES

- [1] Q. BARTHELEMY, A. LARUE, AND J. I. MARS, *Color Sparse Representations for Image Processing: Review, Models, and Prospects*, IEEE Transactions on Image Processing, 24 (2015), pp. 3978–3989, <https://doi.org/10.1109/TIP.2015.2458175>.
- [2] R. BRO, R. A. HARSHMAN, N. D. SIDIROPOULOS, AND M. E. LUNDY, *Modeling multi-way data with linearly dependent loadings*, Journal of Chemometrics, 23 (2009), pp. 324–340, <https://doi.org/https://doi.org/10.1002/cem.1206>.
- [3] H. CHEN, W. WANG, AND W. LIU, *Augmented Quaternion ESPRIT-Type DOA Estimation With a Crossed-Dipole Array*, IEEE Communications Letters, 24 (2020), pp. 548–552, <https://doi.org/10.1109/LCOMM.2019.2962463>.
- [4] Y. CHEN, X. XIAO, AND Y. ZHOU, *Low-Rank Quaternion Approximation for Color Image Processing*, IEEE Transactions on Image Processing, 29 (2020), pp. 1426–1439, <https://doi.org/10.1109/TIP.2019.2941319>.
- [5] P. COMON, *Tensors: a brief introduction*, IEEE Signal Processing Magazine, 31 (2014), pp. 44–53.
- [6] A. L. F. DE ALMEIDA, G. FAVIER, AND J. C. M. MOTA, *A constrained factor decomposition with application to mimo antenna systems*, IEEE Transactions on Signal Processing, 56 (2008), pp. 2429–2442, <https://doi.org/10.1109/TSP.2008.917026>.
- [7] A. L. F. DE ALMEIDA, X. LUCIANI, A. STEGEMAN, AND P. COMON, *Confac decomposition approach to blind identification of underdetermined mixtures based on generating function derivatives*, IEEE Transactions on Signal Processing, 60 (2012), pp. 5698–5713, <https://doi.org/10.1109/TSP.2012.2208956>.
- [8] L. DE LATHAUWER, *Decompositions of a Higher-Order Tensor in Block Terms—Part I: Lemmas for Partitioned Matrices*, SIAM Journal on Matrix Analysis and Applications, 30 (2008), pp. 1022–1032, <https://doi.org/10.1137/060661685>.
- [9] L. DE LATHAUWER, *Decompositions of a Higher-Order Tensor in Block Terms—Part II: Definitions and Uniqueness*, SIAM Journal on Matrix Analysis and Applications, 30 (2008), pp. 1033–1066, <https://doi.org/10.1137/070690729>.
- [10] G. FAVIER AND A. L. DE ALMEIDA, *Overview of constrained parafac models*, EURASIP Journal on Advances in Signal Processing, 2014 (2014), pp. 1–25.
- [11] J. FLAMANT, S. MIRON, AND D. BRIE, *Quaternion Non-Negative Matrix Factorization: Definition, Uniqueness, and Algorithm*, IEEE Transactions on Signal Processing, 68 (2020), pp. 1870–1883, <https://doi.org/10.1109/TSP.2020.2974651>.
- [12] J. FLAMANT, S. MIRON, AND D. BRIE, *A general framework for constrained convex quaternion optimization*, IEEE Transactions on Signal Processing, 70 (2021), pp. 254–267.
- [13] K. GILMAN, D. A. TARZANAGH, AND L. BALZANO, *Grassmannian optimization for online tensor completion and tracking with the t-svd*, IEEE Transactions on Signal Processing, 70 (2022), pp. 2152–2167.
- [14] M. HOBIGER, N. LE BIHAN, C. CORNOU, AND P. BARD, *Multicomponent Signal Processing for Rayleigh Wave Ellipticity Estimation: Application to Seismic Hazard Assessment*, IEEE Signal Processing Magazine, 29 (2012), pp. 29–39, <https://doi.org/10.1109/MSP.2012.2184969>.
- [15] O. IMHOGIEMHE, J. FLAMANT, X. LUCIANI, Y. ZNIYED, AND S. MIRON, *Low-rank tensor decompositions for quaternion multiway arrays*, in ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2023, pp. 1–5.
- [16] J. E. JAMISON, *Extension of some theorems of complex functional analysis to linear spaces over the quaternions and Cayley numbers*, PhD thesis, University of Missouri–Rolla, 1970.
- [17] M. E. KILMER, K. BRAMAN, N. HAO, AND R. C. HOOVER, *Third-order tensors as operators on matrices: A theoretical and computational framework with applications in imaging*, SIAM Journal on Matrix Analysis and Applications, 34 (2013), pp. 148–172.
- [18] M. E. KILMER AND C. D. MARTIN, *Factorization strategies for third-order tensors*, Linear Algebra and its Applications, 435 (2011), pp. 641–658.

- [19] T. G. KOLDA AND B. W. BADER, *Tensor decompositions and applications*, SIAM review, 51 (2009), pp. 455–500.
- [20] J. B. KRUSKAL, *Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics*, Linear Algebra and its Applications, 18 (1977), pp. 95–138, [https://doi.org/10.1016/0024-3795\(77\)90069-6](https://doi.org/10.1016/0024-3795(77)90069-6).
- [21] L.-H. LIM, *Tensors in computations*, Acta Numerica, 30 (2021), pp. 555–764.
- [22] J. MIAO AND K. I. KOU, *Quaternion tensor singular value decomposition using a flexible transform-based approach*, Signal Processing, 206 (2023), p. 108910.
- [23] J. MIAO, K. I. KOU, H. CAI, AND L. LIU, *Quaternion tensor left ring decomposition and application for color image inpainting*, Journal of Scientific Computing, 101 (2024), p. 1.
- [24] S. MIRON, J. FLAMANT, N. LE BIHAN, P. CHAINAIS, AND D. BRIE, *Quaternions in Signal and Image Processing: A comprehensive and objective overview*, IEEE Signal Processing Magazine, 40 (2023), pp. 26–40, <https://doi.org/10.1109/MSP.2023.3278071>.
- [25] J. PAN AND M. K. NG, *Separable quaternion matrix factorization for polarization images*, SIAM Journal on Imaging Sciences, 16 (2023), pp. 1281–1307.
- [26] J. PAN AND M. K. NG, *Block-diagonalization of quaternion circulant matrices with applications*, SIAM Journal on Matrix Analysis and Applications, 45 (2024), pp. 1429–1454.
- [27] Z. QIN, Z. MING, AND L. ZHANG, *Singular value decomposition of third order quaternion tensors*, Applied Mathematics Letters, 123 (2022), p. 107597.
- [28] L. RODMAN, *Topics in quaternion linear algebra*, vol. 45, Princeton University Press, 2014.
- [29] D. SCHULZ AND R. S. THOMÄ, *Using Quaternion-Valued Linear Algebra*, arXiv:1311.7488 [math], (2014). arXiv: 1311.7488.
- [30] K. SHOEMAKE, *Animating rotation with quaternion curves*, in Proceedings of the 12th annual conference on Computer graphics and interactive techniques, 1985, pp. 245–254.
- [31] C. TOOK AND D. MANDIC, *The Quaternion LMS Algorithm for Adaptive Filtering of Hypercomplex Processes*, IEEE Transactions on Signal Processing, 57 (2009), pp. 1316–1327, <https://doi.org/10.1109/TSP.2008.2010600>.
- [32] A. USCHMAJEV, *Local convergence of the alternating least squares algorithm for canonical tensor approximation*, SIAM Journal on Matrix Analysis and Applications, 33 (2012), pp. 639–652.
- [33] N. A. VALOUS, E. HITZER, S. VITABILE, S. BERNSTEIN, C. LAVOR, D. ABBOTT, M. E. LUNAE-LIZARRARÁS, AND W. LOPES, *Hypercomplex signal and image processing: Part 1 [from the guest editors]*, IEEE Signal Processing Magazine, 41 (2024), pp. 11–13, <https://doi.org/10.1109/MSP.2024.3378129>.
- [34] N. A. VALOUS, E. HITZER, S. VITABILE, S. BERNSTEIN, C. LAVOR, D. ABBOTT, M. E. LUNAE-LIZARRARÁS, AND W. LOPES, *Hypercomplex signal and image processing: Part 2 [from the guest editors]*, IEEE Signal Processing Magazine, 41 (2024), pp. 18–20, <https://doi.org/10.1109/MSP.2024.3415248>.
- [35] D. XU, C. JAHANCHAH, C. C. TOOK, AND D. P. MANDIC, *Enabling quaternion derivatives: the generalized HR calculus*, Royal Society Open Science, 2 (2015), p. 150255, <https://doi.org/10.1098/rsos.150255>.
- [36] D. XU AND D. P. MANDIC, *The Theory of Quaternion Matrix Derivatives*, IEEE Transactions on Signal Processing, 63 (2015), p. 14.
- [37] L. YANG, K. I. KOU, J. MIAO, Y. LIU, AND P. M. HOI, *Quaternion tensor completion with sparseness for color video recovery*, Applied Soft Computing, 154 (2024), p. 111322.
- [38] F. ZHANG, *Quaternions and matrices of quaternions*, Linear Algebra and its Applications, 251 (1997), pp. 21–57, [https://doi.org/10.1016/0024-3795\(95\)00543-9](https://doi.org/10.1016/0024-3795(95)00543-9).
- [39] C. ZOU, K. I. KOU, AND Y. WANG, *Quaternion Collaborative and Sparse Representation With Application to Color Face Recognition*, IEEE Transactions on Image Processing, 25 (2016), pp. 3287–3302, <https://doi.org/10.1109/TIP.2016.2567077>.
- [40] C. ZOU, K. I. KOU, Y. WANG, AND Y. Y. TANG, *Quaternion block sparse representation for signal recovery and classification*, Signal Processing, 179 (2021), p. 107849, <https://doi.org/10.1016/j.sigpro.2020.107849>.