

Quaternions in Signal and Image Processing

Sebastian Miron, Julien Flamant, Nicolas Le Bihan, Pierre Chainais, David Brie

Abstract

Quaternions are still largely misunderstood and often considered as an “exotic” signal representation without much practical utility, despite the fact that they have been around the signal and image processing community for more than thirty years now. The main aim of this article is to counter this misconception and to demystify the use of quaternion algebra for solving problems in signal and image processing. To this end, we propose a comprehensive and objective overview of the key aspects of quaternion representations, models and methods and illustrate our journey through the literature with flagship applications. We conclude this work by an outlook on the remaining challenges and open problems in quaternion signal and image processing.

Index Terms

Quaternion algebra, hypercomplex signal processing, quaternion representations, quaternion singular value decomposition, $\mathbb{H}\mathbb{R}$ -calculus, properness of quaternion random variables

I. HISTORY, BACKGROUND AND AIM OF THE PAPER

Quaternions were first introduced by Irish mathematician Sir William Rowan Hamilton in 1843 as a result of his dedication to generalize complex numbers in more than two dimensions. He spent many years trying in vain to define a three-dimensional algebra based on a system of triplets. Story has it that he discovered the quaternions when walking along the Royal Canal in Dublin, Ireland on the 16th of October 1843, and immediately carved the fundamental equation for quaternion algebra in the stone of the nearby Brougham Bridge:

$$i^2 = j^2 = k^2 = ijk = -1, \quad (1)$$

S. Miron, J. Flamant, and D. Brie are with the Université de Lorraine, CNRS, CRAN, F-54000 Nancy, France. N. Le Bihan is with the CNRS/GIPSA-Lab, F-38000 Grenoble, France. P. Chainais is with the Université de Lille, CNRS, Centrale Lille, UMR 9189 CRISTAL, F-59000 Lille, France. This work has been supported in part by the ANR through 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](https://creativecommons.org/licenses/by/4.0/).

where i, j and k define imaginary units. While the carving has now disappeared, a plaque honoring Hamilton's memory can be found at the same place today. Hamilton devoted his last 20 years to the study of his quaternions which culminated in his book, *Elements of quaternions*. After his death in 1865 quaternions remained fashionable for some time, but they were rapidly superseded by the advent of linear algebra as we know it today through the work of Gibbs and Heaviside at the end of the 19th century¹. Still, Hamilton was a precursor in many aspects and influenced many. For instance, he invented the term *vector* well before the advent of modern linear algebra: at the time, it simply referred to the three dimensional imaginary part of a quaternion.

The set of quaternions is usually denoted by \mathbb{H} as a tribute to Hamilton's discovery. Just like complex numbers are well known to describe algebraically the geometry of the 2D plane, quaternion algebra permits straightforward descriptions of geometric transformations in 3D and 4D spaces. Formally, a quaternion is defined by a real (or scalar) part and an imaginary (or vector) part made of three components along imaginary units i, j and k . This close relationship between purely imaginary (or simply pure) quaternions and vectors in \mathbb{R}^3 is fundamental. In fact, the triplet of imaginary units (i, j, k) can be identified with the canonical Cartesian basis of \mathbb{R}^3 given by $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_1 \times \mathbf{e}_2)$ where \times denotes the cross-product between vectors of \mathbb{R}^3 . Remarkably, quaternion algebra encodes the cross-product operation in a natural way since $ij = k$, $jk = i$ or $ki = j$. More generally, the product of two quaternions involves 3D scalar products and cross-products. This also explains why the multiplication of two quaternions is non-commutative: it results from the well-known non-commutativity of the cross-product, and translates the fact that geometric transformations in 3D and higher dimensions lack commutativity as well. For later reference, Table I collects essential definitions, sets, properties and polar forms related to quaternion algebra.

Perhaps one of the most striking examples of quaternion utilization in today's applications lies in their ability to represent 3D rotations. Representing a 3D rotation with a single unit quaternion has many benefits over standard Euler angles rotation matrices: a lower number of parameters, no *gimbal lock* singularities², nice interpolation properties between rotations. These advantages have been acknowledged for a long time in robotics [1] and computer graphics [2], where the use of quaternions is well-established. On the contrary, the use of quaternions in signal and image processing is still blooming, with first works dating back to the early 1990s [3].

¹To learn more about the fascinating history of quaternions and linear algebra, we recommend the reading of *A History of Vector Analysis: Evolution of the Idea of a Vectorial System*, by Michael J. Crowe, Dover Publications, 1994.

²This refers to the loss of one degree of freedom that can occur when using Euler angles to parameterize 3D rotations, causing important practical issues when representing a sequence of rotations.

TABLE I
HANDBOOK OF QUATERNION ALGEBRA

Basic definitions	Canonical basis	$\mathbb{H} = \text{span}\{1, \mathbf{i}, \mathbf{j}, \mathbf{k}\}$	
	Elementary relations	$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{i}\mathbf{j}\mathbf{k} = -1, \quad \mathbf{i}\mathbf{j} = -\mathbf{j}\mathbf{i} = \mathbf{k}$ $\mathbf{k}\mathbf{i} = -\mathbf{i}\mathbf{k} = \mathbf{j}, \quad \mathbf{j}\mathbf{k} = -\mathbf{k}\mathbf{j} = \mathbf{i}$	
	Cartesian representation	$q = a + \mathbf{i}b + \mathbf{j}c + \mathbf{k}d, \quad a, b, c, d \in \mathbb{R}$	
	Real and imaginary parts	$\text{Re}(q) = a, \quad \text{Im}_{\mathbf{i}}(q) = b,$ $\text{Im}_{\mathbf{j}}(q) = c, \quad \text{Im}_{\mathbf{k}}(q) = d$	
	Scalar and vector parts	$\mathcal{S}(q) = a, \quad \mathcal{V}(q) = \mathbf{i}b + \mathbf{j}c + \mathbf{k}d$	
	Conjugation	$\bar{q} = a - \mathbf{i}b - \mathbf{j}c - \mathbf{k}d = \mathcal{S}(q) - \mathcal{V}(q)$	
	Modulus	$ q = \sqrt{q\bar{q}} = \sqrt{\bar{q}q} = \sqrt{a^2 + b^2 + c^2 + d^2}$	
	Inverse	$q^{-1} = \frac{\bar{q}}{ q ^2}, \quad q \neq 0$	
	Involution	$q^\mu = -\mu q \mu, \quad \mu^2 = -1$	
	Sets	Pure quaternions	$\mathcal{V}(\mathbb{H}) = \{q \in \mathbb{H} \mid \text{Re}(q) = \mathcal{S}(q) = 0\}$
Unit quaternions		$\text{Sp}(1) = \{q \in \mathbb{H} \mid q = 1\}$	
Complex subfields of \mathbb{H} , $\mu \in \mathcal{V}(\mathbb{H}) \cap \text{Sp}(1)$		$\mathbb{C}_\mu = \{\alpha + \mu\beta \mid \alpha, \beta \in \mathbb{R}\}$	
Properties $p, q \in \mathbb{H}$	Addition	$\text{Re}(p + q) = \text{Re}(p) + \text{Re}(q)$ $\text{Im}_\mu(p + q) = \text{Im}_\mu(p) + \text{Im}_\mu(q), \quad \mu = \mathbf{i}, \mathbf{j}, \mathbf{k}$	
	Product	$pq = \mathcal{S}(p)\mathcal{S}(q) - \langle \mathcal{V}(p), \mathcal{V}(q) \rangle_{\mathbb{R}^3}$ $+ \mathcal{S}(p)\mathcal{V}(q) + \mathcal{S}(q)\mathcal{V}(p) + \mathcal{V}(p) \times_{\mathbb{R}^3} \mathcal{V}(q)$	
	Compatibility with operations	conjugation	$\overline{(pq)} = \bar{q}\bar{p}$
		involution	$(pq)^\mu = p^\mu q^\mu$
		inverse	$(pq)^{-1} = q^{-1}p^{-1}$
		modulus	$ pq = p q $
Polar forms and geometry	Euler formula (axis-angle representation)	$q = q e^{\mu_q \Phi_q} = q (\cos \phi_q + \mu_q \sin \phi_q)$ axis $\mu_q \in \mathcal{V}(\mathbb{H})$ with $ \mu_q = 1$, angle $\phi_q \in [0, 2\pi]$	
	Euler angle polar form (xzy convention)	$q = q e^{i\theta}e^{-\mathbf{k}\chi}e^{j\phi}$ $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}], \chi \in [-\frac{\pi}{4}, \frac{\pi}{4}], \phi \in [0, 2\pi]$	
	3D rotation by axis μ , angle $\alpha \in [0, \pi]$	$\mathcal{R}_{\mu, \alpha}(q) = \exp(\mu \frac{\alpha}{2})q \exp(-\mu \frac{\alpha}{2})$	

This article aims at providing an overview of the current use of quaternions in signal and image processing, ranging from data representation using quaternions to dedicated quaternion-domain methods and algorithms. It is intended to demystify the field for the newcomers, and make it accessible to the many. We hope to demonstrate that, up to the special care required to extend standard signal and image processing tools to quaternion algebra, the use of quaternion-domain approaches enables a compact, elegant and interpretable way to handle geometric properties of signals and images.

II. REPRESENTING SIGNALS AND IMAGES WITH QUATERNIONS

Many physical phenomena can be probed using (electronic) sensors. In a very broad viewpoint, sensors transform complex physical properties into electrical properties (such as output voltage or current) that can be processed by further electronics. For that reason, physical measurements always boil down to acquiring real values: intensity of light passing through a color filter or variations of amplitude along one direction in an accelerometer, for instance. Data recordings therefore correspond to arrays of real numbers, such as vectors (*e.g.* univariate signals) or matrices (*e.g.* grayscale images). However, even if raw data are intrinsically real-valued, one often takes advantage of other *representations* to facilitate their modeling, analysis or processing. One of the most striking examples is perhaps the use of complex numbers in signal and image processing. They arise naturally when *transforming* raw data using the (complex) Fourier transform. Such a manipulation enables many insights that would have been otherwise (almost) impossible. For instance, complex numbers define unambiguously the essential notions of magnitude and phase, which are pivotal to signal processing practice: spectral analysis, filter design, time-frequency analysis, array processing, etc. They also provide a compact and elegant way to write pairs of signals such as in-phase and quadrature components in communications or functional magnetic resonance imaging. These several convenient properties explain the popularity of complex-valued representations in signal and image processing.

Quaternions are no different in that respect. Just like complex numbers, they offer a novel representation space which exhibits several unique properties such as polar forms and natural handling of 3D geometry, which can be interesting to exploit in applications. More importantly, quaternions define a (skew)-field: this means that except noncommutativity of the quaternion product, quaternions have the same desirable properties as the real and complex fields. This ensures that the mathematical foundations crucial to signal processing (Fourier transform, vector spaces, linear algebra, etc.) can all be defined in a meaningful way. Moreover, the similarity between methodologies developed for quaternion-valued signal and image processing and their real counterparts tend to demonstrate that *noncommutativity* is not an issue in general – provided that it is handled in an adequate manner. Since their introduction in the signal processing

community more than three decades ago, the usage of quaternion-valued representations has focused on two complementary settings, namely the *encoding* of 3D and 4D signals and the construction of interpretable algebraic *embeddings* of signals and images.

A. Encoding 3D and 4D vector signals

This first setting may arguably be seen as the most natural one. The main idea is to encode the components of 3D or 4D vector signals on the three (imaginary only) or four (real and imaginary) parts of a quaternion. This allows to extend the standard arithmetic operations over real numbers (addition, subtraction, multiplication, division) to three and four dimensional real vectors. In the case of the two-dimensional real vectors, this extension is naturally performed by the complex numbers. This way, one can handle vector quantities using algebraic operations in a way similar to what can be done with scalars. This can be very helpful especially for the case when 3D or 4D vector data is acquired with respect to one or two diversities (time, space, wavelength, etc.). As an illustrative example, consider the case of a color image defined by the triplet of real matrices $\{\mathbf{R}, \mathbf{G}, \mathbf{B}\}$ encoding red, green and blue color channels respectively. This triplet can be conveniently represented as the pure quaternion matrix $\mathbf{Q} = i\mathbf{R} + j\mathbf{G} + k\mathbf{B}$. This algebraic representation follows directly from the identification of the imaginary units i, j, k with the canonical Cartesian basis of \mathbb{R}^3 . It permits to separate between the *internal multivariate* nature of the color image (*i.e.* a 3D vector encoding colors at each pixel) and its *external multidimensional* nature (an array of $M \times N$ spatial pixels) in an elegant way. In comparison, the equivalent real-domain representation of such data is often cumbersome, and usually handled by stacking the 3 or 4 components in a single long-vector or matrix. While this stacking procedure is mathematically sound, it may hinder the intimate relationships between the internal components and the geometric properties of such vector data. On the contrary, the algebraic quaternion encoding of 3D and 4D vectors enables natural representations of multi-diversity vector data as quaternion vectors and quaternion matrices. This also means that many fundamental signal processing operations for 3D and 4D vector data can be formulated in terms of quaternion linear algebra operations in a rather straightforward way. Building on these advantages, quaternions have been effectively employed to encode vector measurements in seismology [4], wind and temperature forecasting [5], electromagnetics [6], telecommunications [7], or color image processing [8], [9], to name a few.

B. Algebraic embeddings of signals and images

This second setting is a little bit more intricate. It relies on carefully designed transforms that map signals or images into the quaternion domain. These transforms define *quaternion embeddings* which

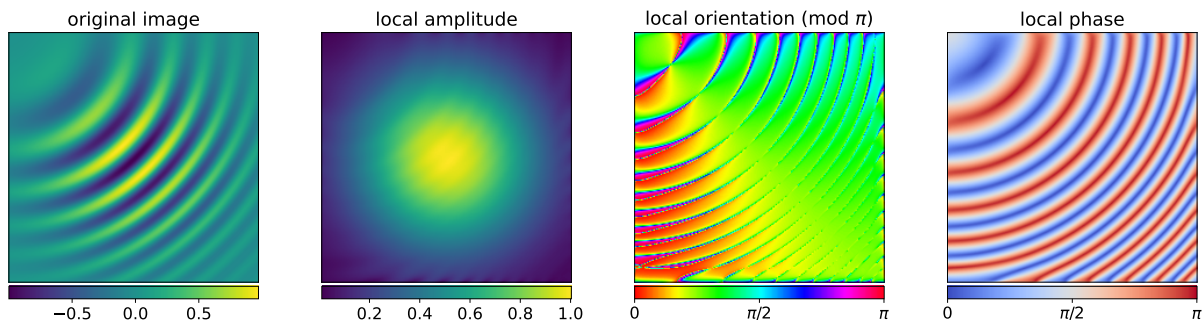


Fig. 1. Monogenic signal analysis of a AM-FM mode. The quaternion polar form enables identification of local amplitude (Gaussian kernel envelope); local orientation (shown as direction for visualization purposes) which gives the orientation of the tangent vector of contour lines; local phase which encodes image lines (0 or π values) and contours ($\pi/2$ values).

facilitate the analysis, understanding and processing of the original data. They ship with highly interpretable parameters, making it possible to decipher geometric features of the original signal or image. So far, most of the research towards interpretable quaternion-valued embeddings has focused on two areas: the construction of quaternion transforms for analyzing local features in grayscale images and the development of a geometric signal processing toolbox for bivariate signals. It is worth noting that albeit being apparently unrelated, both approaches consider generalizations of the analytic signal in higher-dimensions by exploiting quaternion algebra; they also both leverage extensively quaternion polar forms for meaningful interpretations of the embeddings. These two areas are reviewed in detail here below.

Quaternions transforms for grayscale image analysis

The importance of the analytic signal for the understanding and modeling of instantaneous amplitude and phase of real-valued signals has been recognized for a long time in signal processing. This motivated the study of its generalizations in higher dimensions, the most prominent one being the definition of a meaningful 2D-“analytic signal” to analyze the local content of images (grayscale). The main difficulty in directly extending definitions from the 1D case is that higher dimensions (notably 2D) lack a natural multidimensional Hilbert transform. To fill this gap, several approaches have been designed for the 2D case: the most salient ones exploit the higher degrees of freedom offered by the quaternion algebra to formulate meaningful 2D counterparts of the 1D analytic signal. A first approach, proposed by Bülow and Sommer [10], uses a carefully designed 2D Quaternion Fourier Transform (QFT) enjoying desirable symmetry properties for grayscale images. This permits to define a one-to-one mapping between a grayscale image and a quaternion-valued image obtained by restricting its QFT to the first orthant. Further decomposing this quaternion-valued signal using the Euler quaternion polar form (see Table I)

allows identification of a local amplitude and three phases, which are meaningful for texture analysis. This QFT-based approach was further explored in [11] with the design of a dual-tree quaternion wavelet transform for coherent multiscale analysis of grayscale images.

Another line of work, perhaps the most popular one, revolves around the *monogenic signal*. It was first introduced by Felsberg and Sommer [12] as a generalization of the analytic signal to the 2D case. The monogenic signal is a quaternion-valued image, built from the original grayscale image and two Riesz transforms. The interpretation as a 2D-“analytic signal” essentially comes from the intuition that, “the Riesz transform is to the Hilbert transform what the gradient is to the derivative operator”, to quote [13]. Given a grayscale image $f(\mathbf{r})$ with spatial coordinates $\mathbf{r} = (r_1, r_2)$, the Riesz transform $\mathcal{R}f = (\mathcal{R}_1 f, \mathcal{R}_2 f)$ is defined in the spatial domain as

$$\mathcal{R}_i f(\mathbf{r}) = \text{p.v.} \frac{1}{\pi} \iint_{\mathbb{R}^2} \frac{(r_i - r'_i)}{\|\mathbf{r} - \mathbf{r}'\|_2^3} f(\mathbf{r}') d\mathbf{r}', \quad i = 1, 2, \quad (2)$$

where p.v. stands for Cauchy principal value. The Riesz transform is translation- and scale-invariant. It also exhibits a nice compatibility with 2D rotations, a property known as *steerability*. The monogenic signal $\mathcal{M}f$ is constructed in the quaternion domain as $\mathcal{M}f = f + \mathbf{i}\mathcal{R}_1 f + \mathbf{j}\mathcal{R}_2 f$. Being quaternion-valued, it can be uniquely decomposed using the quaternion polar form $q = |q|e^{\boldsymbol{\mu}_q \phi_q}$, where the axis $\boldsymbol{\mu}_q$ is a pure unit quaternion (*i.e.* such that $\boldsymbol{\mu}_q^2 = -1$) and $\phi_q \in [0, \pi)$ is the phase. Applying this polar decomposition to the monogenic signal enables identification of *local features* of the image $f(\mathbf{r})$ in a straightforward way. It reads

$$\mathcal{M}f(\mathbf{r}) = A(\mathbf{r}) \exp(\boldsymbol{\mu}_\theta(\mathbf{r})\phi(\mathbf{r})), \quad (3)$$

where $A(\mathbf{r}) := |\mathcal{M}f(\mathbf{r})|$ defines the *local amplitude*, $\phi(\mathbf{r})$ is the *local phase*, and where the axis $\boldsymbol{\mu}_\theta(\mathbf{r}) = \mathbf{i} \cos \theta(\mathbf{r}) + \mathbf{j} \sin \theta(\mathbf{r})$ defines a *local orientation* $\theta(\mathbf{r}) \in [-\pi, \pi)$. Note that the axis $\boldsymbol{\mu}_\theta(\mathbf{r})$ has no \mathbf{k} -component, as a result of the construction of the monogenic signal $\mathcal{M}f(\mathbf{r})$ along the imaginary axes \mathbf{i} and \mathbf{j} . As a first example, consider a plane wave $f(\mathbf{r}) = A_0 \cos(\boldsymbol{\kappa} \cdot \mathbf{r})$ where $\boldsymbol{\kappa} = (\kappa_1, \kappa_2) \in \mathbb{R}^2$ is the wavenumber vector. Direct computations of the Riesz transform yield the monogenic signal $\mathcal{M}f(\mathbf{r}) = A_0 \exp[(\boldsymbol{\kappa} \cdot \mathbf{r})(\mathbf{i} \cos \theta_0 + \mathbf{j} \sin \theta_0)]$ with $\theta_0 = \arg(k_1 + \mathbf{i}k_2)$. Hence the local amplitude is constant $A(\mathbf{r}) = A_0$, the local phase $\phi(\mathbf{r}) = \boldsymbol{\kappa} \cdot \mathbf{r}$ is directly that of the cosine wave, and local orientation is constant $\theta(\mathbf{r}) = \theta_0$ corresponding to that of the wavenumber vector \mathbf{k} in the 2D plane. Fig. 1 depicts a more sophisticated example corresponding to a 2D AM-FM mode. The monogenic signal permits a direct identification of Gaussian amplitude modulation kernel $A(\mathbf{r})$ and local orientation $\theta(\mathbf{r})$. The local phase $\phi(\mathbf{r})$ allows to detect lines ($\phi(\mathbf{r}) = 0 \pmod{\pi}$), simultaneously with contours ($\phi(\mathbf{r}) = \pi/2$).

The monogenic signal provides key insights into the geometry of 2D grayscale images. However, it suffers from the same limitations as the standard 1D analytic signal approach. It exhibits poor performance

in noisy settings and fails to capture meaningful local features when considering multicomponent 2D signals, such as superposition of 2D AM-FM modes. Therefore, an important line of research has focused on extending the monogenic signal approach towards multiscale or multiresolution analysis of grayscale images. The general idea is to devise multiple filterbanks built from *monogenic wavelets functions* at different scales. Then, at each scale or resolution, one can identify corresponding local features by computing the quaternion polar form of coefficients³. We mention hereafter some of these extensions. One is the monogenic continuous wavelet transform [14], which can be seen as a generalization of the 1D-analytic wavelet transform to the case of grayscale images. In the discrete case, a minimally-redundant monogenic multiresolution analysis was proposed in [13] using so-called Riesz-Laplace wavelets. A generalization of the curvelet transform to the monogenic case, called monogenic curvelet transform, was proposed in [15]. Other proposed approaches include transposing ideas from mode reconstructions in time-frequency analysis to the case of the monogenic signal, leading to the monogenic synchrosqueezing transform [16], or extending monogenic wavelet decompositions to the case of color images [17]. Monogenic signal-based approaches have found many applications (*e.g.* texture segmentation, target recognition, boundary detection) in various domains such as medical imaging, synthetic aperture radar imaging and geophysics.

Quaternion Fourier transform for bivariate signal processing

Bivariate signals appear in a broad range of applications where the joint analysis of two real-valued time series is required: polarized waveforms in seismology and optics, eastward and northward current velocities in oceanography, or even gravitational waves emitted by coalescing compact binaries. In such applications, it is crucial to provide clear and straightforward interpretations of the joint geometric and dynamic properties of the two components $x_1(t)$ and $x_2(t)$ that define the bivariate signal. Formally, a bivariate signal can be represented in two equivalent ways: a 2D time-evolving vector $\mathbf{x}(t) = [x_1(t), x_2(t)] \in \mathbb{R}^2$ or a complex-valued signal $x(t) = x_1(t) + \mathbf{i}x_2(t) \in \mathbb{C}$ encoding the two components on its real and imaginary parts. While the vector representation is generic (meaning that it is not restricted to the bivariate case), it also hinders a natural understanding of the geometric properties of bivariate signals. On the other hand, the complex representation permits the definition of a meaningful quaternion framework for bivariate signals relying on: (i) a dedicated Quaternion Fourier Transform

³For completeness, it is worth noting that not all works subsequent to the original seminal paper [12] use an explicit quaternion formulation of the monogenic signal. However, they largely make use of local angle and axis features, which are naturally connected to the quaternion polar form (3) of the monogenic signal. Therefore, these approaches can be labeled as *quaternion-based*, in a broad sense.

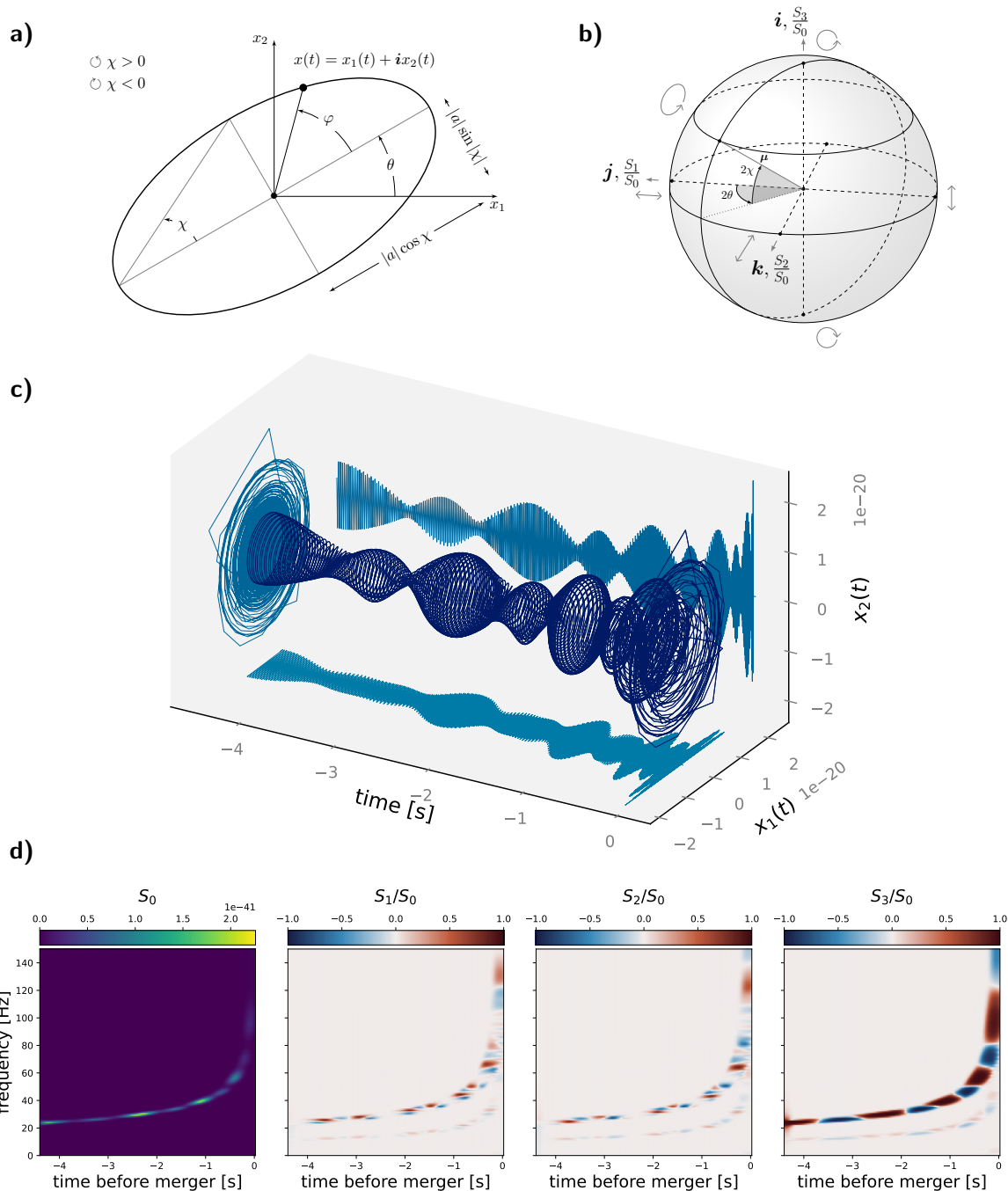


Fig. 2. **a)** Instantaneous ellipse parameters $[a, \theta, \chi, \phi]$ associated to the bivariate signal $x(t)$ thanks to the Euler polar form (6) of its quaternion embedding. **b)** Poincaré sphere of polarization states, a visual tool appearing naturally in the quaternion spectral representation of bivariate signals. Second-order properties of bivariate signals have a straightforward interpretation in terms of physical Stokes parameters S_0, S_1, S_2, S_3 . In turn, normalized Stokes parameters $S_1/S_0, S_2/S_0, S_3/S_0$ correspond to Cartesian coordinates of points on the Poincaré sphere. **c)** Typical gravitational wave bivariate signal emitted by a compact binary black hole system exhibiting precession of the orbital plane. **d)** Quaternion spectrogram of the bivariate signal depicted in c). It describes the time-frequency content of the signal, both in terms of energy (S_0) and polarization properties (normalized Stokes parameters $S_1/S_0, S_2/S_0, S_3/S_0$).

(QFT) and (ii) the extensive use of quaternion calculus (such as polar forms) to extract relevant physical and geometric information.

The key intuition for a quaternion spectral representation of bivariate signals is rather simple. For real-valued (that is, univariate) signals, the use of the standard complex Fourier transform enables a complex-valued spectral representation. This complex embedding of univariate signals is at the heart of definitions of amplitude and phase, which are crucial to many tasks of signal processing such as spectral analysis, filtering or time-frequency analysis. Now, if one represents bivariate signals as complex-valued signals, a quaternion embedding can be constructed in a similar way. First, observe that $x(t) = x_1(t) + ix_2(t) \in \mathbb{C}_i \subset \mathbb{H}$: it is a special case of a quaternion-valued signal. However, contrary to the complex Fourier transform, the QFT has no unique (or canonical) definition. Freedom of definition comes from the position of exponential, which can appear either left or right of the signal $x(t)$, and from the choice of the *axis* μ (a pure unit quaternion such that $\mu^2 = -1$) in the exponential. For instance, by choosing $\mu = i$ with $x(t) \in \mathbb{C}_i$ one recovers the standard complex Fourier transform. For bivariate signals, the right-sided QFT definition with $\mu = j$ is usually adopted [18]

$$X(f) = \int_{\mathbb{R}} x(t) e^{-j2\pi ft} dt, \quad x(t) \in \mathbb{C}_i. \quad (4)$$

The definition (4) exhibits every desirable property of FTs: it is well-defined for typical bivariate signals, it preserves energy and inner products (Parseval-Plancherel theorem) and it can be computed efficiently with 2 FFTs by observing that $X(f) = X_1(f) + iX_2(f)$, where $X_1(f), X_2(f)$ are standard (\mathbb{C}_j -valued) complex Fourier transforms of $x_1(t)$ and $x_2(t)$, respectively. More importantly, for bivariate signals viewed as \mathbb{C}_i -valued signals, it exhibits a Hermitian-like symmetry $X(-f) = -iX(f)i$, meaning that only the positive frequency spectrum carries relevant information. This makes it possible to define the *quaternion embedding* of a bivariate signal by canceling out the negative frequency spectrum. This bivariate analogue of the well-known analytic signal of real-valued univariate signals is defined as

$$x_+(t) = \int_{\mathbb{R}_+} X(f) e^{j2\pi ft} df. \quad (5)$$

The signal $x_+(t)$ is quaternion-valued. Therefore, at each time instant, it can be decomposed thanks to the Euler polar form of a quaternion $q = ae^{i\theta} e^{-k\chi} e^{j\phi}$ which identifies a magnitude $a := |q|$ and three phases corresponding to successive rotations around axes i, j and k . The Euler polar form plays the same role as the standard polar form for the usual analytic signal. It establishes a one-to-one mapping between the original bivariate signal $x(t)$ and a canonical quadruplet of instantaneous parameters $[a(t), \theta(t), \chi(t), \phi(t)]$, obtained by decomposing the quaternion embedding $x_+(t)$ as

$$x_+(t) = a(t) e^{i\theta(t)} e^{-k\chi(t)} e^{j\phi(t)}. \quad (6)$$

Under a classical narrow-band assumption on $x(t)$, it is now possible to attach a very insightful interpretation to the canonical parameters $[a(t), \theta(t), \chi(t), \phi(t)]$.

Fig. 2a) displays the instantaneous ellipse traced out by $x(t)$ in the (x_1, x_2) plane. The ellipse is characterized by its size $a(t)$, orientation $\theta(t)$ and shape $\chi(t)$: the last canonical parameter $\phi(t)$ corresponds to the dynamical phase, *i.e.* the instantaneous position of $x(t)$ within the ellipse. This shows that the instantaneous parameters have a natural geometric interpretation, which also corresponds to the physical notion of polarization in optics. The geometric insights provided by the quaternion representation do not stop there. Fig. 2b) gives the Poincaré sphere of polarization states, where each point of the 2-sphere corresponds to a given polarization state. Strikingly, it can be shown that the two quadratic quantities appearing in the Parseval-Plancherel theorem for the QFT correspond directly to the physical Stokes parameters S_0, S_1, S_2, S_3 , a set of four real-valued parameters widely used in optics to describe the polarization properties of light. The Poincaré sphere representation is particularly helpful, as it shows that normalized Stokes parameters $S_1/S_0, S_2/S_0, S_3/S_0$ are Cartesian coordinates for polarization states. For instance, the poles of the Poincaré sphere correspond to counter-rotating circularly polarized states, whereas its equator contain linear polarization states, with orientation given by the longitude. This key property enables the definition of power spectral densities or time-frequency representations (such as spectrograms or scalograms) that have a straightforward interpretation in terms of physical polarization parameters.

Fig. 2c) and Fig. 2d) showcase the time-frequency analysis of a gravitational wave (GW) bivariate signal thanks to the quaternion Fourier transform framework. Gravitational waves are bivariate signals, with two components $x_1(t)$ and $x_2(t)$ corresponding to *plus* and *cross* gravitational wave polarizations, respectively. Fig. 2c) depicts a simulated GW emitted by a binary black hole, a typical GW astrophysical source. The instantaneous variations of the interrelations between amplitude and phase of each component yields polarization modulation, which provides crucial insights towards the orbital motion of the GW (e.g. precession or orbital eccentricity). Fig. 2d) represents the quaternion spectrogram of the bivariate GW signal depicted in Fig. 2c). The quaternion spectrogram is a natural generalization of the classical spectrogram: it relies on the definition of the quaternion short-term Fourier transform (QSTFT), which follows directly by considering the QFT (4) of windowed, time-translated signal segments $x(t)g(t - \tau)$, where τ is a time-shift and $g(t)$ is usually a real symmetric window with compact support [18]. Second-order conservation properties of the QSTFT lead to representing and interpreting the quaternion spectrogram in terms of time-frequency Stokes parameters. This four time-frequency maps must be analyzed jointly: S_0 is alike the standard well-know spectrogram and gives the energy distribution of the signal in the time-frequency plane; the normalized Stokes parameters $S_1/S_0, S_2/S_0, S_3/S_0$ reveal the

instantaneous polarization state along ridges (*i.e.* lines of maximal energy in S_0).

More generally, the quaternion spectral representation of bivariate signals enables the definition of a full algebraic signal processing framework. It puts *polarization* at the heart of many fundamentals tools of bivariate signal processing, such as general time-frequency or time-scale representations [18], the definition of power spectral densities and their estimation [19] and the design of linear time-invariant filters [20]. Quaternions provide considerable insights into the geometric properties of bivariate signals by leveraging the well-established language of polarization in optics, without sacrificing the usual mathematical guarantees nor the availability of computationally cheap implementations.

III. THE QUATERNION TOOLBOX FOR SIGNAL PROCESSING

The interesting upper mentioned benefits of quaternion-valued representations immediately raise a natural question: is it possible to extend standard signal processing tools (e.g. low-rank decompositions, minimization of cost functions, least mean squares algorithms, etc.) to the quaternion setting in some natural way? If the answer was negative, then quaternion-valued representations would be of little practical use, since the insights gained by choosing quaternion representations would be canceled out by cumbersome processing. Fortunately this is not the case: standard tools of signal and image processing extend nicely over quaternion algebra. Still, this is not straightforward: the non-commutativity of the multiplication of two quaternions usually prevents direct extensions from the real and complex case. Over the last three decades, researchers have dedicated a lot of energy to establish a meaningful signal processing toolbox to deal with quaternion-valued data. We propose hereafter a tour of these important tools, and illustrate their relevance on flagship applications.

A. Low-rank quaternion models

Among the numerous applications of linear algebra in signal processing, low-rank approximation plays a central role and algorithms for performing EigenValue Decomposition (EVD) or Singular Value Decomposition (SVD) of data matrices are used on a daily basis in a wide range of applications. Existence of EVD and SVD for quaternion matrices is known for a long time, even though standard concepts of linear algebra have to be considered with caution because of the non-commutativity of the quaternion product. For instance, one must distinguish between *right* and *left* eigenvalues, depending on the side position of the scalar in eigenvalue problems. Given a square quaternion-valued matrix $\mathbf{A} \in \mathbb{H}^{N \times N}$, its *right* eigenvalues $\lambda_r \in \mathbb{H}$ are defined as solutions of $\mathbf{A}\mathbf{x} = \mathbf{x}\lambda_r$, where $\mathbf{x} \in \mathbb{H}^N$ is a nonzero quaternion vector. Similarly, left eigenvalues $\lambda_l \in \mathbb{H}$ are defined as solutions of $\mathbf{A}\mathbf{x} = \lambda_l\mathbf{x}$. *Right* eigenvalues are well understood and used in all applications of quaternions in data science until now. On

the contrary, *left* eigenvalues are much more cumbersome: their study remains an active field of research in mathematics. For signal processing practice, dealing with right eigenvalues is sufficient: they define natural quaternion counterparts of usual real and complex eigenvalues and can be computed efficiently. They also share some similar properties with their classical counterparts: for instance, a Hermitian matrix \mathbf{A} (*i.e.* $\mathbf{A}^H = \mathbf{A}$, where $(\cdot)^H$ is the usual conjugate-transpose operator) necessarily has real-valued right eigenvalues. Moreover, a necessary and sufficient condition for Hermitian \mathbf{A} to be positive semidefinite (*i.e.* $\mathbf{x}^H \mathbf{A} \mathbf{x} \geq 0$ for any \mathbf{x}) is to have all its right eigenvalues nonnegative.

The quaternion singular value decomposition (Q-SVD) of a rectangular quaternion-valued matrix $\mathbf{A} \in \mathbb{H}^{N \times P}$ is defined as [21]:

$$\mathbf{A} = \mathbf{U} \mathbf{\Delta} \mathbf{V}^H, \quad \text{with } \mathbf{U} \in \mathbb{H}^{N \times N}, \mathbf{V} \in \mathbb{H}^{P \times P}, \mathbf{\Delta} \in \mathbb{R}_+^{N \times P}. \quad (7)$$

The expression of Q-SVD looks very familiar: $\mathbf{\Delta}$ is a rectangular diagonal matrix with non-negative real numbers (the singular values of \mathbf{A}) on its diagonal $\Delta_{ii} = \delta_i$, \mathbf{U} and \mathbf{V} are unitary quaternion matrices, meaning that $\mathbf{U}^H \mathbf{U} = \mathbf{I}_N$ and $\mathbf{V}^H \mathbf{V} = \mathbf{I}_P$ where \mathbf{I} is the identity matrix. Columns of \mathbf{U} define the left singular vectors of \mathbf{A} , whereas columns of \mathbf{V} give its right singular vectors. Moreover, Q-SVD enables the computation of the best rank- ℓ approximation of a quaternion matrix in the Frobenius norm: one simply truncates the Q-SVD by keeping only the first ℓ terms corresponding to the ℓ largest singular values. As an important special case of the Q-SVD, consider the case of an Hermitian matrix $\mathbf{A} \in \mathbb{H}^{N \times N}$. Its Q-SVD reads $\mathbf{A} = \mathbf{U} \mathbf{\Sigma} \mathbf{U}^H$, where $\mathbf{\Sigma} = \text{diag}(|\lambda_1|, \dots, |\lambda_N|)$ and $\lambda_1, \dots, \lambda_N \in \mathbb{R}$ are the (right) eigenvalues of \mathbf{A} . Therefore, if \mathbf{A} is positive semidefinite, then its right eigenvalues are real nonnegative, and the Q-SVD of \mathbf{A} is identical to its quaternion right EVD (Q-EVD). In particular, columns of \mathbf{U} correspond to (right) eigenvector of \mathbf{A} . As a result, computing the Q-EVD of a positive definite Hermitian matrix comes down computing its Q-SVD, for which efficient algorithms exist [22]. This result is particularly important for subspace methods in array processing, which rely on diagonalization of covariance matrices built from quaternion-valued samples, such as quaternion MUSIC (Q-MUSIC). See details below.

The Q-SVD plays a crucial role in the definition of any subspace methods over quaternions. In vector-sensor array processing, for example, Q-SVD allows to separate quaternion data into signal and noise subspaces and to take advantage of vector-sensors measurements to perform source localization [22]–[25]. However, one may question the relevance of using quaternions for subspace estimation. The mathematical soundness of quaternion subspace methods relies on the orthogonality constraint inherited from the scalar product over \mathbb{H}^N . The columns of \mathbf{U} and of \mathbf{V} in (7) are orthogonal, meaning that $\mathbf{u}_\ell^H \mathbf{u}_{\ell'} = 0$ and $\mathbf{v}_\ell^H \mathbf{v}_{\ell'} = 0$ for $\ell \neq \ell'$. This actually imposes that the real and the three imaginary parts of this scalar product cancel out simultaneously. Such a constraint is much stronger than complex orthogonality (considering

pair of complex-valued vectors rather than quaternion vectors). As a consequence, quaternion subspace methods are usually more robust to model errors than their long-vector complex counterparts.

Quaternion-MUSIC for vector-sensor array processing An illustrative example of application of subspace methods in signal processing is the estimation of the Direction Of Arrival (DOA) of polarized sources impinging on a vector-sensor array using the quaternion MUSIC (Q-MUSIC) algorithm introduced in [23]. For 2D vector-sensors (crossed dipoles, geophones, acoustic vector-sensors, etc.), a narrow-band fully polarized source can be represented as the quaternion embedding of a bivariate signal (6) with linear phase $\phi(t) = \omega_0 t$, slowly varying amplitude $a(t)$ and constant polarization parameters $[\theta, \chi]$. Consider a uniform linear array made of M 2D vector-sensors, identically oriented, with equal spacing Δx . Assuming a single far-field source, the response of the m -th sensor reads $x_m(t) = a(t) \exp[i\theta] \exp[-k\chi] \exp[j\omega_0 t] \exp[-j\tau_m(\alpha)]$ where $\tau_m(\alpha) = (2\pi/\lambda)(m-1)\Delta x \sin \alpha$, is the phase delay between the first and m th sensors. Here λ denotes the wavelength and α the DOA of the source in the array plane. Therefore, a snapshot of the array at time instant t can be written as the quaternion-valued vector $\mathbf{x}(t) \in \mathbb{H}^M$ such that

$$\mathbf{x}(t) = x_0(t)\mathbf{v}(\alpha), \text{ with } \mathbf{v}(\alpha) = \left[1, e^{-j\tau_2(\alpha)}, \dots, e^{-j\tau_M(\alpha)}\right]^\top, \quad (8)$$

where $x_0(t)$ is the quaternion-valued signal measured by the first sensor and $(\cdot)^\top$ denotes the transpose of a matrix. Generalizing this model to a noise-corrupted linear superposition of K far-field polarized sources, a snapshot of the output of the sensor-array can be expressed as:

$$\mathbf{y}(t) = \sum_{k=1}^K x_{0k}(t)\mathbf{v}_k(\alpha_k) + \mathbf{n}(t), \quad (9)$$

with \mathbf{v}_k the steering vector associated to source k of DOA α_k and where $\mathbf{n}(t) \in \mathbb{H}^M$ is the noise vector on the array. Following the rationale of subspace methods developed in the real and complex cases, one forms the quaternion-valued covariance matrix of the data $\mathbf{C}_{yy} = E\{\mathbf{y}(t)\mathbf{y}^H(t)\} \in \mathbb{H}^{M \times M}$, where $E\{\cdot\}$ denotes the mathematical expectation. Then, under some commonly accepted assumptions (uncorrelated sources and white noise), the matrix \mathbf{C}_{yy} is shown [23] to be written as:

$$\mathbf{C}_{yy} = \mathbf{D}\mathbf{\Sigma}\mathbf{D}^H + \sigma_n^2 \mathbf{I}_M, \quad (10)$$

where the columns of $\mathbf{D} \in \mathbb{H}^{M \times K}$ correspond to the K source vectors $\mathbf{d}_k(\theta_k, \chi_k, \alpha_k) = e^{i\theta_k} e^{-k\chi_k} \mathbf{v}_k(\alpha_k)$, $\mathbf{\Sigma} = \text{diag}(\sigma_1^2, \dots, \sigma_K^2)$ with $\sigma_k^2 = |x_{0k}(t)|^2$ the k th source power and σ_n^2 is the noise power. Now, classical arguments yield that the signal subspace of the observations is spanned by the K quaternion eigenvectors associated to the K largest eigenvalues of \mathbf{C}_{yy} . The remaining $M - K$ eigenvectors form an orthonormal basis in the noise subspace. Recall that in this case, the Q-EVD of \mathbf{C}_{yy} boils down to the

Q-SVD (since it is Hermitian positive semidefinite). By leveraging the orthogonality between the signal and noise subspaces, the expression of the Q-MUSIC functional is:

$$\mathcal{S}_{\text{Q-MUSIC}}(\theta, \chi, \alpha) = \frac{1}{\mathbf{d}^H(\theta, \chi, \alpha) \mathbf{\Pi} \mathbf{d}(\theta, \chi, \alpha)}, \quad (11)$$

where $\mathbf{d}(\theta, \chi, \alpha) = e^{i\theta} e^{-k\chi} \mathbf{v}(\alpha)$ is the steering vector and $\mathbf{\Pi} = \mathbf{U}_{K+1} \mathbf{U}_{K+1}^H$ is the orthogonal projector on the noise subspace spanned by $\mathbf{U}_{K+1} = [\mathbf{u}_{K+1}, \dots, \mathbf{u}_M]$. Similarly to standard MUSIC algorithm for scalar sensors, the K highest local maxima of (11) correspond to the DOAs and polarization parameters of the K sources.

Classically, when dealing with 2D vector-sensors, the two complex-valued components of size M of the array are simply concatenated in a long-vector of size $2M$. Then, the standard MUSIC algorithm is applied to this enhanced data vector. This is commonly known as the long-vector MUSIC (LV-MUSIC). Compared to LV-MUSIC, where classical orthogonality is imposed globally over the two stacked components of the array, the quaternion-vector orthogonality associated with Q-MUSIC implies stronger orthogonality constraints (component-wise) between signal and noise subspaces. It was shown in [23] that these quaternion constraints enable a better separation between subspaces. This results in improved performance (in terms of estimation error and resolution) compared to LV-MUSIC, while reducing the size of the data covariance representation. Fig. 3 provides a brief illustration of the advantages of Q-MUSIC compared to its long-vector counterpart. Following these results, several other quaternion methods [4], [6], [26], extending the standard subspace-based DOA estimation algorithms to the quaternion framework have been proposed.

B. Solving optimization problems in the quaternion domain

Modern signal and image processing problems can often be formulated as solutions of optimization problems. Quaternion domain problems are no exception. For instance, the quaternion least mean squares (QLMS) algorithm described in Section III-C or the sparse coding of 3D/4D data (see further below) can all be formulated as the problem of minimizing a real-valued cost function f with quaternion arguments. Formally, one is interested in solving optimization problems of the form

$$\hat{q} = \arg \min_{q \in \mathbb{H}} f(q), \quad (12)$$

where $f : \mathbb{H} \rightarrow \mathbb{R}$ is a cost function (e.g. a quadratic form) and \hat{q} is the solution of the minimization problem. This formulation extends easily to the case of quaternion vectors $\mathbf{q} \in \mathbb{H}^M$ and quaternion matrices $\mathbf{Q} \in \mathbb{H}^{M \times N}$ variables. Now, how can one solve (12)? If q was a standard real-valued variable, it would be sufficient to explore the stationary points of f (points where the derivative of f vanishes, at

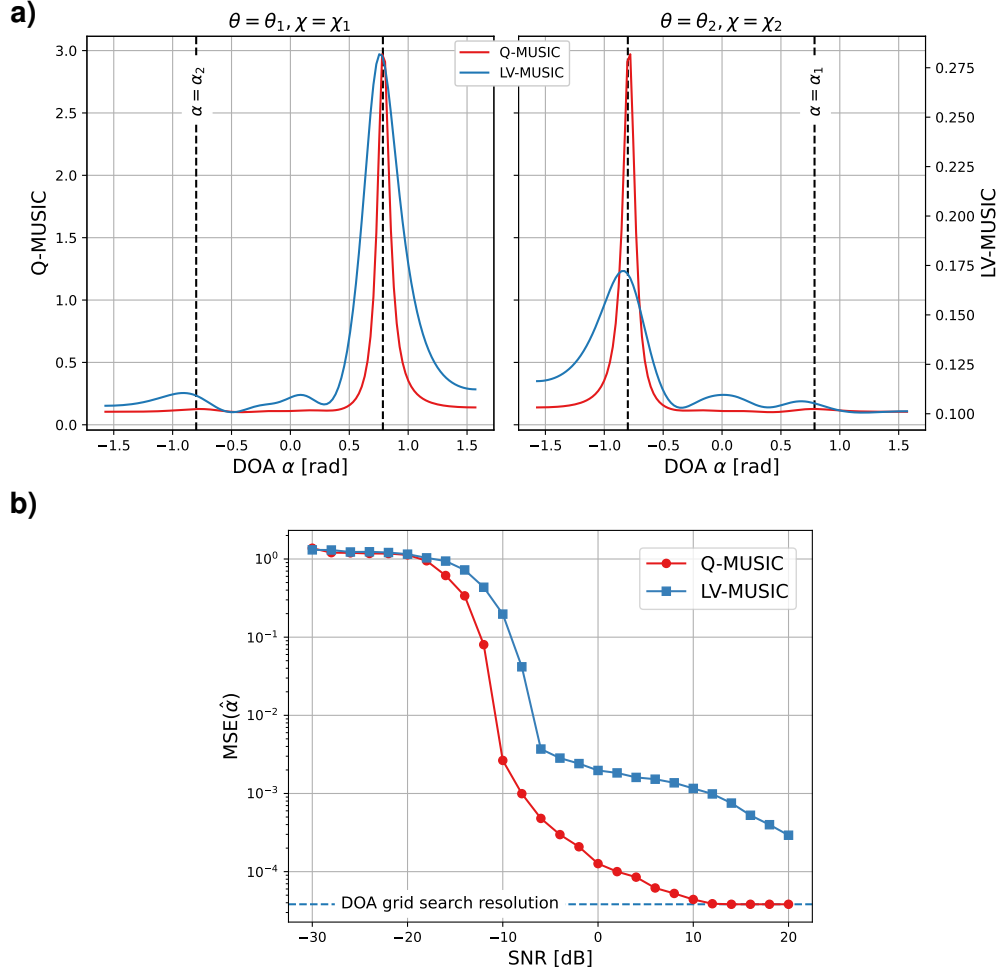


Fig. 3. Comparison between Q-MUSIC and LV-MUSIC for a uniform linear array with $M = 10$ sensors with spacing $\Delta x = \lambda/4$. SNR was set to 0 dB assuming \mathbb{H} -proper Gaussian noise. Sample covariance matrix were computed using 32 samples in all experiments. a) Plots of the Q-MUSIC (red) and LV-MUSIC (blue) functionals for two far fields polarized sources with DOAs α_1 and α_2 . The two panels depicts values of functionals with respect to angle α , along two slices corresponding to polarization parameters of each source (left: $\theta = \pi/3, \chi = \pi/5$, right: $\theta = -\pi/6, \chi = -\pi/5$). b) Performance comparison in terms of MSE on the DOA in a single source scenario. Curves depicts average values computed using 1000 independents runs.

least in the differentiable case) to obtain the different local minima of the cost function f . Unfortunately, the same approach cannot be applied *directly* with quaternion variables. The main reason is that quaternion differentiability (in a quaternion analysis sense) can only be defined for *analytic* functions $f : \mathbb{H} \rightarrow \mathbb{H}$. In particular, cost functions – which are of interest for signal and image processing – are not analytic as they only take values in \mathbb{R} . On the other hand, the real-valued cost function f in (12) of quaternion argument $q = q_a + iq_b + jq_c + kq_d$ can also be considered as a function of its four real-valued

components q_a, q_b, q_c, q_d , such that $f(q) = f(q_a, q_b, q_c, q_d)$, with little abuse of notation. This formal identification suggests that one could compute the derivative of the function f as if it was a function of \mathbb{R}^4 . This apparent contradiction between real-domain and quaternion-domain derivatives is addressed by the theory of (generalized) $\mathbb{H}\mathbb{R}$ -calculus, developed in the first half of the 2010 decade [27]–[30]. In essence, it forms a complete set of calculus rules that permits computations of derivatives of functions of quaternion variables, as if they were functions of their real and imaginary parts. Still, it operates directly in the quaternion domain, meaning that the burden of transforming (12) into an equivalent real-domain optimization problem can be completely avoided.

Before going further, let us mention that the construction of $\mathbb{H}\mathbb{R}$ -calculus is closely related to the framework of $\mathbb{C}\mathbb{R}$ -calculus (also known as *Wirtinger* calculus) for solving optimization problems in complex variables – often encountered in complex-valued signal processing. Still, the derivation of $\mathbb{H}\mathbb{R}$ -calculus is far from being trivial. Many differences and subtleties arise when generalizing from complex variables to quaternion variables, in particular due to the noncommutativity of the quaternion product. While we will only survey the basics of $\mathbb{H}\mathbb{R}$ -calculus here to illustrate its usefulness, we refer the interested reader to the seminal papers [27]–[29] for detailed proofs and extensive results.

Let us consider a cost function $f : \mathbb{H} \rightarrow \mathbb{R}$. We suppose, for convenience, that f is real-differentiable [29], that is the function $f(q_a, q_b, q_c, q_d)$ is differentiable with respect to its four variables q_a, q_b, q_c and q_d . The $\mathbb{H}\mathbb{R}$ -derivatives of f with respect to q and \bar{q} are defined as [30]

$$\frac{\partial f}{\partial q} = \frac{1}{4} \left(\frac{\partial f}{\partial q_a} - \frac{\partial f}{\partial q_b} \mathbf{i} - \frac{\partial f}{\partial q_c} \mathbf{j} - \frac{\partial f}{\partial q_d} \mathbf{k} \right), \quad (13)$$

$$\frac{\partial f}{\partial \bar{q}} = \frac{1}{4} \left(\frac{\partial f}{\partial q_a} + \frac{\partial f}{\partial q_b} \mathbf{i} + \frac{\partial f}{\partial q_c} \mathbf{j} + \frac{\partial f}{\partial q_d} \mathbf{k} \right). \quad (14)$$

More generally, it is also possible to define $\mathbb{H}\mathbb{R}$ -derivatives with respect to canonical involutions q^i, q^j, q^k and their respective conjugates. Equations (13) – (14) show that $\mathbb{H}\mathbb{R}$ -derivatives are quaternion-valued, as they collect partial derivatives of $f(q)$ with respect to each one of the components of q along the canonical basis $\{1, \mathbf{i}, \mathbf{j}, \mathbf{k}\}$. Remark that since we assume f to be real-valued, the two $\mathbb{H}\mathbb{R}$ -derivatives are related through conjugation $\overline{(\partial f / \partial q)} = \partial f / \partial \bar{q}$. However, to turn the fundamental definitions (13) – (14) into a practical framework for computing quaternion derivatives, it is necessary to consider additional derivatives with respect to q^μ and \bar{q}^μ where $\mu \in \mathbb{H}$ and $\mu \neq 0$. This leads to the *generalized* $\mathbb{H}\mathbb{R}$ -calculus [29], which equips (13) – (14) with convenient product and chain rules. It is worth noting that these properties hold for arbitrary quaternion-valued functions $f : \mathbb{H} \rightarrow \mathbb{H}$, *i.e.* the theory is not limited to cost functions. Importantly, the (generalized) $\mathbb{H}\mathbb{R}$ -calculus framework has been extended [27] to compute derivatives with respect to quaternion vector and matrix variables. Vector and matrix counterparts of $\partial f / \partial q$ and $\partial f / \partial \bar{q}$ are denoted in quite a natural way, *i.e.* $\nabla_q f$ and $\nabla_{\bar{q}} f$ denote $\mathbb{H}\mathbb{R}$ -gradients of $f(q)$,

TABLE II

$\mathbb{H}\mathbb{R}$ -DERIVATIVES OF USUAL COST FUNCTIONS OF QUATERNION SCALAR, VECTOR AND MATRIX ARGUMENTS. ADAPTED FROM [27], [29].

Scalar functions $f : \mathbb{H} \rightarrow \mathbb{R}$	$\frac{\partial f}{\partial \bar{q}}$	notes
$f(q) = \text{Re}(\alpha q \beta + \gamma)$	$\frac{1}{4} \bar{\alpha} \bar{\beta}$	$\alpha, \beta, \gamma \in \mathbb{H}$
$f(q) = \alpha q \beta + \gamma $	$\frac{1}{4} \frac{\bar{\alpha}(\alpha q \beta + \gamma) \bar{\beta}}{ \alpha q \beta + \gamma }$	$\alpha, \beta, \gamma \in \mathbb{H}$
$f(q) = \alpha q \beta + \gamma ^2$	$\frac{1}{2} \bar{\alpha} (\alpha q \beta + \gamma) \bar{\beta}$	$\alpha, \beta, \gamma \in \mathbb{H}$
Vector functions $f : \mathbb{H}^M \rightarrow \mathbb{R}$	$\nabla_{\bar{q}} f$	
$f(\mathbf{q}) = \text{Re}(\mathbf{b}^H \mathbf{q})$	$\frac{1}{4} \mathbf{b}$	$\mathbf{b} \in \mathbb{H}^M$
$f(\mathbf{q}) = \mathbf{q}^H \mathbf{A} \mathbf{q}$	$\frac{1}{2} \mathbf{A} \mathbf{q}$	$\mathbf{A} \in \mathbb{H}^{M \times M}, \mathbf{A}^H = \mathbf{A}$
$f(\mathbf{q}) = \ \mathbf{A} \mathbf{q} - \mathbf{b}\ _{\mathbb{H}^M}^2$	$\frac{1}{2} \mathbf{A}^H (\mathbf{A} \mathbf{q} - \mathbf{b})$	$\mathbf{A} \in \mathbb{H}^{M \times N}, \mathbf{b} \in \mathbb{H}^M$
Matrix functions $f : \mathbb{H}^{M \times N} \rightarrow \mathbb{R}$	$\nabla_{\bar{Q}} f$	
$f(\mathbf{Q}) = \text{Re}(\text{trace}\{\mathbf{A} \mathbf{Q} \mathbf{B}\})$	$\frac{1}{4} \mathbf{A}^H \mathbf{B}^H$	$\mathbf{A} \in \mathbb{H}^{L \times M}, \mathbf{B} \in \mathbb{H}^{N \times L}$
$f(\mathbf{Q}) = \ \mathbf{A} - \mathbf{B} \mathbf{Q} \mathbf{C}\ _F^2$	$\frac{1}{2} \mathbf{B}^H (\mathbf{B} \mathbf{Q} \mathbf{C} - \mathbf{A}) \mathbf{C}^H$	$\mathbf{A} \in \mathbb{H}^{L \times P}, \mathbf{B} \in \mathbb{H}^{L \times M}, \mathbf{C} \in \mathbb{H}^{N \times P}$

while $\nabla_{\mathbf{Q}} f$ and $\nabla_{\bar{\mathbf{Q}}} f$ correspond to matrix derivatives of the function $f(\mathbf{Q})$. Table II collects such derivatives for some elementary cost functions in the scalar, vector and matrix case. Only derivatives with respect to conjugates $\bar{q}, \bar{\mathbf{q}}$ and $\bar{\mathbf{Q}}$ are given; since the considered functions are real-valued, it suffices to take the quaternion conjugate to obtain derivatives with respect to q, \mathbf{q} or \mathbf{Q} . Moreover, observe the similarity between $\mathbb{H}\mathbb{R}$ -derivatives and usual real-domain functions: the only noticeable difference being the unusual $1/4$ prefactor arising from definitions (13) – (14).

As explained earlier, solving (12) requires the ability to find stationary points of the cost function. The $\mathbb{H}\mathbb{R}$ -calculus enables their characterization as one would expect: for instance, a vector \mathbf{q}_* is a stationary point of $f(\mathbf{q})$ if the $\mathbb{H}\mathbb{R}$ -gradient vanishes at this point, *i.e.* $\nabla_{\mathbf{q}} f(\mathbf{q}_*) = \nabla_{\bar{\mathbf{q}}} f(\mathbf{q}_*) = \mathbf{0}$. Moreover, the direction of the maximum rate of change of f at point \mathbf{q} is given by $\nabla_{\bar{\mathbf{q}}} f(\mathbf{q})$, that is the $\mathbb{H}\mathbb{R}$ -gradient with respect to $\bar{\mathbf{q}}$. This yields immediately the steepest gradient descent method in the quaternion domain

as iterations

$$\mathbf{q}^{(k+1)} = \mathbf{q}^{(k)} - \eta_k \nabla_{\bar{\mathbf{q}}} f(\mathbf{q}^{(k)}) \quad (15)$$

where $\eta_k > 0$ is the step size at iteration k . Eq. (15) is one of the first building blocks of quaternion-domain optimization. For instance, it represents a cornerstone of the derivation of the quaternion LMS (QLMS) algorithm and its variants – as we shall explain in detail in Section III-C. The generalized $\mathbb{H}\mathbb{R}$ -derivatives framework can be further refined. For instance, it is possible to define a meaningful quaternion Hessian matrix, thus paving the way to extend well-known second-order algorithms such as the Newton method to the quaternion domain. We refer the interested reader to [28] for further details on this topic.

To further illustrate the generalized $\mathbb{H}\mathbb{R}$ -calculus framework, let us consider the quaternion sparse coding problem. Let $\mathbf{y} \in \mathbb{H}^M$ be a vector of observations which one seeks to represent in a dictionary $\mathbf{D} \in \mathbb{H}^{M \times N}$ as $\mathbf{y} \approx \mathbf{D}\mathbf{q}$ where \mathbf{q} is a sparse vector and $N > M$. This problem appears naturally in color imaging, where \mathbf{y} is a color patch encoded as a pure quaternion vector, \mathbf{D} is an overcomplete collection of color atoms and \mathbf{q} is a sparse quaternion vector. Just like in the real and complex case, this problem can be solved using greedy algorithms based on ℓ_0 penalty, such as quaternion orthogonal matching pursuit (Q-OMP) [8], [31]. On the other hand, following standard practice, one can relax the non-convex ℓ_0 -penalty into a convex ℓ_1 -norm regularization leading to the quaternion LASSO (Q-LASSO) problem [32], [33] (or equivalently, quaternion basis pursuit denoising):

$$\arg \min_{\mathbf{q} \in \mathbb{H}^M} \|\mathbf{y} - \mathbf{D}\mathbf{q}\|_2^2 + \lambda \|\mathbf{q}\|_1, \quad \lambda \geq 0. \quad (16)$$

In (16) the quaternion ℓ_1 -norm of vector $\mathbf{q} \in \mathbb{H}^M$ is defined as the sum of moduli of its entries $\|\mathbf{q}\|_1 := \sum_{m=1}^M |q_m|$. Interestingly, since the quaternion ℓ_1 -norm collects a sum of moduli, it can be interpreted as a mixed norm $\ell_{2,1}$ on the real matrix \mathbf{A} obtained by concatenating the four real-valued components of \mathbf{q} such that $\mathbf{A} = [\mathbf{q}_a \ \mathbf{q}_b \ \mathbf{q}_c \ \mathbf{q}_d] \in \mathbb{R}^{M \times 4}$, i.e. $\|\mathbf{q}\|_1 = \|\mathbf{A}\|_{2,1}$. This means that the Q-LASSO (16) can be interpreted in the real domain as a group Lasso with M groups of size 4.

The problem (16) defines a convex optimization problem in quaternion variables. In practice, quaternion convex optimization problems can be solved by translating usual real domain algorithms to the quaternion case by leveraging the generalized $\mathbb{H}\mathbb{R}$ -calculus. Still, this generalization is not trivial and requires special care [34]. In particular, the general form of equality constraints in quaternion convex optimization problems is *widely affine*: i.e. it reads $\mathbf{A}_1\mathbf{q} + \mathbf{A}_2\mathbf{q}^i + \mathbf{A}_3\mathbf{q}^j + \mathbf{A}_4\mathbf{q}^k = \mathbf{b}$, where matrices and vectors are quaternion-valued with appropriate sizes. In comparison, in real-domain convex optimization, only affine equality constraints of the form $\mathbf{A}\mathbf{x} = \mathbf{b}$, where vectors are all real-valued, are considered. To solve the (unconstrained) problem (16), one can adapt the celebrated Iterative Shrinkage-Thresholding

Algorithm (ISTA) to handle quaternion variables. Generalized $\mathbb{H}\mathbb{R}$ -calculus makes it possible to derive the quaternion ISTA iterations in a intuitive way. Letting $f(\mathbf{q}) = \|\mathbf{y} - \mathbf{D}\mathbf{q}\|_2^2$, the iterations read

$$\mathbf{q}^{(k+1)} = \mathcal{T}_{\lambda\eta_k} \left\{ \mathbf{q}^{(k)} - \eta_k \nabla_{\bar{\mathbf{q}}} f(\mathbf{q}) \right\}, \quad (17)$$

where $\eta_k > 0$ is the step size at iteration k and \mathcal{T}_β is the soft-thresholding operator (*i.e.* the proximal operator associated to the quaternion ℓ_1 -norm) given entry-wise by $\mathcal{T}_\beta(q) = \max(0, 1 - \beta/|q|)q$. Note that in this case, the gradient can be directly computed thanks to Table II as $\nabla_{\bar{\mathbf{q}}} f(\mathbf{q}) = \frac{1}{2} \mathbf{D}^H (\mathbf{D}\mathbf{q} - \mathbf{b})$.

Of course, more sophisticated optimization problems can be formulated and solved directly in the quaternion domain. For instance, in the case of 3D data sparse coding (such as color images) that uses pure quaternions, it might be relevant to impose that the solution \mathbf{q}_* satisfies $\text{Re}(\mathbf{D}\mathbf{q}_*) = \mathbf{0}$. This constraint is *widely linear* – meaning that constraining the problem (16) preserves its convex nature. Solving this type of constrained quaternion optimization problems can be carried out within the same general framework, for instance using the quaternion alternating direction of multiplier method (Q-ADMM), as explained in [34].

C. Statistics for quaternion random variables

Statistics for quaternion random variables and vectors has been developed in order to extend classical signal processing algorithms relying on probabilistic models. As 4D variables, quaternions can either be understood as four dimensional real vector-valued random variables, *i.e.* variables in \mathbb{R}^4 or as two dimensional complex random vectors, *i.e.* variables in \mathbb{C}^2 . Just like for complex-valued random variables and vectors, detecting and taking into account symmetries in the probability density function (pdf) of quaternion random variables is essential in devising powerful signal processing tools. The notion of *properness* (also known as *second order circularity*) captures such rotation invariance of the pdf. It was considered in many scenarios and exploited in several algorithms either as an extra parameter in the signal model or as a signature of the absence/presence of a targeted signal hidden in a noisy environment. A solid literature is available for the complex case, (see e.g., [35] and references therein).

In the quaternion case, the original study and definition of *properness* traces back to Vakhania's work [36], before being considered by the signal processing community [37]–[40]. The major difference between the complex and quaternion cases is the existence of different levels of properness over \mathbb{H} while only one level can be identified over \mathbb{C} . The three levels of properness of a quaternion-valued random variable⁴ are denoted \mathbb{R} -properness, $(1, \boldsymbol{\mu})$ -properness (where $\boldsymbol{\mu}$ is a pure unit quaternion, also denoted

⁴The notion can be directly extended to vectors, but we illustrate here the concept on scalar-valued variable for clarity.

\mathbb{C}^μ -properness) and \mathbb{H} -properness [38], [40]. These levels correspond to different invariance properties of the random variable distribution. In the Gaussian case, *properness* means invariance properties of the 4×4 real (augmented) covariance matrix $\mathbf{\Gamma}_{\mathbb{R}} = E\{\mathbf{q}_{\mathbb{R}}\mathbf{q}_{\mathbb{R}}^{\top}\}$, where $\mathbf{q}_{\mathbb{R}} = [q_a, q_b, q_c, q_d]^{\top} \in \mathbb{R}^4$. Thus, symmetries of the quaternion Gaussian pdf can be interpreted as symmetries of the equivalent real covariance matrix of the augmented vector $\mathbf{q}_{\mathbb{R}}$ gathering the real and imaginary parts of $q \in \mathbb{H}$. Fig. 4 gives an illustration of the structure induced by two levels of *properness* on the 4×4 real covariance matrix $\mathbf{\Gamma}_{\mathbb{R}}$ in the case of centered quaternion-valued Gaussian variables. While the \mathbb{H} -properness level is the “simplest” (the four components of the quaternion variables are uncorrelated and have the same variance), the $(1, \mu)$ -proper case has a much more complicated correlation structure (here $\mu = j$ for illustration purposes).

Accounting for properness is essential in quaternion-valued signal processing. It means that the symmetries, or invariances, of the distribution are correctly taken into account. Therefore, an important research effort has been dedicated to develop and formulate statistical tools to deal with proper and improper quaternion signals. This includes extensions of generalized likelihood ratio method to test properness levels of quaternion variables [41] or the development of quaternion independent component analysis based on second-order statistics [42]. In estimation, accounting for properness naturally yields to the notion of widely linear processing [38], [43], which exploits the dependence between a quaternion random variable \mathbf{x} and its three canonical involutions \mathbf{x}^i , \mathbf{x}^j and \mathbf{x}^k to capture the full second-order information, hence providing optimal estimation performance.

Quaternion adaptive filtering for 3D and 4D signals Another important application of quaternion-valued signal processing is provided by the quaternion implementation of the popular Least Mean Square (LMS) adaptive filter. When dealing with 3D and 4D signals, using quaternion representations along with the linear quaternion model allows exploiting the geometric coupling between channels in a natural and structured manner. This is what motivated the development of the *Quaternion Least Mean Squares* (QLMS) algorithm. The first formulation of QLMS was proposed in [5] and applied to temperature and 3D-wind data measurements. Subsequently, various forms of quaternion adaptive filters have been introduced along with their performance analysis [29], [44]–[48].

We provide hereafter, an overview of the derivation of the update expression for the QLMS filter weights. QLMS algorithm is generally used to estimate the quaternion valued vector \mathbf{w} of coefficients of a desired filter, by minimizing the least mean square of the error signal (difference between the desired and the actual signal). If n is the number of the current input sample $\mathbf{x}(n)$, QLMS aims at minimizing the real-valued cost function of quaternion variables $\mathcal{J}(n) = |e(n)|^2$, where $e(n) = d(n) - y(n)$ is the error between the desired signal $d(n)$ and the filter output $y(n) = \mathbf{w}(n)^{\text{H}}\mathbf{x}(n)$. This minimization is

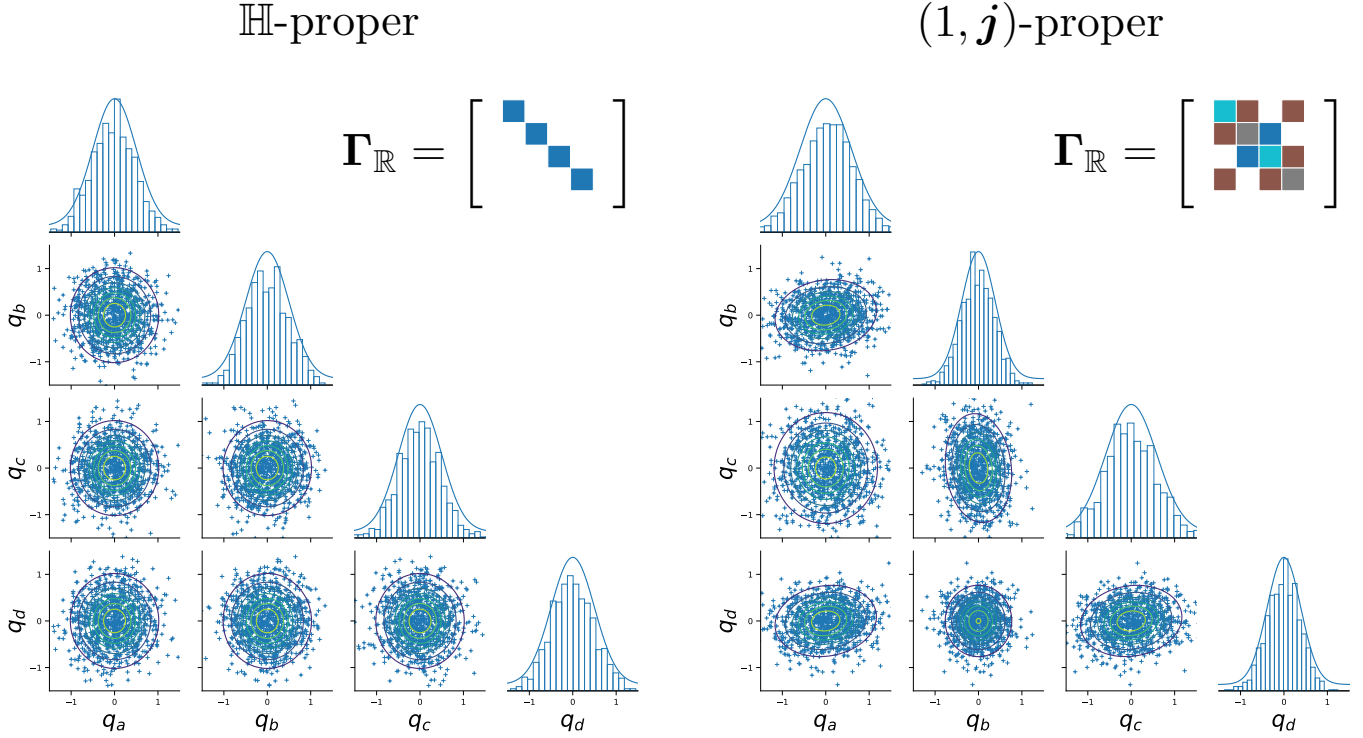


Fig. 4. Two different properness levels for a centered quaternion scalar Gaussian variable. Insets show the corresponding structure of the augmented covariance matrix $\Gamma_{\mathbb{R}}$, where identical colors indicate equal values.

performed using a steepest descent procedure, and the QLMS weights update is given by (see (15)):

$$\mathbf{w}(n+1) = \mathbf{w}(n) - \eta \nabla_{\bar{\mathbf{w}}} \mathcal{J}(n), \quad (18)$$

where $\eta > 0$ is the step size. Using the \mathbb{H} R-calculus rules from Table II and the general quaternion calculus rules, it can be easily shown that $\nabla_{\bar{\mathbf{w}}} \mathcal{J}(n) = -\mathbf{x}(n)\bar{\mathbf{e}}(n)/2$. Thus, absorbing the scalar factor in the step size η , the QLMS update of the weight vector reads:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \eta \mathbf{x}(n)\bar{\mathbf{e}}(n). \quad (19)$$

One can observe that the quaternion update (19) has a similar form to the one obtained in the complex case. It is well-known that the performance of the LMS filter depends on the second-order statistics on the input vector \mathbf{x} . When \mathbf{x} is \mathbb{H} -proper, the standard Hermitian covariance matrix $\mathbf{C}_{\mathbf{x}\mathbf{x}} = E\{\mathbf{x}\mathbf{x}^{\mathbb{H}}\}$ contains the full second-order information of the data. However, when \mathbf{x} is not \mathbb{H} -proper, it is necessary to consider the so-called *augmented statistics*, *i.e.* the correlations between \mathbf{x} and its three involutions $\mathbf{x}^i, \mathbf{x}^j$ and \mathbf{x}^k to capture the full second-order information. This leads to the *widely-linear* QLMS (WL-QLMS) [29], [47], [48], that has a similar form to (19), but for which the input vector \mathbf{x} is replaced by

the augmented input $\mathbf{x}_a = [\mathbf{x}^T, \mathbf{x}^i{}^T, \mathbf{x}^j{}^T, \mathbf{x}^k{}^T]^T$. A detailed analysis of QLMS performance for different types of properness can be found in [47]. Application of QLMS and WL-QLMS on benchmark 3D data (Lorenz attractor) as well as real-world 3D and 4D data [5] has shown that the quaternion approach outperforms LMS and complex LMS (CLMS) in terms of accuracy, proving the capacity of quaternion models to fusion heterogeneous data sources.

IV. REMAINING CHALLENGES AND OPEN PROBLEMS IN QUATERNION-VALUED SIGNAL AND IMAGE PROCESSING

It has been more than three decades now since the first appearance of quaternions in the signal and image processing literature. Still, the use of quaternions in these domains regularly raises scepticism. For instance, the question “Do quaternion-valued approaches perform better than standard real or complex-valued methods?” is perhaps the most encountered. Even though this might be an interesting question to ask, it misses a more fundamental point. Rather, we believe one should be questioning whether *the use of a quaternion-valued representation should be preferred over a more conventional real-valued representation*. Answering explicitly this question is crucial. It justifies the use of quaternions in the first place, which can be motivated *e.g.* by the geometric insights offered by the algebra, or by the ease of formulating compact and explicit models for the data at hand. In essence, one must identify the unique properties offered by quaternion algebra that are relevant to the considered setting. Once the motivation for using *quaternion representations* clearly established, the comparison of *quaternion methods* with other real-valued approaches could be conducted. It is important to keep in mind that *quaternion models* can always be transformed into equivalent (often structured) real-domain models. In that case, comparisons between quaternion and real approaches are usually meaningless, since they describe the same model or methodology, just expressed using different algebra rules. However, one can legitimately wonder whether such equivalent real-domain model could have been directly formulated without the design of a quaternion model in the first place. Therefore, we believe that discussions should focus on pros and cons for the respective modeling approaches.

Quaternion representations call for dedicated tools. Indeed, if one wishes to advocate for the benefits of quaternion-valued representations and models, then one should also develop the associated signal processing methodologies in the quaternion domain. Doing otherwise, *e.g.* resorting to an equivalent real-domain formulation to solve the problem at hand, would imply losing any advantages gained using the original quaternion formulation. The tour of the current quaternion-valued signal and image processing toolbox presented in Section III paves the way for end-to-end quaternion domain approaches. Nonetheless, many questions remain open. In a nutshell, these define exciting research directions for quaternion-valued

signal and image processing: the search for new insightful quaternion embeddings, the design of dedicated quaternion-valued statistical models, novel decompositions for large quaternion datasets, and machine learning with quaternions. These promising avenues of research are briefly outlined below:

- *Novel models in the quaternion domain* Often, models for quaternion signal and data are built by mimicking standard linear algebra. This is the case, *e.g.* for the linear mixing model used in many quaternion signal processing scenarios. However, the intrinsic link between quaternions and 3D / 4D geometry, such as rotations or reflections, allow for much more diverse modeling opportunities. Future research should take into account this specificity of quaternion algebra to develop original and innovative models capable of revealing further geometric insights into the studied phenomena. This includes, for instance, models capable of identifying and separating features in polarimetric or color imaging that are invariant to some geometric transformation.
- *Statistical quaternion signal processing* While there exists an important literature in complex-valued statistical signal processing [35], the quaternion domain counterpart still exhibits plenty of opportunities for theoretical and methodological developments. Indeed, until now, most of the research effort has been concentrated towards establishing the notions of *properness* and widely linear filtering. As a result, fundamental building blocks of quaternion statistical signal processing are still missing. For example, the definition of general quaternion probability distributions beyond the Gaussian case is still a largely open question. Constructing a satisfying and complete probabilistic framework for quaternions would pave the way for meaningful end-to-end quaternion domain statistical approaches such as Bayesian inference, hypothesis testing, performance bounds for estimators, etc.
- *Structured decompositions of multidimensional quaternion arrays* We have illustrated in this article the fact that orthogonal low-rank decompositions of quaternion-valued matrices (Q-SVD, Q-EVD) have been successfully used in various applications such as color image compression or vector-sensor array processing. However, these orthogonal decompositions quickly reach their limits when we must deal with either structured (*e.g.* Hankel matrices) or constrained (*e.g.* non-negative) data, as it is often the case in machine learning applications, for example. There is a need for new versatile low-rank decomposition tool allowing to account for these various constraints. Several authors already proposed such structured decompositions for dictionary learning [8], sparse coding [32], [33] or for learning features from polarized imaging data [49], but this research area is still in its infancy. Another interesting research direction is the extension of multilinear (tensor) algebra models and decompositions to quaternions. Because of the non-commutativity of quaternion multiplication, the definition of multilinearity in the quaternion framework is a challenging but important theoretical

question, for which there are already some promising leads.

- *Quaternion neural networks* Motivated by the success of quaternion encoding of 3D/4D signals, the recent years have witnessed the development of quaternion (convolutional, recurrent) neural networks (NNs) [50]. By leveraging the benefits of quaternion algebra that permits efficient accounting for inter-channel relations, such NNs allow improved feature learning from 3D and 4D data. This usually results in overall better performance than standard real-valued NNs in terms of classification / prediction rates and in a lower number of learned parameters. This is the case, for example, for the convolutional NNs that exploit correlations between the red, green and blue channels of a color image seen as a quaternion matrix. However, there is still a long way to go before using these NNs on a large scale. One of the main bottlenecks for these approaches is the hardware implementation of the quaternion product which is a powerful but expensive operation in the context of quaternion NNs. Optimized GPU implementations for this operation would allow to drastically improve computation efficiency for these systems, and therefore their practical utility. An exciting prospect lies in devising quaternion NNs that take full advantage of quaternion-based optimization (the $\mathbb{H}\mathbb{R}$ -calculus) and which are likely to reduce the computational burden and improve the efficiency of learning algorithms in applications relying on quaternion representations.

Almost 180 years after their discovery nearby Brougham bridge, quaternions are still surrounded by a hint of mystery for the newcomer. We hope the present article will help to demystify their use and benefits to researchers in signal and image processing. If a large set of works have already illustrated the power of quaternions to propose elegant and interpretable solutions to difficult problems, there remains some necessary effort to even better understand and exploit their potential. Quaternions will certainly continue to intrigue and inspire the signal processing community for the years to come.

REFERENCES

- [1] J. Chou, "Quaternion kinematic and dynamic differential equations," *IEEE Transactions on Robotics and Automation*, vol. 8, no. 1, pp. 53–64, 1992. [Online]. Available: <https://ieeexplore.ieee.org/document/127239>
- [2] K. Shoemake, "Animating rotation with quaternion curves," in *Proceedings of the 12th annual conference on Computer graphics and interactive techniques - SIGGRAPH '85*. ACM Press, 1985, pp. 245–254. [Online]. Available: <http://portal.acm.org/citation.cfm?doid=325334.325242>
- [3] H.-D. Schutte and J. Wenzel, "Hypercomplex numbers in digital signal processing," in *IEEE International Symposium on Circuits and Systems*. New Orleans, LA, USA: IEEE, 1990, pp. 1557–1560. [Online]. Available: <http://ieeexplore.ieee.org/document/112431/>
- [4] 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*, vol. 29, no. 3, pp. 29–39, May 2012. [Online]. Available: <http://ieeexplore.ieee.org/document/6179823/>

- [5] C. Took and D. Mandic, "The Quaternion LMS Algorithm for Adaptive Filtering of Hypercomplex Processes," *IEEE Transactions on Signal Processing*, vol. 57, no. 4, pp. 1316–1327, Apr. 2009. [Online]. Available: <http://ieeexplore.ieee.org/document/4696053/>
- [6] H. Chen, W. Wang, and W. Liu, "Augmented Quaternion ESPRIT-Type DOA Estimation With a Crossed-Dipole Array," *IEEE Communications Letters*, vol. 24, no. 3, pp. 548–552, Mar. 2020. [Online]. Available: <https://ieeexplore.ieee.org/document/8943327/>
- [7] J.-F. Gu and K. Wu, "Quaternion Modulation for Dual-Polarized Antennas," *IEEE Communications Letters*, vol. 21, no. 2, pp. 286–289, Feb. 2017. [Online]. Available: <http://ieeexplore.ieee.org/document/7579660/>
- [8] Q. Barthelemy, A. Larue, and J. I. Mars, "Color Sparse Representations for Image Processing: Review, Models, and Prospects," *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 3978–3989, Nov. 2015. [Online]. Available: <http://ieeexplore.ieee.org/document/7161332/>
- [9] Y. Chen, X. Xiao, and Y. Zhou, "Low-Rank Quaternion Approximation for Color Image Processing," *IEEE Transactions on Image Processing*, vol. 29, pp. 1426–1439, 2020. [Online]. Available: <https://ieeexplore.ieee.org/document/8844978/>
- [10] T. Bulow and G. Sommer, "Hypercomplex signals—a novel extension of the analytic signal to the multidimensional case," *IEEE Transactions on Signal Processing*, vol. 49, no. 11, pp. 2844–2852, Nov. 2001. [Online]. Available: <http://ieeexplore.ieee.org/document/960432/>
- [11] W. L. Chan, H. Choi, and R. Baraniuk, "Coherent Multiscale Image Processing Using Dual-Tree Quaternion Wavelets," *IEEE Transactions on Image Processing*, vol. 17, no. 7, pp. 1069–1082, Jul. 2008. [Online]. Available: <http://ieeexplore.ieee.org/document/4526699/>
- [12] M. Felsberg and G. Sommer, "The monogenic signal," *IEEE Transactions on Signal Processing*, vol. 49, no. 12, pp. 3136–3144, Dec. 2001. [Online]. Available: <http://ieeexplore.ieee.org/document/969520/>
- [13] M. Unser, D. Sage, and D. Van De Ville, "Multiresolution Monogenic Signal Analysis Using the Riesz–Laplace Wavelet Transform," *IEEE Transactions on Image Processing*, vol. 18, no. 11, pp. 2402–2418, Nov. 2009. [Online]. Available: <http://ieeexplore.ieee.org/document/5164973/>
- [14] S. Olhede and G. Metikas, "The Monogenic Wavelet Transform," *IEEE Transactions on Signal Processing*, vol. 57, no. 9, pp. 3426–3441, Sep. 2009. [Online]. Available: <http://ieeexplore.ieee.org/document/4956988/>
- [15] M. Storath, "Directional Multiscale Amplitude and Phase Decomposition by the Monogenic Curvelet Transform," *SIAM Journal on Imaging Sciences*, vol. 4, no. 1, pp. 57–78, Jan. 2011. [Online]. Available: <http://epubs.siam.org/doi/10.1137/100803924>
- [16] M. Clausel, T. Oberlin, and V. Perrier, "The monogenic synchrosqueezed wavelet transform: a tool for the decomposition/demodulation of AM–FM images," *Applied and Computational Harmonic Analysis*, vol. 39, no. 3, pp. 450–486, Nov. 2015. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S1063520314001274>
- [17] R. Souillard and P. Carré, "Elliptical Monogenic Wavelets for the Analysis and Processing of Color Images," *IEEE Transactions on Signal Processing*, vol. 64, no. 6, pp. 1535–1549, Mar. 2016. [Online]. Available: <http://ieeexplore.ieee.org/document/7347476/>
- [18] J. Flamant, N. Le Bihan, and P. Chainais, "Time–frequency analysis of bivariate signals," *Applied and Computational Harmonic Analysis*, vol. 46, no. 2, pp. 351–383, Mar. 2019. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S1063520317300507>
- [19] —, "Spectral Analysis of Stationary Random Bivariate Signals," *IEEE Transactions on Signal Processing*, vol. 65, no. 23, pp. 6135–6145, Dec. 2017. [Online]. Available: <http://ieeexplore.ieee.org/document/8003393/>
- [20] J. Flamant, P. Chainais, and N. Le Bihan, "A Complete Framework for Linear Filtering of Bivariate Signals,"

- IEEE Transactions on Signal Processing*, vol. 66, no. 17, pp. 4541–4552, Sep. 2018. [Online]. Available: <https://ieeexplore.ieee.org/document/8410469/>
- [21] F. Zhang, “Quaternions and matrices of quaternions,” *Linear Algebra and its Applications*, vol. 251, pp. 21–57, Jan. 1997. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/0024379595005439>
- [22] N. Le Bihan and J. Mars, “Singular value decomposition of quaternion matrices: a new tool for vector-sensor signal processing,” *Signal Processing*, vol. 84, no. 7, pp. 1177–1199, Jul. 2004. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0165168404000647>
- [23] S. Miron, N. Le Bihan, and J. Mars, “Quaternion-MUSIC for vector-sensor array processing,” *IEEE Transactions on Signal Processing*, vol. 54, no. 4, pp. 1218–1229, Apr. 2006. [Online]. Available: <http://ieeexplore.ieee.org/document/1608539/>
- [24] X. Gong, Y. Xu, and Z. Liu, “Quaternion ESPRIT for direction finding with a polarization sensitive array,” in *2008 9th International Conference on Signal Processing*. Beijing, China: IEEE, Oct. 2008, pp. 378–381. [Online]. Available: <http://ieeexplore.ieee.org/document/4697149/>
- [25] A. Sajeve and G. Menanno, “Characterisation and extraction of a Rayleigh-wave mode in vertically heterogeneous media using quaternion SVD,” *Signal Processing*, vol. 136, pp. 43–58, Jul. 2017. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0165168416303255>
- [26] X. Gong, Y. Xu, and Z. Liu, “Quaternion ESPRIT for direction finding with a polarization sensitive array,” in *2008 9th International Conference on Signal Processing*. IEEE, 2008, pp. 378–381. [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/4697149>
- [27] D. Xu and D. P. Mandic, “The Theory of Quaternion Matrix Derivatives,” *IEEE Transactions on Signal Processing*, vol. 63, no. 6, p. 14, 2015. [Online]. Available: <https://ieeexplore.ieee.org/document/7029694/>
- [28] D. Xu, Y. Xia, and D. P. Mandic, “Optimization in Quaternion Dynamic Systems: Gradient, Hessian, and Learning Algorithms,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 2, pp. 249–261, Feb. 2016. [Online]. Available: <http://ieeexplore.ieee.org/document/7124503/>
- [29] D. Xu, C. Jahanchahi, C. C. Took, and D. P. Mandic, “Enabling quaternion derivatives: the generalized HR calculus,” *Royal Society Open Science*, vol. 2, no. 8, p. 150255, Aug. 2015. [Online]. Available: <https://royalsocietypublishing.org/doi/10.1098/rsos.150255>
- [30] D. P. Mandic, C. Jahanchahi, and C. C. Took, “A Quaternion Gradient Operator and Its Applications,” *IEEE Signal Processing Letters*, vol. 18, no. 1, pp. 47–50, Jan. 2011. [Online]. Available: <http://ieeexplore.ieee.org/document/5623300/>
- [31] Y. Xu, L. Yu, H. Xu, H. Zhang, and T. Nguyen, “Vector Sparse Representation of Color Image Using Quaternion Matrix Analysis,” *IEEE Transactions on Image Processing*, vol. 24, no. 4, pp. 1315–1329, Apr. 2015. [Online]. Available: <http://ieeexplore.ieee.org/document/7024169/>
- [32] C. Zou, K. I. Kou, and Y. Wang, “Quaternion Collaborative and Sparse Representation With Application to Color Face Recognition,” *IEEE Transactions on Image Processing*, vol. 25, no. 7, pp. 3287–3302, Jul. 2016. [Online]. Available: <http://ieeexplore.ieee.org/document/7468463/>
- [33] C. Zou, K. I. Kou, Y. Wang, and Y. Y. Tang, “Quaternion block sparse representation for signal recovery and classification,” *Signal Processing*, vol. 179, p. 107849, Feb. 2021. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0165168420303935>
- [34] J. Flamant, S. Miron, and D. Brie, “A General Framework for Constrained Convex Quaternion Optimization,” *IEEE Transactions on Signal Processing*, vol. 70, pp. 254–267, 2022. [Online]. Available: <https://ieeexplore.ieee.org/document/9662233/>
- [35] P. J. Schreier and L. L. Scharf, *Statistical Signal Processing of Complex-Valued Data: The Theory of Improper and*

- Noncircular Signals*. Cambridge University Press, 2010. [Online]. Available: <https://doi.org/10.1017/CBO9780511815911>
- [36] N. N. Vakhania, “Random Vectors with Values in Quaternion Hilbert Spaces,” *Theory of Probability & Its Applications*, vol. 43, no. 1, pp. 99–115, Jan. 1999. [Online]. Available: <http://epubs.siam.org/doi/10.1137/S0040585X97976696>
- [37] P.-O. Amblard and N. Le Bihan, “On Properness Of Quaternion Valued Random Variables,” in *IMA Conference on Mathematics in Signal Processing*, 2004.
- [38] J. Via, D. Ramirez, and I. Santamaria, “Properness and Widely Linear Processing of Quaternion Random Vectors,” *IEEE Transactions on Information Theory*, vol. 56, no. 7, pp. 3502–3515, Jul. 2010. [Online]. Available: <http://ieeexplore.ieee.org/document/5484995/>
- [39] C. Cheong Took and D. P. Mandic, “Augmented second-order statistics of quaternion random signals,” *Signal Processing*, vol. 91, no. 2, pp. 214–224, Feb. 2011. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0165168410002719>
- [40] N. Le Bihan, “The geometry of proper quaternion random variables,” *Signal Processing*, vol. 138, pp. 106–116, Sep. 2017. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0165168417301093>
- [41] J. Via, D. P. Palomar, and L. Vielva, “Generalized Likelihood Ratios for Testing the Properness of Quaternion Gaussian Vectors,” *IEEE Transactions on Signal Processing*, vol. 59, no. 4, pp. 1356–1370, Apr. 2011. [Online]. Available: <http://ieeexplore.ieee.org/document/5672618/>
- [42] J. Via, D. P. Palomar, L. Vielva, and I. Santamaria, “Quaternion ICA From Second-Order Statistics,” *IEEE Transactions on Signal Processing*, vol. 59, no. 4, pp. 1586–1600, Apr. 2011. [Online]. Available: <http://ieeexplore.ieee.org/document/5672617/>
- [43] J. Navarro-Moreno, R. M. Fernandez-Alcala, and J. C. Ruiz-Molina, “A Quaternion Widely Linear Model for Nonlinear Gaussian Estimation,” *IEEE Transactions on Signal Processing*, vol. 62, no. 24, pp. 6414–6424, Dec. 2014. [Online]. Available: <http://ieeexplore.ieee.org/document/6936344/>
- [44] C. C. Took and D. P. Mandic, “A Quaternion Widely Linear Adaptive Filter,” *IEEE Transactions on Signal Processing*, vol. 58, no. 8, pp. 4427–4431, 2010. [Online]. Available: <https://ieeexplore.ieee.org/document/5447699>
- [45] B. Che Ujang, C. C. Took, and D. P. Mandic, “Quaternion-Valued Nonlinear Adaptive Filtering,” *IEEE Transactions on Neural Networks*, vol. 22, no. 8, pp. 1193–1206, Aug. 2011. [Online]. Available: <http://ieeexplore.ieee.org/document/5934421/>
- [46] F. Ortolani, D. Communiello, M. Scarpiniti, and A. Uncini, “Frequency domain quaternion adaptive filters: Algorithms and convergence performance,” *Signal Processing*, vol. 136, pp. 69–80, Jul. 2017. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0165168416302948>
- [47] M. Xiang, S. Kanna, and D. P. Mandic, “Performance Analysis of Quaternion-Valued Adaptive Filters in Nonstationary Environments,” *IEEE Transactions on Signal Processing*, vol. 66, no. 6, pp. 1566–1579, 2018. [Online]. Available: <https://ieeexplore.ieee.org/document/8239856>
- [48] M. Xiang, Y. Xia, and D. P. Mandic, “Performance Analysis of Deficient Length Quaternion Least Mean Square Adaptive Filters,” *IEEE Transactions on Signal Processing*, vol. 68, pp. 65–80, 2020. [Online]. Available: <https://ieeexplore.ieee.org/document/8911458/>
- [49] J. Flamant, S. Miron, and D. Brie, “Quaternion Non-Negative Matrix Factorization: Definition, Uniqueness, and Algorithm,” *IEEE Transactions on Signal Processing*, vol. 68, pp. 1870–1883, 2020. [Online]. Available: <https://ieeexplore.ieee.org/document/9001262/>
- [50] T. Parcollet, M. Morchid, and G. Linarès, “A survey of quaternion neural networks,” *Artificial Intelligence Review*, vol. 53, no. 4, pp. 2957–2982, Apr. 2020. [Online]. Available: <http://link.springer.com/10.1007/s10462-019-09752-1>