Signal Processing 91 (2011) 214–224

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Augmented second-order statistics of quaternion random signals

Clive Cheong Took *, Danilo P. Mandic

Department of Electrical and Electronic Engineering, Imperial College, London SW7 2AZ, UK

ARTICLE INFO

Article history:
Received 21 December 2009
Received in revised form
7 May 2010
Accepted 25 June 2010
Available online 7 July 2010

Keywords:
Quaternion random variables
Augmented quaternion statistics
$\mathbb{Q}$-properness
Widely linear modelling
Multivariate Gaussian distribution
Quadrivariate entropy
Quadrivariate random processes
Second order circularity
Widely linear quaternion Wiener model

ABSTRACT

Second order statistics of quaternion random variables and signals are revisited in order to exploit the complete second order statistical information available. The conditions for $\mathbb{Q}$-proper (second order circular) random processes are presented, and to cater for the non-vanishing pseudocovariance of such processes, the use of $\imath$-$\jmath$-κ-covariances is investigated. Next, the augmented statistics and the corresponding widely linear model are introduced, and a generic multivariate Gaussian distribution is subsequently derived for both $\mathbb{Q}$-proper and $\mathbb{Q}$-improper processes. The maximum entropy bound and an extension of mutual information to multivariate processes are derived in order to provide a complete description of joint information theoretic properties of general quaternion valued processes. A comparative analysis with the corresponding second order statistics of quadrivariate real valued processes supports the approach.

1. Introduction

Standard techniques employed in statistical multichannel processing typically do not fully exploit the 'coupled' nature of the available information within the channels. Most practical approaches are based on channelwise processing—this is often inadequate as the components of a multichannel process are typically correlated. On the other hand, the quaternion domain $\mathbb{H}$ facilitates modelling of three- and four-dimensional signals, and accounts for the mutual information between the data channels in a natural way; this has been reflected in an increasing number of recent applications based on quaternion modelling. In the signal processing community, quaternions have been employed in Kalman filtering [1], MUSIC spectrum estimation [2], singular value decomposition for vector sensing [3], and the least-mean-square estimation [4]. However, these applications have also revealed some problems in using standard second order statistics for general quaternionic signals, especially for processes with of different powers in data channels, such as in wind modelling [4]. For instance, it is clear that in most scenarios the two horizontal wind components will have much large dynamics than the vertical wind component, leading to noncircular three-dimensional signal. Recently, there has been a large effort to introduce complex-valued algorithms suitable for the processing of both circular and noncircular signals [5]. However, despite quaternions being a natural generalisation of complex numbers (their hypercomplex extension), the developments in the 'augmented' statistics of general processes (both second order circular and noncircular) in the quaternion domain are still in their infancy.[1]

It is therefore natural to investigate whether the recent developments in so-called augmented complex statistics and widely linear modelling in the complex domain can

* Corresponding author.
E-mail addresses: c.cheong-took@ic.ac.uk (C.C. Took), d.mandic@ic.ac.uk (D.P. Mandic).

[1] Second order circular signals are termed proper, whereas second order noncircular signals are termed improper.

0165-1684/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.sigpro.2010.06.024

be extended to the quaternion domain, in order to provide theoretical rigour and enhanced practical accuracy. One of the pioneering results in augmented complex statistics is the work by Neeser and Massey, who provide a comprehensive account of the concept of properness (second order circularity, rotation invariant probability distribution). They demonstrated that the covariance matrix $E\{\mathbf{z}\mathbf{z}^H\}$ of a complex random vector $\mathbf{z}$ alone is not adequate to describe a complete second order statistical information [6] for general signals and that the pseudocovariance matrix or complementary covariance matrix $E\{\mathbf{z}\mathbf{z}^T\}$ also needs to be considered. Further, Van Den Bos formulated a generic multivariate Gaussian distribution of both proper and improper complex processes, to show that the traditional definition of the complex Gaussian distribution (based on the covariance) is only a special case, applicable to proper processes only [7]. These foundations have been successfully used to design novel algorithms in adaptive signal processing [5], autoregressive moving average (ARMA) modelling [8], and independent component analysis [9].

Existing statistical signal processing approaches in $\mathbb{H}$ typically take into account only the information contained in the quaternion-valued covariance [10,1,2,11]; by analogy with the complex domain, this is not guaranteed to maximise the use of the available second order statistical information. In this direction, Vakhania extended the concept of 'properness' to the quaternion domain; however, his definition of $\mathbb{Q}$-properness is restricted to the invariance of the probability density function (pdf) under some specific rotations around angle of $\pi/2$ [12]. Amblard and Le Bihan further relaxed the conditions of $\mathbb{Q}$-properness to an arbitrary axis and angle of rotation φ, that is [13]

$$q \triangleq e^{\nu\varphi} q \quad \forall \varphi \tag{1}$$

for any pure unit quaternion ν (whose real part vanishes); symbol $\triangleq$ denotes equality in terms of pdf. These authors formulated the Gaussian distribution for single quaternion-valued variables in the complex domain, based on the Cayley–Dickson representation, whereby, a quaternion variable q is represented as a pair of two complex variables z_1 and z_2, that is, $q = z_1 + \jmath z_2$ [14]. Buchholz and Le Bihan also employed the Cayley–Dickson representation to give further insight into the complex-valued statistics for quaternion variables [15]. These results provide an initial insight into the statistics of quaternion variables; however, they are lacking generality, as they either consider single quaternion variables or are formulated indirectly via the complex domain. This makes them not straightforward to apply to multivariate quaternion-valued random vectors, or to provide a unifying framework for the second order statistical modelling of general quaternion signals.

This work aims to provide a unifying framework for the second order statistics of quaternion variables together with deriving the conditions for complete second order statistical description of both second order circular and noncircular signals. We demonstrate that in order to exploit complete second order information, it is necessary to incorporate complementary covariance matrices, thus accounting for a possible improperness of quaternion processes. The benefits of such an approach are thus likely to be analogous to the advantages that the augmented statistics provides for noncircular complex-valued data [5,16]. The analysis shows that the basis for augmented quaternion statistics should comprise quaternion involutions. The so-introduced augmented covariance matrix contains all the necessary second order statistical information, and paves the way for widely linear modelling in $\mathbb{H}$. Next, multivariate Gaussian distribution is revisited in order to cater for general quaternion processes, leading to enhanced entropy based descriptors. Finally, conditions for $\mathbb{Q}$-properness (second order circularity) are presented, and it is shown that $\mathbb{Q}$-proper Gaussian processes attain maximum entropy.

The organisation of the paper is as follows: in Section 2 we briefly review the elements of quaternion algebra. In Section 3, novel statistical measures for quaternion-valued variables are introduced and the duality with their quadrivariate real domain counterparts is addressed. Next, Section 4 revisits the fundamentals of $\mathbb{Q}$-properness and illustrates its implications on quaternion statistics. Section 5 illustrates an application of the augmented quaternion statistics in adaptive filtering. Section 6 formulates a generic Gaussian distribution to cater for both $\mathbb{Q}$-proper and $\mathbb{Q}$-improper signals. In Section 7, the upper bound of the entropy of a multivariate quaternion-valued data is derived, and it is shown that it is attained for $\mathbb{Q}$-proper signals. Further, the so-called interaction information, an extension of mutual information to multivariate processes, is introduced. We conclude this work in Section 8. For convenience, a summary of notations is given in Table 1.

Table 1
Summary of notations for quaternion variables $q = q_a + \imath q_b + \jmath q_c + \kappa q_d$.

Notations	Description
$\Re\{\cdot\}$	Scalar real part q_a
$\Im\{\cdot\}$	Vector imaginary part $\imath q_b + \jmath q_c + \kappa q_d$
$\Im_{\imath,\jmath,\kappa}\{\cdot\}$	$\imath$-,$\jmath$-,κ-component of vector imaginary part of q
$\times$	Cross-product
$q^{\imath,\jmath,\kappa}$	$\imath$-$\jmath$-κ-involution given in (6)
$(\cdot)^*$	Quaternion conjugate operator
$(\cdot)^T$	Quaternion transpose operator
$(\cdot)^H$	Quaternion conjugate transpose operator
$\mathbf{A}$	$4N \times 4N$ mapping matrix $\mathbb{H} \rightarrow \mathbb{R}^4$ given in (10)
$\mathbf{q}^a$	Augmented quaternion-valued vector defined in (10)
$\mathbf{q}^r$	Quadrivariate real-valued vector defined in (10)
$\mathcal{C}_{\mathbf{qq}}$	Standard covariance matrix defined in (13)
$\mathcal{C}_{\mathbf{q}\imath}$	$\imath$-covariancematrix defined in (14)
$\mathcal{C}_{\mathbf{q}\jmath}$	$\jmath$-covariance matrix defined in (15)
$\mathcal{C}_{\mathbf{q}\kappa}$	κ-covariance matrix defined in (16)
σ^2	Variance
$\mathcal{I}(\cdot)$	Interaction information

2. Properties of quaternion random vectors

2.1. Quaternion algebra

Quaternion domain provides a natural framework for a unified treatment of three- and four-dimensional processes and can be regarded as a non-commutative extension of complex numbers [17]. A quaternion variable $q \in \mathbb{H}$ comprises a real part $\Re\{\cdot\}$ (denoted by subscript a) and a vector-part, also called a pure quaternion $\Im\{\cdot\}$, consisting of three imaginary components (denoted by subscripts b, c, and d), and can be expressed as

$$\begin{aligned} q &= \Re\{q\}+\Im\{q\} \\ &= \Re\{q\}+\imath\Im_\imath\{q\}+\jmath\Im_\jmath\{q\}+\kappa\Im_k\{q\} \\ &= q_a+\imath q_b+\jmath q_c+\kappa q_d \in \mathbb{H} \end{aligned} \tag{2}$$

The orthogonal unit vectors, $\imath,\jmath,\kappa$ not only describe the three axes of the part of a quaternion, but are also imaginary numbers; their relationships are given by

$$\begin{aligned} &\imath\jmath=\kappa \quad \jmath\kappa=\imath \quad \kappa\imath=\jmath \\ &\imath\jmath\kappa=\imath^2=\jmath^2=\kappa^2=-1 \end{aligned} \tag{3}$$

For every $q_1,q_2 \in \mathbb{H}$, quaternion multiplication is defined as

$$q_1q_2=\Re\{q_1q_2\}+\Im\{q_1q_2\}$$

$$\text{where} \quad \Re\{q_1q_2\}=q_{1,a}q_{2,a}+q_{1,b}q_{2,b}+q_{1,c}q_{2,c}+q_{1,d}q_{2,d}$$

$$\Im\{q_1q_2\}=q_{1,a}\Im\{q_2\}+q_{2,a}\Im\{q_1\}+\Im\{q_1\}\times\Im\{q_2\} \tag{4}$$

where the symbol '$\times$' denotes the vector product; observe that $q_1q_2=q_2q_1-2\Im\{q_2\}\times\Im\{q_1\}\neq q_2q_1$. The non-commutativity of the quaternion product is a consequence of the vector product in (4). The quaternion conjugate is defined as

$$q^*=\Re\{q\}-\Im\{q\}=q_a-\imath q_b-\jmath q_c-\kappa q_d \tag{5}$$

2.2. Quaternion involutions and the augmented basis vector

Complex calculus allows for the real and imaginary part of a complex number $z=z_a+\imath z_b$ to be calculated as $z_a=\frac{1}{2}(z+z^*)$ and $z_b=1/2\imath(z-z^*)$. The necessity to use both z and z^* to describe the elements of the corresponding bivariate signal in $\mathbb{R}^2$ is used as a basis for the augmented complex statistics, where the 'augmented' basis vector is $z^a=[z,\ z^*]^T$. However, the quaternion domain does not permit such convenient manipulation, and a correspondence between the elements of a quadrivariate vector in $\mathbb{R}^4$ and the elements of a quaternion valued variable in $\mathbb{H}$ is not straightforward to establish. To deal with this issue, we employ the three perpendicular quaternion involutions (self-inverse mappings), given by

$$\begin{aligned} q^\imath &= -\imath q\imath = q_a+\imath q_b-\jmath q_c-\kappa q_d \\ q^\jmath &= -\jmath q\jmath = q_a-\imath q_b+\jmath q_c-\kappa q_d \\ q^\kappa &= -\kappa q\kappa = q_a-\imath q_b-\jmath q_c+\kappa q_d \end{aligned} \tag{6}$$

whose conjugates $q^{\imath*}$, $q^{\jmath*}$ and $q^{\kappa*}$are defined as

$$\begin{aligned} q^{\imath*} &= q_a-\imath q_b+\jmath q_c+\kappa q_d \\ q^{\jmath*} &= q_a+\imath q_b-\jmath q_c+\kappa q_d \\ q^{\kappa*} &= q_a+\imath q_b+\jmath q_c-\kappa q_d \end{aligned} \tag{7}$$

The four components of the quaternion q can now be expressed as [18]

$$\begin{aligned} &q_a=\frac{1}{2}(q+q^*) \quad q_b=\frac{1}{2\imath}(q-q^{\imath*}) \\ &q_c=\frac{1}{2\jmath}(q-q^{\jmath*}) \quad q_d=\frac{1}{2\kappa}(q-q^{\kappa*}) \end{aligned} \tag{8}$$

Notice that the quaternion conjugate operator $(\cdot)^*$ is also an involution, that is

$$q^*=\frac{1}{2}(q^\imath+q^\jmath+q^\kappa-q) \tag{9}$$

In analogy to the complex domain, to make the augmented statistics in $\mathbb{H}$ suitable for dealing with both second order circular and noncircular signals, we need to first establish a one-to-one correspondence between the components of a quadrivariate real variable (q_a, q_b, q_c, q_d) and its quaternionic counterpart $q=q_a+\imath q_b+\jmath q_c+\kappa q_d$. To this end, following on [19] (see pp. 118–119), the augmented quaternion signal $\mathbf{q}^a \in \mathbb{H}^{4N\times1}$ is next considered, and is related to real valued quadrivariate vectors in $\mathbb{R}^{4N\times1}$. Based on (6)–(9), for convenient manipulation of the components of quaternion variables, we can use a combination[2] of $\{q,q^*,q^\imath,q^\jmath,q^\kappa\}$ to define the augmented quaternion vector $\mathbf{q}^a=[\mathbf{q}^T\ \mathbf{q}^{\imath T}\ \mathbf{q}^{\jmath T}\ \mathbf{q}^{\kappa T}]^T$, which is related to its vectorial counterpart $\mathbf{q}^r \in \mathbb{R}^{4N}$ as

$$\mathbf{q}^a=\mathbf{A}\mathbf{q}^r$$

$$\begin{bmatrix} \mathbf{q} \\ \mathbf{q}^\imath \\ \mathbf{q}^\jmath \\ \mathbf{q}^\kappa \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \imath\mathbf{I} & \jmath\mathbf{I} & \kappa\mathbf{I} \\ \mathbf{I} & \imath\mathbf{I} & -\jmath\mathbf{I} & -\kappa\mathbf{I} \\ \mathbf{I} & -\imath\mathbf{I} & \jmath\mathbf{I} & -\kappa\mathbf{I} \\ \mathbf{I} & -\imath\mathbf{I} & -\jmath\mathbf{I} & \kappa\mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{q}_a \\ \mathbf{q}_b \\ \mathbf{q}_c \\ \mathbf{q}_d \end{bmatrix} \tag{10}$$

where $\mathbf{I} \in \mathbb{R}^{N\times N}$ is the identity matrix, and $\mathbf{q}=[q_1,q_2,\ldots,q_N]^T \in \mathbb{H}^{N\times1}$; similar description also applies to $\mathbf{q}^\imath,\mathbf{q}^\jmath,\mathbf{q}^\kappa \in \mathbb{H}^{N\times1}$, and $\mathbf{q}_a$, $\mathbf{q}_b$, $\mathbf{q}_c$ and $\mathbf{q}_d \in \mathbb{R}^{N\times1}$. The $4N\times4N$ matrix $\mathbf{A}$ provides an invertible mapping between the augmented quaternion valued signal $\mathbf{q}^a \in \mathbb{H}^{4N\times1}$ and the quadrivariate 'composite' real valued vector $\mathbf{q}^r=[\mathbf{q}_a^T\ \mathbf{q}_b^T\ \mathbf{q}_c^T\ \mathbf{q}_d^T]^T \in \mathbb{R}^{4N\times1}$. The inverse mapping from $\mathbb{R}^{4N}$ to $\mathbb{H}^{4N}$ is performed using

$$\mathbf{A}^{-1}=\frac{1}{4}\mathbf{A}^H \tag{11}$$

thus yielding $\mathbf{q}^r=\frac{1}{4}\mathbf{A}^H\mathbf{q}^a$. The determinant of $\mathbf{A}$ can be calculated as a product of its singular values, and so, e.g. for N=1, the singular values[3] of $\mathbf{A}$ are {2,2,2,2} and thus det($\mathbf{A}$)=16. For an arbitrary vector length N, the determinant of matrix $\mathbf{A}$ therefore becomes

$$\det(\mathbf{A})=16^N \tag{12}$$

The basis $\{q,q^\imath,q^\jmath,q^\kappa\}$ in (10) has been selected, so that the matrix $\mathbf{A}$ satisfies (11), thus facilitating its algebraic manipulation. In the sequel, we show that due to the

---

[2] Any four of $\{q,q^*,q^\imath,q^\jmath,q^\kappa\}$ or their conjugates can be used with the same effect.

[3] See for instance the work of Zhang [20]; the computation of singular values can be performed conveniently using the quaternion MATLAB toolbox [21].

relation (9), any other combination of four elements of $\{q,q^*,q^\imath,q^\jmath,q^\kappa\}$, for instance $\{q,q^*,q^{\imath*},q^{\jmath*}\}$ also represents a valid basis, however, these bases do not guarantee the elegant inverse property as in (11).

3. Augmented quaternion statistics

3.1. Preliminaries

The standard covariance matrix $\mathcal{C}_{\mathbf{qq}}$ of a quaternion random vector $\mathbf{q}=[q_1,\ldots,q_N]^T$ is given by

$$\mathcal{C}_{\mathbf{qq}}=E\{\mathbf{qq}^H\} \tag{13}$$

and its structure is detailed in Table 3. Observe that the real and imaginary parts of $\mathcal{C}_{\mathbf{qq}}$ are linear functions of the real-valued covariance and cross-covariance matrices of the component vectors $\mathbf{q}_a$, $\mathbf{q}_b$, $\mathbf{q}_c$ and $\mathbf{q}_d \in \mathbb{R}^{N\times 1}$. From Table 3, the cross-correlation matrices have special symmetry properties, e.g. $\mathcal{C}_{\mathbf{q}_b\mathbf{q}_a}=\mathcal{C}^T_{\mathbf{q}_a\mathbf{q}_b}$, and it thus becomes apparent that $\Re\{\mathcal{C}_{\mathbf{qq}}\}$ is symmetric, whereas $\Im\{\mathcal{C}_{\mathbf{qq}}\}$ is skew-symmetric, thus explaining the Hermitian property of $\mathcal{C}_{\mathbf{qq}}$.

Based on (8) and (10), the real-valued componentwise correlation matrices of the components $\mathbf{q}_a$, $\mathbf{q}_b$, $\mathbf{q}_c$ and $\mathbf{q}_d$ cannot be estimated from the quaternion-valued covariance matrix $\mathcal{C}_{\mathbf{q}}$ alone. Hence, second order information within the quaternion-valued vector $\mathbf{q}$ cannot be characterised completely by the covariance matrix, and complementary correlation matrices: the $\imath$-covariance $\mathcal{C}_{\mathbf{q}^\imath}$, the $\jmath$-covariance $\mathcal{C}_{\mathbf{q}^\jmath}$, and the κ-covariance $\mathcal{C}_{\mathbf{q}^\kappa}$ need to be used. Based on $q^{\imath H}=[q_1^{\imath*},\ldots,q_N^{\imath*}]^T$, $q^{\jmath H}=[q_1^{\jmath*},\ldots,q_N^{\jmath*}]^T$ and $q^{\kappa H}=[q_1^{\kappa*},\ldots,q_N^{\kappa*}]^T$, these complementary covariance matrices are defined in (14)–(16), and will be used to augment the information within the covariance.

$$\mathcal{C}_{\mathbf{q}^\imath}=E\{\mathbf{qq}^{\imath H}\} \tag{14}$$

$$\mathcal{C}_{\mathbf{q}^\jmath}=E\{\mathbf{qq}^{\jmath H}\} \tag{15}$$

$$\mathcal{C}_{\mathbf{q}^\kappa}=E\{\mathbf{qq}^{\kappa H}\} \tag{16}$$

The structures of the real and imaginary parts of $\mathcal{C}_{\mathbf{q}^\imath}$, $\mathcal{C}_{\mathbf{q}^\jmath}$, and $\mathcal{C}_{\mathbf{q}^\kappa}$ are given in Tables 3 and 4.

Observe that, e.g. all the components of the $\imath$-covariance $\mathcal{C}_{\mathbf{q}^\imath}$ are symmetric, except for the $\imath$-component $\Im_\imath\{\mathcal{C}_{\mathbf{q}^\imath}\}$ which is skew-symmetric, giving rise to its $\imath$-Hermitian property. Similarly, the $\jmath$-covariance $\mathcal{C}_{\mathbf{q}^\jmath}$ and the κ-covariance $\mathcal{C}_{\mathbf{q}^\kappa}$ are, respectively, $\jmath$-Hermitian and κ-Hermitian, that is

$$\mathcal{C}_{\mathbf{q}^\imath}=\mathcal{C}^{\imath H}_{\mathbf{q}^\imath}$$

$$\mathcal{C}_{\mathbf{q}^\jmath}=\mathcal{C}^{\jmath H}_{\mathbf{q}^\jmath}$$

$$\mathcal{C}_{\mathbf{q}^\kappa}=\mathcal{C}^{\kappa H}_{\mathbf{q}^\kappa} \tag{17}$$

These properties are specific to the quaternion domain; they do not arise in the statistics of complex valued random variables [9,5], thus illustrating a non-trivial nature of the extension of augmented complex statistics to its quaternion counterpart.

3.2. Duality between the quaternionic and quadrivariate statistics

To justify the need for augmented complex statistics, where the covariance matrix alone is not adequate to describe general complex-valued random vectors[4] $\mathbf{z}=\mathbf{z}_a+\imath\mathbf{z}_b$, it was shown that the complete correlation structure, catering for both proper and improper signals, can be obtained if the covariance matrices of the 'composite' bivariate real vector can be computed from their complex valued counterparts (see pp. 118–119 [19]). In Section 2, we have already shown that components of a composite quadrivariate real variable corresponding to the quaternion variable q cannot be completely expressed based on only q and q^*, and to be able to introduce augmented statistics in $\mathbb{H}$, we need to consider an augmented basis also comprising the involutions $q^\imath$,$q^\jmath$, and q^κ. Following on these results, we show that a complete second order statistical description in $\mathbb{H}$ can be obtained, provided that the quadrivariate real-valued correlation matrices of each single component $\mathbf{q}_a$, $\mathbf{q}_b$, $\mathbf{q}_c$ and $\mathbf{q}_d$ of the quaternion random vector $\mathbf{q}$ can be expressed in terms of the quaternion-valued covariance and the complementary covariance matrices,[5] that is

$$\mathcal{C}_{\mathbf{q}_a}=\frac{1}{4}\Re\{\mathcal{C}_{\mathbf{qq}}+\mathcal{C}_{\mathbf{q}^\imath}+\mathcal{C}_{\mathbf{q}^\jmath}+\mathcal{C}_{\mathbf{q}^\kappa}\}$$
$$\mathcal{C}_{\mathbf{q}_b}=\frac{1}{4}\Re\{\mathcal{C}_{\mathbf{qq}}+\mathcal{C}_{\mathbf{q}^\imath}-\mathcal{C}_{\mathbf{q}^\jmath}-\mathcal{C}_{\mathbf{q}^\kappa}\}$$

$$\mathcal{C}_{\mathbf{q}_c}=\frac{1}{4}\Re\{\mathcal{C}_{\mathbf{qq}}-\mathcal{C}_{\mathbf{q}^\imath}+\mathcal{C}_{\mathbf{q}^\jmath}-\mathcal{C}_{\mathbf{q}^\kappa}\}$$
$$\mathcal{C}_{\mathbf{q}_d}=\frac{1}{4}\Re\{\mathcal{C}_{\mathbf{qq}}-\mathcal{C}_{\mathbf{q}^\imath}-\mathcal{C}_{\mathbf{q}^\jmath}+\mathcal{C}_{\mathbf{q}^\kappa}\}$$

$$\mathcal{C}_{\mathbf{q}_b\mathbf{q}_a}=\frac{1}{4}\Im_\imath\{\mathcal{C}_{\mathbf{qq}}+\mathcal{C}_{\mathbf{q}^\imath}+\mathcal{C}_{\mathbf{q}^\jmath}+\mathcal{C}_{\mathbf{q}^\kappa}\}$$
$$\mathcal{C}_{\mathbf{q}_c\mathbf{q}_a}=\frac{1}{4}\Im_\jmath\{\mathcal{C}_{\mathbf{qq}}+\mathcal{C}_{\mathbf{q}^\imath}+\mathcal{C}_{\mathbf{q}^\jmath}+\mathcal{C}_{\mathbf{q}^\kappa}\}$$

$$\mathcal{C}_{\mathbf{q}_d\mathbf{q}_a}=\frac{1}{4}\Im_\kappa\{\mathcal{C}_{\mathbf{qq}}+\mathcal{C}_{\mathbf{q}^\imath}+\mathcal{C}_{\mathbf{q}^\jmath}+\mathcal{C}_{\mathbf{q}^\kappa}\}$$
$$\mathcal{C}_{\mathbf{q}_c\mathbf{q}_b}=\frac{1}{4}\Im_\kappa\{\mathcal{C}_{\mathbf{qq}}+\mathcal{C}_{\mathbf{q}^\imath}-\mathcal{C}_{\mathbf{q}^\jmath}-\mathcal{C}_{\mathbf{q}^\kappa}\}$$

$$\mathcal{C}_{\mathbf{q}_d\mathbf{q}_b}=-\frac{1}{4}\Im_\jmath\{\mathcal{C}_{\mathbf{qq}}+\mathcal{C}_{\mathbf{q}^\imath}-\mathcal{C}_{\mathbf{q}^\jmath}-\mathcal{C}_{\mathbf{q}^\kappa}\}$$

[4] In the complex domain, both the covariance $\mathcal{C}_{\mathbf{z}}=E\{\mathbf{zz}^H\}$ and the pseudocovariance $\mathcal{P}_{\mathbf{z}}=E\{\mathbf{zz}^T\}$ should be used, that is

$$\mathcal{C}_{\mathbf{z}_a}=\frac{1}{2}\Re\{\mathcal{C}_{\mathbf{z}}+\mathcal{P}_{\mathbf{z}}\} \quad \mathcal{C}_{\mathbf{z}_b}=\frac{1}{2}\Re\{\mathcal{C}_{\mathbf{z}}-\mathcal{P}_{\mathbf{z}}\}$$

$$\mathcal{C}_{\mathbf{z}_a\mathbf{z}_b}=\frac{1}{2}\Im_\imath\{\mathcal{P}_{\mathbf{z}}-\mathcal{C}_{\mathbf{z}}\} \quad \mathcal{C}_{\mathbf{z}_b\mathbf{z}_a}=\mathcal{C}^T_{\mathbf{z}_a\mathbf{z}_b}$$

where $\mathcal{C}_{\mathbf{z}_a}$ and $\mathcal{C}_{\mathbf{z}_b}$ are, respectively, the componentwise covariance matrices of the real part $\mathbf{z}_a$ and the imaginary part $\mathbf{z}_b$, whereas $\mathcal{C}_{\mathbf{z}_a\mathbf{z}_b}$ and $\mathcal{C}_{\mathbf{z}_b\mathbf{z}_a}$ denote the cross-covariance matrices.

[5] If a different basis, e.g. $\{q,q^*,q^{\imath*},q^{\jmath*}\}$ is chosen, the full description of the second order statistics is still obtained; this also applies to any other combination of quadruples taken from $\{q,q^*,q^\imath,q^\jmath,q^\kappa\}$.

$$\mathcal{C}_{\mathbf{q}_d\mathbf{q}_c} = \frac{1}{4}\Im_\imath\{\mathcal{C}_{\mathbf{qq}} - \mathcal{C}_{\mathbf{q}^\imath} + \mathcal{C}_{\mathbf{q}^\jmath} - \mathcal{C}_{\mathbf{q}^\kappa}\} \tag{18}$$

The augmented quaternion-valued covariance matrix of an augmented random vector $\mathbf{q}^a = [\mathbf{q}^T \ \mathbf{q}^{\imath T} \ \mathbf{q}^{\jmath T} \ \mathbf{q}^{\kappa T}]^T$ (see also (10)), then becomes

$$\mathcal{C}^a_{\mathbf{q}} = E\{\mathbf{q}^a\mathbf{q}^{aH}\} = \begin{bmatrix} \mathcal{C}_{\mathbf{qq}} & \mathcal{C}_{\mathbf{q}^\imath} & \mathcal{C}_{\mathbf{q}^\jmath} & \mathcal{C}_{\mathbf{q}^\kappa} \\ \mathcal{C}^H_{\mathbf{q}^\imath} & \mathcal{C}_{\mathbf{q}^\imath\mathbf{q}^\imath} & \mathcal{C}_{\mathbf{q}^\imath\mathbf{q}^\jmath} & \mathcal{C}_{\mathbf{q}^\imath\mathbf{q}^\kappa} \\ \mathcal{C}^H_{\mathbf{q}^\jmath} & \mathcal{C}_{\mathbf{q}^\jmath\mathbf{q}^\imath} & \mathcal{C}_{\mathbf{q}^\jmath\mathbf{q}^\jmath} & \mathcal{C}_{\mathbf{q}^\jmath\mathbf{q}^\kappa} \\ \mathcal{C}^H_{\mathbf{q}^\kappa} & \mathcal{C}_{\mathbf{q}^\kappa\mathbf{q}^\imath} & \mathcal{C}_{\mathbf{q}^\kappa\mathbf{q}^\jmath} & \mathcal{C}_{\mathbf{q}^\kappa\mathbf{q}^\kappa} \end{bmatrix} \tag{19}$$

where the submatrices within (19) are calculated according to (13)–(16), and $\mathcal{C}_{\boldsymbol{\alpha\beta}} = E\{\boldsymbol{\alpha\beta}^H\}\forall\boldsymbol{\alpha},\boldsymbol{\beta} \in \{\mathbf{q},\mathbf{q}^\imath,\mathbf{q}^\jmath,\mathbf{q}^\kappa\}$. To verify that the augmented covariance matrix in (19) provides a complete second order statistical description, we need to show that it permits a static invertible one-to-one mapping (isomorphism) with the corresponding real valued quadrivariate covariance matrix $\mathcal{C}_R$, defined as

$$\mathcal{C}_R = E\{\mathbf{q}^r\mathbf{q}^{rT}\} = \begin{bmatrix} \mathcal{C}_{\mathbf{q}_a} & \mathcal{C}_{\mathbf{q}_a\mathbf{q}_b} & \mathcal{C}_{\mathbf{q}_a\mathbf{q}_c} & \mathcal{C}_{\mathbf{q}_a\mathbf{q}_d} \\ \mathcal{C}_{\mathbf{q}_b\mathbf{q}_a} & \mathcal{C}_{\mathbf{q}_b} & \mathcal{C}_{\mathbf{q}_b\mathbf{q}_c} & \mathcal{C}_{\mathbf{q}_b\mathbf{q}_d} \\ \mathcal{C}_{\mathbf{q}_c\mathbf{q}_a} & \mathcal{C}_{\mathbf{q}_c\mathbf{q}_b} & \mathcal{C}_{\mathbf{q}_c} & \mathcal{C}_{\mathbf{q}_c\mathbf{q}_d} \\ \mathcal{C}_{\mathbf{q}_d\mathbf{q}_a} & \mathcal{C}_{\mathbf{q}_d\mathbf{q}_b} & \mathcal{C}_{\mathbf{q}_d\mathbf{q}_c} & \mathcal{C}_{\mathbf{q}_d} \end{bmatrix} \tag{20}$$

Based on the relationship between the augmented quaternion-valued vector $\mathbf{q}^a$ and the corresponding real valued 'composite' vector $\mathbf{q}^r$ in (10), and since from (11) $\mathbf{q}^r = \mathbf{A}^{-1}\mathbf{q}^a = \frac{1}{4}\mathbf{A}^H\mathbf{q}^a$, the real valued quadrivariate covariance matrix can be expressed in terms of the augmented quaternion valued covariance matrix in (19) as

$$\mathcal{C}_R = \mathbf{A}^{-1}\mathcal{C}^a_{\mathbf{q}}\mathbf{A}^{-H} = \frac{1}{16}\mathbf{A}^H\mathcal{C}^a_{\mathbf{q}}\mathbf{A} \tag{21}$$

where $\mathbf{A}^{-H} = (\mathbf{A}^{-1})^H$. The so-introduced augmented quaternion statistics provides a general tool for unified second order modelling, of both proper and improper quaternion random processes.

3.3. Second order stationarity

Recall that a real-valued quadrivariate variable is wide-sense stationary if all its four components are jointly wide-sense stationary [22]. Thus, to define stationarity in $\mathbb{H}$, it is sufficient to consider the correlation matrices $\mathcal{C}_{\mathbf{qq}}$, $\mathcal{C}_{\mathbf{q}^\imath}$, $\mathcal{C}_{\mathbf{q}^\jmath}$, and $\mathcal{C}_{\mathbf{q}^\kappa}$, as from (18) they provide the complete description of quaternion second order statistics.[6]

We can now state that a quaternion-valued random process $q(n)$ is wide-sense stationary, provided

1. The mean is constant, that is, $\mu = E\{q(n)\} = K, \quad \forall n$
2. The covariance and its complementary matrices are function of only the lag τ, that is

$$\mathcal{C}_{\mathbf{qq}}(\tau) = E\{\mathbf{q}(n)\mathbf{q}^H(n+\tau)\}$$

$$\mathcal{C}_{\mathbf{q}^\imath}(\tau) = E\{\mathbf{q}(n)\mathbf{q}^{\imath H}(n+\tau)\}$$

$$\mathcal{C}_{\mathbf{q}^\jmath}(\tau) = E\{\mathbf{q}(n)\mathbf{q}^{\jmath H}(n+\tau)\}$$

$$\mathcal{C}_{\mathbf{q}^\kappa}(\tau) = E\{\mathbf{q}(n)\mathbf{q}^{\kappa H}(n+\tau)\}$$

3. All entries of the covariance matrix $\mathcal{C}_{\mathbf{qq}} = E\{\mathbf{q}(n)\mathbf{q}^H(n)\}$ are finite.

Notice that the wide-sense stationarity of $\mathcal{C}_{\mathbf{qq}}$, $\mathcal{C}_{\mathbf{q}^\imath}$, $\mathcal{C}_{\mathbf{q}^\jmath}$, and $\mathcal{C}_{\mathbf{q}^\kappa}$ also implies the wide-sense stationarity of the corresponding real-valued quadrivariate cross-covariances such as $\mathcal{C}_{\mathbf{q}_b\mathbf{q}_a}$ in (18).

4. Second order circularity in $\mathbb{H}$ and $\mathbb{Q}$-properness

The notion of second order circularity (or properness) in the complex domain refers to complex-valued variables having rotation-invariant probability distributions, and consequently a vanishing pseudocovariance [23]. The two conditions imposed on a complex variable $z = z_a + \imath z_b$ to be proper ($\mathbb{C}$-proper) are therefore

$$\sigma^2_{z_a} = \sigma^2_{z_b}$$

$$E\{z_a z_b\} = 0 \tag{22}$$

that is, the real and imaginary part are of equal power and not correlated, which amounts to a vanishing pseudocovariance matrix $\mathcal{P} = E\{\mathbf{zz}^T\}$.

By continuity, a quaternion-valued second order circular ($\mathbb{Q}$-proper) variable should satisfy the two conditions in (22) of a $\mathbb{C}$-proper variable for the six pairs of axes: $\{1,\imath\},\{1,\jmath\},\{1,\kappa\},\{\imath,\jmath\},\{\kappa,\jmath\}$ and $\{\kappa,\imath\}$, where '1' represents the real axis and $\imath,\jmath,\kappa$ denote the corresponding imaginary axes. In other words, the probability distribution of a $\mathbb{Q}$-proper variable is rotation-invariant with respect to all these six pairs of axes, leading to the properties of a $\mathbb{Q}$-proper variable summarised in Table 2 [12].

The first property, P1, states that all the four components of a $\mathbb{Q}$-proper variable have equal powers. The property P2 implies that the components of q are uncorrelated. Property P3 indicates that the pseudocovariance matrix does not vanish for $\mathbb{Q}$-proper signals, in contrast to the complex case. Finally, the fourth property illustrates that the covariance of a quaternion variable is a sum of the covariances of the process components. Notice that properties P1 and P2 imply properties P3 and P4.

For quaternion random vectors $\mathbf{q}^\vartheta$ and $\mathbf{q}^\upsilon$ to be jointly proper, the composite random vector having $\mathbf{q}^\vartheta$ and $\mathbf{q}^\upsilon$ as subvectors[7] also has to be proper. In addition, to guarantee joint $\mathbb{Q}$-properness, each element of the vectors $\mathbf{q}^\vartheta$ and $\mathbf{q}^\upsilon$ should satisfy properties P1 and P2 in Table 2, and they should be uncorrelated (their joint $\imath$-$\jmath$-κ-covariance matrices vanish). This is discussed in more detail below.

4.1. Augmented statistics and $\mathbb{Q}$-properness

Following on the complex properness (as detailed in Section IIIA of [6]), we shall now consider $\mathbb{Q}$-proper random vectors $\mathbf{q} = [q_1, q_2, \ldots, q_N]^T \in \mathbb{H}^{N\times 1}$. As shown in [13], by

[6] Notice that if another basis was chosen (for instance $\{q,q^*,q^{\imath *},q^{\jmath *}\}$), then another set of covariance matrices ($\mathcal{C}_{\mathbf{qq}}, \mathcal{P}_{\mathbf{q}} = E\{\mathbf{qq}^T\}$, $\mathcal{P}^\imath_{\mathbf{q}} = E\{\mathbf{qq}^{\imath T}\}$, $\mathcal{P}^\jmath_{\mathbf{q}} = E\{\mathbf{qq}^{\jmath T}\}$) would be employed to define stationarity.

[7] Any subvector of a proper random vector is also proper.

Table 2
Properties of a $\mathbb{Q}$-proper random variable.

Property	Mathematical description	
P1	$E\{q_\delta^2\}=\sigma^2$	$\forall\delta=a,b,c,d$
P2	$E\{q_\delta q_\varepsilon\}=0$	$\forall\delta,\varepsilon=a,b,c,d$ and $\delta\neq\varepsilon$
P3	$E\{qq\}=-2E\{q_\delta^2\}=-2\sigma^2$	$\forall\delta=a,b,c,d$
P4	$E\{\lvert q\rvert^2\}=4E\{q_\delta^2\}=4\sigma^2$	$\forall\delta=a,b,c,d$

analogy to single quaternion variables, $\mathbb{Q}$-properness implies that the quaternion vector $\mathbf{q}$ is not correlated with its vector involutions $\mathbf{q}^\imath$, $\mathbf{q}^\jmath$, $\mathbf{q}^\kappa$, that is,[8]

$$E\{\mathbf{q}\mathbf{q}^{\imath H}\}=\mathbf{0}\quad E\{\mathbf{q}\mathbf{q}^{\jmath H}\}=\mathbf{0}\quad E\{\mathbf{q}\mathbf{q}^{\kappa H}\}=\mathbf{0} \tag{23}$$

and thus has vanishing complementary covariance matrices (specified in Tables 3 and 4). A $\mathbb{Q}$-proper random vector is invariant under a linear or affine transformation, similar to the complex case (see Lemma 3 [6]). Based on (17), it is straightforward to show the condition of vanishing $\imath$-covariance matrix, $\mathcal{C}_{\mathbf{q}^\imath}=\mathbf{0}$, in (14) is equivalent to the conditions

$$\mathcal{C}_{\mathbf{q}_b\mathbf{q}_a}=\mathcal{C}^T_{\mathbf{q}_b\mathbf{q}_a}\quad \mathcal{C}_{\mathbf{q}_c\mathbf{q}_d}=\mathcal{C}^T_{\mathbf{q}_c\mathbf{q}_d}$$

$$\mathcal{C}_{\mathbf{q}_c\mathbf{q}_a}=-\mathcal{C}^T_{\mathbf{q}_c\mathbf{q}_a}\quad \mathcal{C}_{\mathbf{q}_d\mathbf{q}_b}=-\mathcal{C}^T_{\mathbf{q}_d\mathbf{q}_b}$$

$$\mathcal{C}_{\mathbf{q}_d\mathbf{q}_a}=-\mathcal{C}^T_{\mathbf{q}_d\mathbf{q}_a}\quad \mathcal{C}_{\mathbf{q}_b\mathbf{q}_c}=-\mathcal{C}^T_{\mathbf{q}_b\mathbf{q}_c} \tag{24}$$

and the vanishing $\jmath$-covariance matrix, $\mathcal{C}_{\mathbf{q}^\jmath}=\mathbf{0}$, in (15) implies

$$\mathcal{C}_{\mathbf{q}_b\mathbf{q}_a}=-\mathcal{C}^T_{\mathbf{q}_b\mathbf{q}_a}\quad \mathcal{C}_{\mathbf{q}_c\mathbf{q}_d}=-\mathcal{C}^T_{\mathbf{q}_c\mathbf{q}_d}$$

$$\mathcal{C}_{\mathbf{q}_c\mathbf{q}_a}=\mathcal{C}^T_{\mathbf{q}_c\mathbf{q}_a}\quad \mathcal{C}_{\mathbf{q}_d\mathbf{q}_b}=\mathcal{C}^T_{\mathbf{q}_d\mathbf{q}_b}$$

$$\mathcal{C}_{\mathbf{q}_d\mathbf{q}_a}=-\mathcal{C}^T_{\mathbf{q}_d\mathbf{q}_a}\quad \mathcal{C}_{\mathbf{q}_b\mathbf{q}_c}=-\mathcal{C}^T_{\mathbf{q}_b\mathbf{q}_c} \tag{25}$$

whereas, the vanishing the κ-covariance matrix, $\mathcal{C}_{\mathbf{q}^\kappa}=\mathbf{0}$, in (16) yields

$$\mathcal{C}_{\mathbf{q}_b\mathbf{q}_a}=-\mathcal{C}^T_{\mathbf{q}_b\mathbf{q}_a}\quad \mathcal{C}_{\mathbf{q}_c\mathbf{q}_d}=-\mathcal{C}^T_{\mathbf{q}_c\mathbf{q}_d}$$

$$\mathcal{C}_{\mathbf{q}_c\mathbf{q}_a}=-\mathcal{C}^T_{\mathbf{q}_c\mathbf{q}_a}\quad \mathcal{C}_{\mathbf{q}_d\mathbf{q}_b}=-\mathcal{C}^T_{\mathbf{q}_d\mathbf{q}_b}$$

$$\mathcal{C}_{\mathbf{q}_d\mathbf{q}_a}=\mathcal{C}^T_{\mathbf{q}_d\mathbf{q}_a}\quad \mathcal{C}_{\mathbf{q}_b\mathbf{q}_c}=\mathcal{C}^T_{\mathbf{q}_b\mathbf{q}_c} \tag{26}$$

Since from above $\mathcal{C}_{\mathbf{q}_b\mathbf{q}_a}=-\mathcal{C}^T_{\mathbf{q}_b\mathbf{q}_a}$ for (25) and (26), whereas $\mathcal{C}_{\mathbf{q}_b\mathbf{q}_a}=\mathcal{C}^T_{\mathbf{q}_b\mathbf{q}_a}$ for (24), this means that $\mathcal{C}_{\mathbf{q}_b\mathbf{q}_a}=\mathbf{0}$. Similar observations can be made for the other componentwise real-valued cross-correlation matrices, meaning that for a $\mathbb{Q}$-proper signal, all the real-valued cross-correlation matrices of the components $\mathbf{q}_a$, $\mathbf{q}_b$, $\mathbf{q}_c$, and $\mathbf{q}_d$ need to vanish. This, in turn, implies that

- All the four individual components of a quaternion variable q_ℓ are uncorrelated (property P2 in Table 2).
- The components of q_ℓ and q_ϱ are uncorrelated for $\ell\neq\varrho$ (in contrast to the complex case [6,23]).

Table 3
Structures of the quaternion-valued covariance matrices in terms of their quadrivariate real-valued counterparts.

Covariance matrix	$\mathcal{C}_{\mathbf{q}\mathbf{q}}=E\{\mathbf{q}\mathbf{q}^H\}$	$\mathcal{C}_{\mathbf{q}^\imath}=E\{\mathbf{q}\mathbf{q}^{\imath H}\}$
$\Re\{\cdot\}$	$\mathcal{C}_{\mathbf{q}_a}+\mathcal{C}_{\mathbf{q}_b}+\mathcal{C}_{\mathbf{q}_c}+\mathcal{C}_{\mathbf{q}_d}$	$\mathcal{C}_{\mathbf{q}_a}+\mathcal{C}_{\mathbf{q}_b}-\mathcal{C}_{\mathbf{q}_c}-\mathcal{C}_{\mathbf{q}_d}$
$\Im_\imath\{\cdot\}$	$\mathcal{C}_{\mathbf{q}_b\mathbf{q}_a}-\mathcal{C}_{\mathbf{q}_a\mathbf{q}_b}+\mathcal{C}_{\mathbf{q}_d\mathbf{q}_c}-\mathcal{C}_{\mathbf{q}_c\mathbf{q}_d}$	$\mathcal{C}_{\mathbf{q}_b\mathbf{q}_a}-\mathcal{C}_{\mathbf{q}_a\mathbf{q}_b}+\mathcal{C}_{\mathbf{q}_c\mathbf{q}_d}-\mathcal{C}_{\mathbf{q}_d\mathbf{q}_c}$
$\Im_\jmath\{\cdot\}$	$\mathcal{C}_{\mathbf{q}_c\mathbf{q}_a}-\mathcal{C}_{\mathbf{q}_a\mathbf{q}_c}+\mathcal{C}_{\mathbf{q}_b\mathbf{q}_d}-\mathcal{C}_{\mathbf{q}_d\mathbf{q}_b}$	$\mathcal{C}_{\mathbf{q}_a\mathbf{q}_c}+\mathcal{C}_{\mathbf{q}_c\mathbf{q}_a}-\mathcal{C}_{\mathbf{q}_d\mathbf{q}_b}-\mathcal{C}_{\mathbf{q}_b\mathbf{q}_d}$
$\Im_\kappa\{\cdot\}$	$\mathcal{C}_{\mathbf{q}_d\mathbf{q}_a}-\mathcal{C}_{\mathbf{q}_a\mathbf{q}_d}+\mathcal{C}_{\mathbf{q}_c\mathbf{q}_b}-\mathcal{C}_{\mathbf{q}_b\mathbf{q}_c}$	$\mathcal{C}_{\mathbf{q}_d\mathbf{q}_a}+\mathcal{C}_{\mathbf{q}_a\mathbf{q}_d}+\mathcal{C}_{\mathbf{q}_b\mathbf{q}_d}+\mathcal{C}_{\mathbf{q}_d\mathbf{q}_b}$

Table 4
Structures of the quaternion-valued covariance matrices in terms of their quadrivariate real-valued counterparts.

Covariance matrix	$\mathcal{C}_{\mathbf{q}^\jmath}=E\{\mathbf{q}\mathbf{q}^{\jmath H}\}$	$\mathcal{C}_{\mathbf{q}^\kappa}=E\{\mathbf{q}\mathbf{q}^{\kappa H}\}$
$\Re\{\cdot\}$	$\mathcal{C}_{\mathbf{q}_a}-\mathcal{C}_{\mathbf{q}_b}+\mathcal{C}_{\mathbf{q}_c}-\mathcal{C}_{\mathbf{q}_d}$	$\mathcal{C}_{\mathbf{q}_a}-\mathcal{C}_{\mathbf{q}_b}-\mathcal{C}_{\mathbf{q}_c}+\mathcal{C}_{\mathbf{q}_d}$
$\Im_\imath\{\cdot\}$	$\mathcal{C}_{\mathbf{q}_b\mathbf{q}_a}+\mathcal{C}_{\mathbf{q}_a\mathbf{q}_b}+\mathcal{C}_{\mathbf{q}_d\mathbf{q}_c}+\mathcal{C}_{\mathbf{q}_c\mathbf{q}_d}$	$\mathcal{C}_{\mathbf{q}_b\mathbf{q}_a}+\mathcal{C}_{\mathbf{q}_a\mathbf{q}_b}-\mathcal{C}_{\mathbf{q}_c\mathbf{q}_d}-\mathcal{C}_{\mathbf{q}_d\mathbf{q}_c}$
$\Im_\jmath\{\cdot\}$	$\mathcal{C}_{\mathbf{q}_c\mathbf{q}_a}-\mathcal{C}_{\mathbf{q}_a\mathbf{q}_c}+\mathcal{C}_{\mathbf{q}_d\mathbf{q}_b}-\mathcal{C}_{\mathbf{q}_b\mathbf{q}_d}$	$\mathcal{C}_{\mathbf{q}_a\mathbf{q}_c}+\mathcal{C}_{\mathbf{q}_c\mathbf{q}_a}+\mathcal{C}_{\mathbf{q}_d\mathbf{q}_b}+\mathcal{C}_{\mathbf{q}_b\mathbf{q}_d}$
$\Im_\kappa\{\cdot\}$	$\mathcal{C}_{\mathbf{q}_d\mathbf{q}_a}+\mathcal{C}_{\mathbf{q}_a\mathbf{q}_d}-\mathcal{C}_{\mathbf{q}_b\mathbf{q}_c}-\mathcal{C}_{\mathbf{q}_c\mathbf{q}_b}$	$\mathcal{C}_{\mathbf{q}_d\mathbf{q}_a}-\mathcal{C}_{\mathbf{q}_a\mathbf{q}_d}+\mathcal{C}_{\mathbf{q}_b\mathbf{q}_c}-\mathcal{C}_{\mathbf{q}_c\mathbf{q}_b}$

- The augmented covariance matrix $\mathcal{C}^a_{\mathbf{q}}$ of a $\mathbb{Q}$-proper random vector $\mathbf{q}$ is real-valued, positive definite, and symmetric.

For a $\mathbb{Q}$-proper random vector, it follows from properties P2 and P4 in Table 2, that the covariance matrices (13)–(16) are real-valued and diagonal, and the covariance matrix of a $\mathbb{Q}$-proper process is positive definite, leading to a simpler structure of the augmented covariance matrix $\mathcal{C}^a_{\mathbf{q}}$, given by

$$\mathcal{C}^a_{\mathbf{q}}=\begin{bmatrix}\mathcal{C}_{\mathbf{q}\mathbf{q}} & \mathbf{0} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathcal{C}_{\mathbf{q}^\imath\mathbf{q}^\imath} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathcal{C}_{\mathbf{q}^\jmath\mathbf{q}^\jmath} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathcal{C}_{\mathbf{q}^\kappa\mathbf{q}^\kappa}\end{bmatrix}=4\sigma^2\mathbf{I} \tag{27}$$

Notice that the cross-covariance matrices $\mathcal{C}_{\boldsymbol{\alpha\beta}}$ also vanish and the determinant can be readily expressed as $\det(\mathcal{C}^a_{\mathbf{q}})=(4\sigma^2)^{4N}$.

Another kind of properness in $\mathbb{H}$ is the so-called $\mathbb{C}$-properness [13], whereby the variable q is correlated with one of $\{q^\imath,q^\jmath,q^\kappa\}$, but it is not correlated with the remaining two perpendicular involutions. However, when applied to general signals, the concept of $\mathbb{C}$-properness is restrictive, as it implies a rotation of $\pi/2$ in two specific two-dimensional planes. For instance, $\mathbb{C}^\imath$-properness involves only the planes spanned by $\{1,\imath\}$ and $\{\jmath,\kappa\}$, and due to its insufficient number of degrees of freedom, it is not further discussed,[9] highlighting the fact that it is not straightforward to extend the augmented statistics from $\mathbb{C}$ to $\mathbb{H}$. Instead, we perform our analysis directly in $\mathbb{H}$, addressing complete 'augmented' second order statistics for the generality of multivariate quaternion-valued random vectors including $\mathbb{Q}$-proper, $\mathbb{C}$-proper or improper random vectors.

[8] Similarly, for a complex-valued random vector $\mathbf{z}$, $\mathbb{C}$-properness means that $\mathbf{z}$ is not correlated with $\mathbf{z}^*$ in 'complex sense', because $E\{\mathbf{z}(\mathbf{z}^*)^H\}=E\{\mathbf{z}\mathbf{z}^T\}=\mathbf{0}$.

[9] For more details on $\mathbb{C}$-properness for single quaternion variables, see for instance [13].

5. The quaternion widely linear model

To exploit the complete second order statistics of quaternion valued signals in linear mean-squared error (MSE) estimation, we need to consider a filtering model similar to the widely linear model developed for the complex case [24]. To this end, consider the MSE estimator of a real-valued signal y in terms of another observation x, that is, $\hat{y}=E[y|x]$; for zero mean, jointly normal real valued y and x, the solution is a linear model

$$\hat{y}=\mathbf{h}^T\mathbf{x} \tag{28}$$

Standard widely used linear models in $\mathbb{H}$ are assumed to be simple extensions of the real-valued MSE estimator in (28), that is, $\hat{y}=\mathbf{h}^H\mathbf{x}$. Based on the augmented quaternion statistics introduced in Section 3, it is not sufficient to simply replace the real variables in (28) by quaternion variables; instead it is important to realise that the estimator $\hat{y}=E[y|x]$ must be applied to every component (the real and the three imaginary parts) of quaternion variables, that is

$$\hat{y}_\beta=E[y_\beta|x_a,x_b,x_c,x_d], \quad \beta\in\{a,b,c,d\}$$

The conditional estimator of a quaternion variable thus becomes

$$\hat{y}=E[y_a|x_a,x_b,x_c,x_d]+\imath E[y_b|x_a,x_b,x_c,x_d] \\ +\jmath E[y_c|x_a,x_b,x_c,x_d]+\kappa E[y_d|x_a,x_b,x_c,x_d] \tag{29}$$

and the use of involutions in (8) allows us to replace the components $x_a,\ldots,x_d$ in (29) by full quaternions (variable x and its involutions), and thus express the quaternion estimator in the form

$$\hat{y}=E[y|x,x^\imath,x^\jmath,x^\kappa]+\imath E[y^\imath|x,x^\imath,x^\jmath,x^\kappa] \\ +\jmath E[y^\jmath|x,x^\imath,x^\jmath,x^\kappa]+\kappa E[y^\kappa|x,x^\imath,x^\jmath,x^\kappa] \tag{30}$$

The widely linear model for general quaternion signals therefore becomes [25]

$$y=\mathbf{w}^{aH}\mathbf{x}^a=\mathbf{g}^H\mathbf{x}+\mathbf{h}^H\mathbf{x}^\imath+\mathbf{u}^H\mathbf{x}^\jmath+\mathbf{v}^H\mathbf{x}^\kappa \tag{31}$$

The Wiener solution which minimises the MSE $E\{|y-d|^2\}$ based on the QWL model (31) is then given by[10]

$$\mathbf{w}^a=E\{\mathbf{x}^a\mathbf{x}^{aH}\}^{-1}E\{\mathbf{x}^a d^*\} \tag{32}$$

Observe that the QWL Wiener solution has the same general form as the standard solution, but is based on the augmented covariance matrix $\mathcal{C}^a_{\mathbf{x}_a}$ in (19). For $\mathbb{Q}$-proper signals, the augmented covariance matrix simplifies to (27), leading **to** the solution

$$\mathbf{w}^a=\frac{1}{4\sigma^2}E\{\mathbf{x}^a d^*\}$$

The corresponding real-valued quadrivariate model relies on the real-valued covariance matrix in (20) [26], and to provide insight into the duality between the quadrivariate and quaternion second order modelling, we need to establish the relationship between the eigenproperties of $\mathcal{C}_R$ and $\mathcal{C}_{\mathbf{x}_a}$. Based on the roots of $\mathcal{C}_R-\lambda\mathbf{I}=\mathbf{0}$, the relationship (21), and the fact that $\mathbf{I}=\mathbf{A}^{-1}\mathbf{A}=\mathbf{A}^H\mathbf{A}/4$, we obtain

$$\mathcal{C}_R-\lambda\mathbf{I}=\frac{1}{16}\mathbf{A}^H\mathcal{C}^a_{\mathbf{x}_a}\mathbf{A}-\lambda\frac{\mathbf{A}^H\mathbf{A}}{4}=\frac{1}{16}\mathbf{A}^H\underbrace{[\mathcal{C}^a_{\mathbf{x}_a}-4\lambda\mathbf{I}]}_{\aleph}\mathbf{A} \tag{33}$$

Notice that $\aleph=[\mathcal{C}^a_{\mathbf{x}_a}-4\lambda\mathbf{I}]$, and thus the eigenvalues of the augmented quaternion covariance matrix (19) are four times those of the quadrivariate real-valued covariance matrix (20). Thus, for instance, if the quaternion least mean square (QLMS) adaptive filtering algorithm exploits the widely linear model, it will converge four times faster than its multichannel quadrivariate counterpart, for the same learning rate (see also [4]). Another example of the use of augmented statistics in an information-theoretic context is the derivation of the Gaussian distribution and its relevance to the maximum entropy theorem in $\mathbb{H}$, similar to the complex case as described in [6]; this is next discussed in Section 6.

6. A multivariate Gaussian distribution for $\mathbb{Q}$-proper and $\mathbb{Q}$-improper variables

In the complex domain, based on the duality between a complex variable $z=z_a+\imath z_b\in\mathbb{C}$ and a corresponding composite real variable $\omega=[z_a,z_b]\in\mathbb{R}^2$, Van Den Bos proposed a generic complex-valued Gaussian distribution to cater for both $\mathbb{C}$-proper and $\mathbb{C}$-improper processes [7]; this was further elaborated by Picinbono [27]. In the same spirit, we shall address the expressions for probability distributions of both proper and improper processes in $\mathbb{H}$, and will next introduce a generic Gaussian distribution for multivariate quaternion valued random signals.

A quaternion valued random variable is Gaussian if all its components are jointly normal, and their joint Gaussian probability distribution is given by

$$p(\mathbf{q}_a,\mathbf{q}_b,\mathbf{q}_c,\mathbf{q}_d)=\frac{1}{(2\pi)^{2N}\det(\mathcal{C}_R)^{1/2}}\exp\{-\tfrac{1}{2}f(\mathbf{q}_a,\mathbf{q}_b,\mathbf{q}_c,\mathbf{q}_d)\} \tag{34}$$

where

$$f(\mathbf{q}_a,\mathbf{q}_b,\mathbf{q}_c,\mathbf{q}_d)=\mathbf{q}^{rT}\mathcal{C}_R^{-1}\mathbf{q}^r=\mathbf{q}^{rH}\mathcal{C}_R^{-1}\mathbf{q}^r \tag{35}$$

It is assumed that $\mathbf{q}_a$, $\mathbf{q}_b$, $\mathbf{q}_c$ and $\mathbf{q}_d\in\mathbb{R}^{N\times1}$ have zero mean, but this does not restrict the generality of the results. To make the Gaussian distribution (34) cater for both $\mathbb{Q}$-proper and $\mathbb{Q}$-improper signals, we need to involve the augmented quaternion valued vector $\mathbf{q}^a$ in (10). This is achieved based on the investigation of the determinant of the quadrivariate covariance $\mathcal{C}_R$ and the quadratic function in (35).

The determinant of $\mathcal{C}_R$ can be expressed as a function of $\mathcal{C}^a_{\mathbf{q}}$, that is

$$\det(\mathcal{C}_R)=\det(\mathbf{A}^{-1}\mathcal{C}^a_{\mathbf{q}}\mathbf{A}^{-H})=\det(\mathbf{A}^{-1})\det(\mathcal{C}^a_{\mathbf{q}})\det(\mathbf{A}^{-H}) \tag{36}$$

where $\mathbf{A}$ is given in (10). From (12), $\det(\mathbf{A})$=16^N and since $\det(A^{-1})$=$\det(A)^{-1}$, the above expression can be further simplified to

$$\det(\mathcal{C}_R)=(\tfrac{1}{16})^{2N}\det(\mathcal{C}^a_{\mathbf{q}}) \tag{37}$$

The quadratic function (35) can be also expressed as a function of the augmented quaternion-valued random

[10] For clarity the derivation is given in Appendix A.1. Symbol d denotes the desired signal.

Table 5
Summary of complex-valued and quaternion-valued algebra and statistics.

The complex domain $\mathbb{C}$	The quaternion domain $\mathbb{H}$
$\mathbf{z}=\mathbf{z}_a+\imath\mathbf{z}_b$	$\mathbf{q}=\mathbf{q}_a+\imath\mathbf{q}_b+\jmath\mathbf{q}_c+\kappa\mathbf{q}_d$
$\mathbf{z}_a=\frac{1}{2}(\mathbf{z}+\mathbf{z}^*)$	$\mathbf{q}_a=\frac{1}{2}(\mathbf{q}+\mathbf{q}^*)$
$\mathbf{z}_b=\frac{1}{2\imath}(\mathbf{z}+\mathbf{z}^*)$	$\mathbf{q}_b=\frac{1}{2\imath}(\mathbf{q}-\mathbf{q}^{\imath*})$
	$\mathbf{q}_c=\frac{1}{2\jmath}(\mathbf{q}-\mathbf{q}^{\jmath*})$
	$\mathbf{q}_d=\frac{1}{2\kappa}(\mathbf{q}-\mathbf{q}^{\kappa*})$
$\imath\mathbf{z}=\mathbf{z}\imath$	$\imath\mathbf{q}\neq\mathbf{q}\imath \quad \jmath\mathbf{q}\neq\mathbf{q}\jmath \quad \kappa\mathbf{q}\neq\mathbf{q}\kappa$
Augmented $\mathbb{C}$-variable z^a={z,z^*}	Augmented $\mathbb{Q}$-variable $q^a=\{q,q^\imath,q^\jmath,q^\kappa\}$
$\mathcal{C}_\mathbf{z}=E\{\mathbf{z}\mathbf{z}^H\} \quad \mathcal{P}_\mathbf{z}=E\{\mathbf{z}\mathbf{z}^T\}$	$\mathcal{C}_\mathbf{q}=E\{\mathbf{q}\mathbf{q}^H\} \quad \mathcal{C}_\mathbf{z}^\imath=E\{\mathbf{q}\mathbf{q}^{\imath H}\}$
	$\mathcal{C}_\mathbf{q}^\jmath=E\{\mathbf{q}\mathbf{q}^{\jmath H}\} \quad \mathcal{C}_\mathbf{z}^\kappa=E\{\mathbf{q}\mathbf{q}^{\kappa H}\}$
Properness does not imply that $E\{\Re\{z_\ell\}\Im\{z_k\}\}=0 \quad \forall k\neq\ell$	$\mathbb{Q}$-Properness implies that $E\{\Re\{q_\ell\}\Im\{q_k\}\}=0 \quad \forall k\neq\ell$
For a proper signal, $E\{\mathbf{z}\mathbf{z}^T\}$=$\mathbf{0}$	For a $\mathbb{Q}$-proper signal, $E\{\mathbf{q}\mathbf{q}^T\}\neq\mathbf{0}$
Complex-valued Gaussian distribution	Quaternion-valued Gaussian distribution
$p(\mathbf{z}^a)=\frac{1}{\pi^N\det(\mathcal{C}_\mathbf{z}^a)^{1/2}}\exp\left\{-\frac{1}{2}\mathbf{z}^{aH}\mathcal{C}_\mathbf{z}^{a^{-1}}\mathbf{z}^a\right\}$	$p(\mathbf{q}^a)=\frac{1}{(\pi^2/4)^N\det(\mathcal{C}_\mathbf{q}^a)^{1/2}}\exp\left\{-\frac{1}{2}\mathbf{q}^{aH}\mathcal{C}_\mathbf{q}^{a^{-1}}\mathbf{q}^a\right\}$
Mutual information	Interaction information

vector $\mathbf{q}^a$, that is

$$f(\mathbf{q}_a,\mathbf{q}_b,\mathbf{q}_c,\mathbf{q}_d)=\mathbf{q}^{rH}\mathcal{C}_R^{-1}\mathbf{q}^r=(\mathbf{q}^{aH}\mathbf{A}^{-H})(\mathbf{A}^H\mathcal{C}_\mathbf{q}^{a^{-1}}\mathbf{A})(\mathbf{A}^{-1}\mathbf{q}^a)$$
$$=\mathbf{q}^{aH}\mathcal{C}_\mathbf{q}^{a^{-1}}\mathbf{q}^a=p(\mathbf{q},\mathbf{q}^\imath,\mathbf{q}^\jmath,\mathbf{q}^\mathcal{K}) \qquad (38)$$

Substitute (37) and (38) into (34) to express the Gaussian probability density function for an augmented multivariate quaternion-valued random vector $\mathbf{q}^a$ as

$$p(\mathbf{q}^a)=p(\mathbf{q},\mathbf{q}^\imath,\mathbf{q}^\jmath,\mathbf{q}^\mathcal{K})=\frac{1}{(\pi^2/4)^N\det(\mathcal{C}_\mathbf{q}^a)^{1/2}}\exp\{-\tfrac{1}{2}\mathbf{q}^{aH}\mathcal{C}_\mathbf{q}^{a^{-1}}\mathbf{q}^a\} \qquad (39)$$

For a $\mathbb{Q}$-proper vector, it can be shown (using (27)) that the Gaussian distribution (39) simplifies to

$$p(\mathbf{q},\mathbf{q}^\imath,\mathbf{q}^\jmath,\mathbf{q}^\mathcal{K})=\frac{1}{(2\pi\sigma^2)^{2N}}\exp\{-\tfrac{1}{2\sigma^2}\mathbf{q}^H\mathbf{q}\} \qquad (40)$$

that is, the argument in the exponential is a real function of only $|\mathbf{q}|=\mathbf{q}^H\mathbf{q}$, thus highlighting the correspondence with the real and proper complex Gaussian distributions [5].

7. A note on information theoretic measures

7.1. Differential entropy for quaternion-valued random vectors

Based on Section 6 and the results in [6], we can now generalise the maximum entropy principle to the quaternion-valued multivariate case [6,28]. The differential entropy of a generic ($\mathbb{Q}$-proper or $\mathbb{Q}$-improper) quaternion-valued Gaussian random vector can be expressed as (the derivation is included in Appendix A.2)

$$H(\mathbf{q})=\log[(\pi e/2)^{2N}\det(\mathcal{C}_\mathbf{q}^a)^{1/2}] \qquad (41)$$

The upper bound on the differential entropy of a quaternion valued random vector $\mathbf{q}$ is given by

$$H(\mathbf{q})\leq 2N\log[(2\pi e\sigma^2)]=H_{\text{proper}} \qquad (42)$$

The equality holds for a centered $\mathbb{Q}$-proper Gaussian random vector $\mathbf{q}$ (as shown in Appendix A.3). It is straightforward to show[11] that the differential entropy of a quaternion random vector with arbitrary probability density function $p_A(\mathbf{q})$ cannot be greater than that of a vector with the Gaussian distribution (39) with the same augmented covariance matrix, thus confirming that a $\mathbb{Q}$-proper Gaussian process attains the upper entropy limit. The difference in entropy values is due to the improperness of a quaternion-valued Gaussian random vector and can be quantified by the difference between (42) and (41).

7.2. Beyond mutual information–interaction information

Standard mutual information (MI) considers only two variables, and its generalisation to higher dimensions is provided using the so-called 'interaction information' $\mathcal{I}$ [30]. Unlike mutual information, the interaction information $\mathcal{I}$ can be negative; physical meaning of a positive $\mathcal{I}$ can be interpreted as the consequence of an increase in the degree of association between the variates of a multivariate quantity, when one variable is kept constant. The reverse applies for $\mathcal{I}<0$ [30]. The interaction information $\mathcal{I}$ between quaternion-valued Gaussian random vectors $\mathbf{q}$, $\mathbf{q}^\imath$, $\mathbf{q}^\jmath$ and $\mathbf{q}^\kappa$ can be measured as

$$\begin{aligned}\mathcal{I}(\mathbf{q};\mathbf{q}^\imath;\mathbf{q}^\jmath;\mathbf{q}^\kappa)&=\mathcal{I}(\mathbf{q}_a;\mathbf{q}_b;\mathbf{q}_c;\mathbf{q}_d)\\&=\underbrace{H(\mathbf{q}_a)+H(\mathbf{q}_b)+H(\mathbf{q}_c)+H(\mathbf{q}_d)-H(\mathbf{q}_a,\mathbf{q}_b,\mathbf{q}_c,\mathbf{q}_d)}_{H_{\text{proper}}-H(\mathbf{q})}\\&\quad+H(\mathbf{q}_a,\mathbf{q}_b,\mathbf{q}_c)+H(\mathbf{q}_a,\mathbf{q}_b,\mathbf{q}_d)+H(\mathbf{q}_a,\mathbf{q}_c,\mathbf{q}_d)\\&\quad+H(\mathbf{q}_c,\mathbf{q}_b,\mathbf{q}_d)-H(\mathbf{q}_a,\mathbf{q}_b)-H(\mathbf{q}_a,\mathbf{q}_c)-H(\mathbf{q}_a,\mathbf{q}_d)\\&\quad-H(\mathbf{q}_b,\mathbf{q}_c)-H(\mathbf{q}_b,\mathbf{q}_d)-H(\mathbf{q}_c,\mathbf{q}_d)=\log\left[\frac{(8\sigma^4)^N}{\det(\mathcal{C}_\mathbf{q}^a)^{1/2}}\right]\\&\quad+H(\mathbf{q}_a,\mathbf{q}_b,\mathbf{q}_c)+H(\mathbf{q}_a,\mathbf{q}_b,\mathbf{q}_d)+H(\mathbf{q}_a,\mathbf{q}_c,\mathbf{q}_d)\\&\quad+H(\mathbf{q}_c,\mathbf{q}_b,\mathbf{q}_d)-H(\mathbf{q}_a,\mathbf{q}_b)-H(\mathbf{q}_a,\mathbf{q}_c)-H(\mathbf{q}_a,\mathbf{q}_d)\\&\quad-H(\mathbf{q}_b,\mathbf{q}_c)-H(\mathbf{q}_b,\mathbf{q}_d)-H(\mathbf{q}_c,\mathbf{q}_d)\end{aligned} \qquad (43)$$

[11] The proof is given on p. 336 of [29].

attaining the value of $\mathcal{I}=0$, for $\mathbb{Q}$-proper signals. Table 5 gives a comparative overview of both second order statistics and information theoretic measures in $\mathbb{C}$ and $\mathbb{H}$.

8. Concluding remarks

Second order statistics and information theoretic measures for quaternion-valued random variables and processes have been revisited. To make use of complete information within quaternion-valued second order statistics, complementary statistical descriptors the $\imath$-covariance, the $\jmath$-covariance, and the κ-covariance matrices have been employed. The so-introduced augmented statistics has served as a basis for a widely linear quaternion model and the widely linear Wiener solution, and the concept of $\mathbb{Q}$-properness (second order circularity) has been addressed based on the properties of the augmented covariance matrix. Further, the generic Gaussian multivariate distribution has been extended to quaternion-valued data, so as to cater for both $\mathbb{Q}$-proper and $\mathbb{Q}$-improper variables and vectors. The upper bound on the entropy of multivariate quaternion-valued processes has been provided, and it has been shown that this bound is attained for $\mathbb{Q}$-proper processes. Comparative analysis with real quadrivariate statistics supports the analysis.

Whereas Le Bihan et al. demonstrated how complex-valued statistics can be employed to give an insight into the statistics of quaternion-valued signals, we set out to introduce the augmented statistics directly in $\mathbb{H}$, thus providing a unifying framework for the analysis of second order circular and noncircular signals based on the isomorphism with quadrivariate real processes. The choice of the preferred approach depends on the particular application. A related work [31,32] uses a different basis to address quaternion statistics and develop closed form solutions in the context of principal component analysis and canonical correlation analysis. In this work and our companion articles [33,25], we have focused on statistical signal processing aspects such as widely linear model, stationarity, 'interaction information', and the Wiener filter.

Acknowledgement

We wish to acknowledge an anonymous reviewer, who at the time of publication drew our attention to a recent related article in [31].

Appendix A

A.1. Derivation of the widely linear Wiener solution in $\mathbb{H}$

Consider the standard real valued mean square error (MSE) function, that is,

$$\begin{aligned}\jmath &= E\{ee^*\}=E\{[d-y][d^*-y^*]\}\\ &=E\{dd^*\}+E\{yy^*\}-E\{yd^*\}-E\{dy^*\}\end{aligned} \tag{44}$$

The derivative of the cost function (44) can be expressed as

$$\begin{aligned}\nabla_{\mathbf{w}^a}\jmath &= E\{(\nabla_{\mathbf{w}^a}y)y^*+y(\nabla_{\mathbf{w}^a}y^*)-(\nabla_{\mathbf{w}^a}y)d^*-d(\nabla_{\mathbf{w}^a}y^*)\}\\ &=\underbrace{E\{4[\mathbf{x}^a y^*-\mathbf{x}^a d^*]\}}_{\Theta}+\underbrace{E\{2[d(n)\mathbf{x}^a-y\mathbf{x}^a]\}}_{\Phi}\end{aligned} \tag{45}$$

To obtain the Wiener solution, the expectations of the terms Θ and Φ in (45) are both set to zero. In the complex domain $\mathbb{C}$, the terms Θ and Φ in (45) are summed up; however, due to the non-commutativity of the quaternion product, we need to treat the terms in (45) separately, yielding the Wiener solution

$$\Theta : \mathbf{w}^a = E\{\mathbf{x}^a\mathbf{x}^{aH}\}^{-1}E\{\mathbf{x}^a d^*\} \tag{46}$$

$$\Phi : \mathbf{w}^a = E\{\mathbf{x}^{a*}\mathbf{x}^{aH}\}^{-1}E\{\mathbf{x}^{a*} d^*\} \tag{47}$$

The first term (46) requires the inversion of the augmented covariance $\mathcal{C}^a_{\mathbf{x}}=E\{\mathbf{x}^a\mathbf{x}^{aH}\}$, whereas the second condition (47) also relies on the conjugate of pseudocovariance matrix of the augmented vector $\mathbf{x}^a$, which conforms with the observation in [4] that the quaternion domain accounts inherently for the information contained in pseudocovariance.

A.2. Derivation of the maximum entropy of a quaternion-valued random vector

Let $p_A(\mathbf{q})$ be an arbitrary probability density function and $p(\mathbf{q})$ the Gaussian distribution (39). For convenience (with a slight abuse of notation), we denote $\iint\iint\cdots\iint\iint$ by a single integration symbol $\int$ and $dq_{a,1}dq_{b,1}dq_{c,1}dq_{d,1}\cdots dq_{a,N}dq_{b,N}dq_{c,N}dq_{d,N}$ by $d\mathbf{q}$

$$\begin{aligned}&\int p_A(\mathbf{q})\log\left[\frac{1}{p(\mathbf{q})}\right]d\mathbf{q}\\ &=\int p_A(\mathbf{q})\log\left[(\pi^2/4)^N\det(\mathcal{C}^a_{\mathbf{q}})^{1/2}\exp\{\tfrac{1}{2}\mathbf{q}^{a^H}\mathcal{C}^{a^{-1}}_{\mathbf{q}}\mathbf{q}^a\}\right]d\mathbf{q}\\ &\approx\int p_A(\mathbf{q})\log\left[(\pi^2/4)^N\det(\mathcal{C}^a_{\mathbf{q}})^{1/2}\exp\{2N\}\right]d\mathbf{q}\\ &\approx\log\left[(\pi^2e^2/4)^N\det(\mathcal{C}^a_{\mathbf{q}})^{1/2}\right]\int p_A(\mathbf{q})d\mathbf{q}\\ &\approx\log\left[(\pi e/2)^{2N}\det(\mathcal{C}^a_{\mathbf{q}})^{1/2}\right]\end{aligned} \tag{48}$$

Notice that the approximation in (48) arises because of the approximation of $\exp\{\frac{1}{2}\mathbf{q}^{a^H}\mathcal{C}^{a^{-1}}_{\mathbf{q}}\mathbf{q}^a\}\approx\exp\{\frac{1}{2}\,trace\{\mathcal{C}^{a^{-1}}\mathbf{q}^a\mathbf{q}^{a^H}\}\}=\exp\{2N\}$. For a $\mathbb{Q}$-proper Gaussian random vector, the augmented covariance matrix has the special structure (27), its determinant is $\det(\mathcal{C}^a_{\mathbf{q}})=(4\sigma^2)^{4N}$, and the expression (48) can be further simplified into

$$H_{\text{proper}}=2N\log[(2\pi e\sigma^2)] \tag{49}$$

A.3. Maximisation of entropy for a $\mathbb{Q}$-proper random variable

To demonstrate that the entropy of $q=q_a+\imath q_b+\jmath q_c+\kappa q_d\in\mathbb{H}$ is maximised for a $\mathbb{Q}$-proper random variable, we first address the maximum entropy of the corresponding real-valued quadrivariate vector $\mathbf{q}^r_s$=$[q_a\ q_b\ q_c\ q_d]^T$. According to the maximum entropy principle, the entropy of $\mathbf{q}^r_s$ satisfies (see p. 234 [28])

$$H(\mathbf{q}^r_s)\le\frac{1}{2}\log[(2\pi e)^4\det(\mathcal{C}_R)] \tag{50}$$

where the equality holds, iff $\mathbf{q}^r_s$ is a centered Gaussian random vector. Upon evaluating the corresponding entropies for N=1, observe that the real quadrivariate covariance matrix $\mathcal{C}_R$ in (20) is positive definite and has a special block structure

$$\mathcal{C}_R = \begin{bmatrix} \mathbf{\Gamma} & \mathbf{B} \\ \mathbf{B}^T & \mathbf{C} \end{bmatrix} \quad (51)$$

which implies that (see [34, p. 478])

$$\det(\mathcal{C}_R) = \det(\mathbf{\Gamma})\det(\mathbf{C} - \mathbf{B}^T\mathbf{\Gamma}^{-1}\mathbf{B}) \le \det(\mathbf{\Gamma})\det(\mathbf{C}) \quad (52)$$

and is maximised (equality holds) when $\mathbf{B}=\mathbf{0}$, yielding

$$E\{q_a q_c\} = E\{q_a q_d\} = E\{q_b q_c\} = E\{q_b q_d\} = 0 \quad (53)$$

Since for the two 2×2 matrices $\det(\mathbf{\Gamma}) = E\{q_a^2\}E\{q_b^2\} - E\{q_a q_b\}^2$ and $\det(\mathbf{C}) = E\{q_c^2\}E\{q_d^2\} - E\{q_c q_d\}^2$, the determinant $\det(\mathcal{C}_R)$ satisfies

$$\begin{aligned} \det(\mathcal{C}_R) &\le [E\{q_a^2\}E\{q_b^2\} - E\{q_a q_b\}^2][E\{q_c^2\}E\{q_d^2\} - E\{q_c q_d\}^2] \\ &= E\{q_a^2\}E\{q_b^2\}E\{q_c^2\}E\{q_d^2\} + E\{q_a q_b\}^2 E\{q_c q_d\}^2 \\ &\quad - E\{q_a q_b\}^2 E\{q_c^2\}E\{q_d^2\} - E\{q_c q_d\}^2 E\{q_a^2\}E\{q_b^2\} \\ &\le E\{q_a^2\}E\{q_b^2\}E\{q_c^2\}E\{q_d^2\} + E\{q_a q_b\}^2 E\{q_c q_d\}^2 \end{aligned} \quad (54)$$

The equality holds if and only if

$$E\{q_a q_b\}^2 = E\{q_c q_d\}^2 = 0 \quad (55)$$

Eqs. (53) and (55) satisfy property P2 of a $\mathbb{Q}$-proper variable in Table 2. Therefore, the determinant of $\mathcal{C}_R$ is upper bounded by

$$\det(\mathcal{C}_R) \le E\{q_a^2\}E\{q_b^2\}E\{q_c^2\}E\{q_d^2\} \quad (56)$$

Using constrained equality based optimisation, we show below that inequality (56) is maximised when the condition P1 of $\mathbb{Q}$-properness in Table 2 is satisfied, yielding

$$\det(\mathcal{C}_R) \le \left(\frac{E\{|q|^2\}}{4}\right)^4 \quad (57)$$

This optimisation problem can be posed as

$$\begin{aligned} &\max \quad \{\det(\mathcal{C}_R)\} = \max\{E\{q_a^2\}E\{q_b^2\}E\{q_c^2\}E\{q_d^2\}\} \\ &\text{subject to} \quad E\{q_a^2\} + E\{q_b^2\} + E\{q_c^2\} + E\{q_d^2\} = E\{|q|^2\} \end{aligned}$$

and can be solved using Lagrange multipliers as

$$\begin{aligned} f(E\{q_a^2\}, E\{q_b^2\}, E\{q_c^2\}, E\{q_d^2\}, \lambda) &= E\{q_a^2\}E\{q_b^2\}E\{q_c^2\}E\{q_d^2\} \\ &\quad + \lambda(E\{q_a^2\} + E\{q_b^2\} + E\{q_c^2\} + E\{q_d^2\} - E\{|q|^2\}) \end{aligned} \quad (58)$$

Set the derivative df=0, to yield the system of equations

$$\frac{\partial f}{\partial E\{q_a^2\}} = E\{q_b^2\}E\{q_c^2\}E\{q_d^2\} + \lambda = 0 \quad (59)$$

$$\frac{\partial f}{\partial E\{q_b^2\}} = E\{q_a^2\}E\{q_c^2\}E\{q_d^2\} + \lambda = 0 \quad (60)$$

$$\frac{\partial f}{\partial E\{q_c^2\}} = E\{q_a^2\}E\{q_b^2\}E\{q_d^2\} + \lambda = 0 \quad (61)$$

$$\frac{\partial f}{\partial E\{q_d^2\}} = E\{q_a^2\}E\{q_b^2\}E\{q_c^2\} + \lambda = 0 \quad (62)$$

$$\frac{\partial f}{\partial \lambda} = E\{q_a^2\} + E\{q_b^2\} + E\{q_c^2\} + E\{q_d^2\} - E\{|q|^2\} = 0 \quad (63)$$

Solving the Eqs. (59)–(62) leads to

$$E\{q_a^2\} = E\{q_b^2\} = E\{q_c^2\} = E\{q_d^2\} \quad (64)$$

which when replaced in (63) yields the solution

$$E\{q_a^2\} = E\{q_b^2\} = E\{q_c^2\} = E\{q_d^2\} = \frac{E\{|q|^2\}}{4} \quad (65)$$

Since the function $\log(\cdot)$ is monotically increasing, we can substitute the maximum value of $\det(\mathcal{C}_R)$ from (57) into (50), to obtain the upper entropy bound in the form

$$H(\mathbf{q}_s^r) \le \log\left[\frac{(\pi^2 e^2 E\{|q|^2\}^2}{4}\right] \le \log[4\pi^2 e^2 \sigma^4] \quad (66)$$

This upper bound is equivalent to the entropy of a $\mathbb{Q}$-proper Gaussian quaternion random variable (42) when N=1, thus illustrating that the entropy of a quaternion variable q is maximised for $\mathbb{Q}$-proper random variables. This also confirms the validity of the introduced form of probability density function (39) for quaternion random variables.

References

[1] D. Choukkroun, I.Y. Bar-Itzhack, Y. Oshman, Novel quaternion Kalman filter, IEEE Transactions on Aerospace and Electronics Systems 42 (1) (2006) 174–190.

[2] S. Miron, N. Le Bihan, J. Mars, Quaternion-MUSIC for vector-sensor array processing, IEEE Transactions on Signal Processing 54 (4) (2006) 1218–1229.

[3] N. Le Bihan, J. Mars, Singular value decomposition of quaternion matrices: a new tool for vector-sensor signal processing, Signal Processing 84 (7) (2004) 1177–1199.

[4] C. Cheong Took, D.P. Mandic, The quaternion LMS algorithm for adaptive filtering of hypercomplex real world processes, IEEE Transactions on Signal Processing 57 (4) (2009) 1316–1327.

[5] D.P. Mandic, V.S.L. Goh, Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models, John Wiley and Sons Ltd, 2009.

[6] F. Neeser, J. Massey, Proper complex random processes with applications to information theory, IEEE Transactions on Information Theory 39 (4) (1993) 1293–1302.

[7] A. Van Den Bos, The multivariate complex normal distribution—a generalization, IEEE Transactions on Information Theory 41 (2) (1995) 537–539.

[8] J. Navarro-Moreno, ARMA prediction of widely linear systems by using the innovations algorithm, IEEE Transactions on Signal Processing 56 (7) (2008) 3061–3068.

[9] J. Eriksson, V. Koivunen, Complex random vectors and ICA models: identifiability uniqueness and separability, IEEE Transactions on Information Theory 52 (3) (2006) 1017–1029.

[10] L. Fortuna, G. Muscato, M.G. Xibilia, A comparison between HMLP and HRBF for attitude control, IEEE Transactions on Neural Networks 12 (2) (2001) 318–328.

[11] C.E. Moxey, S.J. Sangwine, T.A. Ell, Hypercomplex correlation techniques for vector images, IEEE Transactions on Signal Processing 51 (7) (2003) 1941–1953.

[12] N.N. Vakhania, Random vectors with values in quaternion Hilbert spaces, Theory of Probability and its Applications 43 (1) (1999) 99–115.

[13] P.O. Amblard, N. Le Bihan, On properness of quaternion valued random variables, in: Proceedings of International Conference on Mathematics (IMA) in Signal Processing, Cirencester, 2004, pp. 23–26.

[14] N. Le Bihan, P.O. Amblard, Detection and estimation of Gaussian proper quaternion valued random processes, IMA Conference on Mathematics in Signal Processing, Cirencester, U.K., 2006, pp. 23–26.

[15] S. Buchholz, N. Le Bihan, Optimal separation of polarized signals by quaternionic neural networks, in: Proceedings of 14th European Signal Processing Conference, EUSIPCO, 2006.

[16] P.J. Schreier, L.L. Scharf, Statistical Signal Processing of Complex-Valued Data: The Theory of Improper and Noncircular Signals, Cambridge University Press, 2010.

[17] J.P. Ward, Quaternions and Cayley Numbers: Algebra and Applications, Kluwer Academic Publisher, 1997.

[18] T.A. Ell, S.J. Sangwine, Quaternion involutions and anti-involutions, Computers & Mathematics with Applications 53 (1) (2007) 137–143.

[19] B. Picinbono, Random Signals and Systems, Prentice-Hall, 1993.

[20] F. Zhang, Quaternions and matrices of quaternions, Linear Algebra and its Applications 251 (1997) 21–57.
[21] S.J. Sangwine, N. Le Bihan, Quaternion toolbox for Matlab [online], Available: 〈http://qtfm.sourceforge.net/〉.
[22] W.A. Gardner, Introduction to Random Processes with Applications to Signals and Systems, second ed., McGraw-Hill, New York, 1990.
[23] B. Picinbono, On circularity, IEEE Transactions on Signal Processing 42 (12) (1994) 3473–3482.
[24] B. Picinbono, P. Chevalier, Widely linear estimation with complex data, IEEE Transactions on Signal Processing 43 (8) (1995) 2030–2033.
[25] C. Cheong Took, D.P. Mandic, New developments in statistical signal processing of quaternion random variables with applications in wind forecasting, in: Proceedings of the International Symposium on Evolving Intelligent Systems at the AISB 2010 convention, 2010, pp. 58–62.
[26] Y. Huang, J. Benesty, Audio Signal Processing for Next Generation Multimedia Communication Systems, Kluwer Academic Publishers, 2004.
[27] B. Picinbono, Second-order complex vectors and normal distributions, IEEE Transactions on Signal Processing 44 (10) (1996) 2637–2640.
[28] T.M. Cover, J.A. Thomas, Elements of Information Theory, Wiley, New York, 1991.
[29] R.G. Gallager, Information Theory and Reliable Communications, Wiley, New York, 1968.
[30] W.J. McGill, Multivariate information transmission, IEEE Transactions on Information Theory 4 (4) (1954) 93–111.
[31] J. Via, D. Ramirez, I. Santamaria, Propreness and widely linear processing of quaternion random vectors, IEEE Transactions on Information Theory 56 (7) (2010) 3502–3515.
[32] J. Vía, D. Ramírez, I. Santamaría, L. Vielva, Widely and semi-widely linear processing of quaternion vectors, in: Proceedings of International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2010, pp. 3946–3949.
[33] C. Cheong Took, D.P. Mandic, A quaternion widely linear adaptive filter, IEEE Transactions on Signal Processing 58 (8) (2010) 1–5.
[34] R.A. Horn, Matrix Analysis, Cambridge University Press, 1990.