Technical Report: TR-ICL-EP-D061709-03-NOV-09

Electrical and Electronic Department,

Imperial College London, London SW7 2AZ, United Kingdom.

URL: http://www.commsp.ee.ic.ac.uk/~mandic/Tech_Rep.htm

Second-Order Statistics of Quaternion Random Processes[1]

Clive Cheong Took and Danilo Mandic

E-mail:{c.cheong-took, d.mandic}@imperial.ac.uk

Electrical and Electronic Department, South Kensington Campus,

Imperial College London, London SW7 2AZ, United Kingdom.

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 a $\mathbb{Q}$-proper (second order circular) random process are presented, and to cater for the non-vanishing pseudocovariance matrix 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 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 quaternion valued processes. A rigorous comparison with the corresponding second order statistics of quadrivariate real valued processes supports the approach.

Keywords: Quaternion random variables, $\mathbb{Q}$-properness, multivariate Gaussian distribution, quadri-

variate entropy, quadrivariate random processes.

I. Introduction

Standard techniques employed in statistical multichannel processing typically do not fully exploit the ‘coupled’ nature of the available information within the channels. In other words, 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 the modelling of three- and four-dimensional signals. The resurgence in quaternion signal processing is due to the advantages of quaternion algebra over real-valued quadrivariate vector algebra in the modelling of such data. Applications of quaternions include those in information forensics [1], instrumentation [2], communications [3], robotics [4], neural networks [5], and seismics & oceanics [6]. In the signal processing community, quaternions have been employed in Kalman filtering [7], the well-known MUSIC technique [8], singular value decomposition for vector sensing [9] and the least-mean-square estimation [10].

As quaternions are a hypercomplex extension of complex numbers, it is natural to investigate whether the recent developments in so called augmented complex statistics can be extended to the quaternion domain, in order to cater for the generality of random signals. One of the pioneering results in augmented complex statistics is the work by Neeser and Massey, who introduced the concept of properness (second order circularity, rotation invariant probability distribution) into complex-valued statistics. They demonstrated that the covariance $E\{\mathbf{z}\mathbf{z}^H\}$ of a complex random vector $\mathbf{z}$ alone is not adequate to provide a complete second order statistical description [11] and that the pseudocovariance $E\{\mathbf{z}\mathbf{z}^T\}$ also needs to be considered in order to cater for improper signals. Their work was followed by Picinbono [12] and Van Den Bos, who formulated a generic 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 [13]. These foundations have been successfully used to design novel algorithms in adaptive signal processing [14], communications [15], autoregressive moving average (ARMA) modelling [16], and independent component analysis [17]. By virtue of augmented complex statistics, all these results are applicable to the generality of complex signals, both second order circular and noncircular.

The results of work on quaternion-valued second order statistics are still emerging and are scattered in the literature [18], [19]. The existing approaches typically take into account only the information contained in quaternion-valued covariance [5], [7], [8], [20] and by analogy with the complex domain, they are bound not to maximise the use of available statistical information. However, despite quaternions being a natural generalisation of complex numbers (their hyper-complex 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. In this connection, Vakhania extended the concept of 'properness' to the quaternion domain, however, his definition of $\mathbb{Q}$-properness was restricted to the invariance of the probability density function (pdf) under some specific rotations around angle of $\pi/2$ [18]. Amblard and Le Bihan relaxed the conditions of $\mathbb{Q}$-properness to an arbitrary axis and angle of rotation φ, that is [19]

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

for any pure unit quaternion ν (whose real part vanishes); symbol $\triangleq$ denotes equality in terms of pdf. Although these results provide an initial insight into $\mathbb{Q}$-properness, they are restricted to single quaternion variables and it is not straightforward to apply them to quaternion-valued processes.

The augmented statistics of complex variables and signals was addressed in detail in [21], [22]. We here extend this analysis to cater for the quaternion domain and derive conditions for complete second order statistical description of such signals. To that end, this work introduces a generic framework for second order statistical analysis of the generality of quaternion-valued random variables and vectors, both second order circular and noncircular. It is demonstrated 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 [14], [23]. Our analysis shows that the basis for augmented quaternion statistics should also include quaternion involutions, and that the so introduced augmented covariance matrix contains all necessary second order statistical information, also leading to the introduction of widely linear modelling in $\mathbb{H}$. Next, multivariate

Gaussian distribution is revisited 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 II we briefly review the elements of quaternion algebra. In Section III, novel statistical measures for quaternion-valued variables are introduced and the duality with their quadrivariate real domain counterparts is addressed. Next, Section VI revisits the fundamentals of $\mathbb{Q}$-properness and illustrates its implications for quaternion statistics. Section VII formulates a generic Gaussian distribution to cater for both $\mathbb{Q}$-proper and $\mathbb{Q}$-improper signals. In Section VIII, 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 IX.

II. Properties of Quaternion Random Vectors

A. Quaternion Algebra

The 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 [24]. 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:

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

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

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

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

$$\begin{aligned} q_1 q_2 &= \Re\{q_1 q_2\} + \Im\{q_1 q_2\} \\ \text{where} \quad \Re\{q_1 q_2\} &= q_{1,a}q_{2,a} + q_{1,b}q_{2,b} + q_{1,c}q_{2,c} + q_{1,d}q_{2,d} \\ \Im\{q_1 q_2\} &= q_{1,a}\Im\{q_2\} + q_{2,a}\Im\{q_1\} + \Im\{q_1\} \times \Im\{q_2\} \end{aligned} \tag{4}$$

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

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

B. 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 $z_b = \frac{1}{2\imath}(z - z^*)$. 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 $[z \;\; z^*]^T$. However, the quaternion domain does not permit such convenient manipulation and the 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 circumvent this problem, we propose to employ the three perpendicular quaternion involutions (self-inverse mappings), given by

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

The four components of the quaternion variable q can now be expressed as [25]

$$\begin{aligned} q_a &= \frac{1}{2}(q + q^*) \qquad & q_b &= \frac{1}{2\imath}(q - q^{\imath *}) \\ q_c &= \frac{1}{2\jmath}(q - q^{\jmath *}) \qquad & q_d &= \frac{1}{2\kappa}(q - q^{\kappa *}) \end{aligned} \tag{7}$$

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

$$q^* = \frac{1}{2}(q^\imath + q^\jmath + q^\kappa - q) \tag{8}$$

By introducing the augmented quaternion statistics, we aim to establish the duality between the second order statistics of 'augmented' quaternion processes $\mathbf{q}^a \in \mathbb{H}^{4N\times 1}$ and quadrivariate real valued vectors in $\mathbb{R}^{4N\times 1}$. To make the augmented statistics in $\mathbb{H}$ suitable for the description of both second order circular and noncircular signals, following on (see pp. 118-119 [26]), we need to establish a one-to-one correspondence between the components of a quadrivariate real variable and its quaternionic counterpart. For convenient manipulation of the components of quaternion variables, we shall use a combination[2] of $\{q,\ q^*,\ q^\imath,\ q^\jmath,\ q^\kappa\}$, and thus define the augmented quaternion vector $\mathbf{q}^a = [\mathbf{q}^T\ \mathbf{q}^{\imath T}\ \mathbf{q}^{\jmath T}\ \mathbf{q}^{\kappa T}]^T$ as

$$\mathbf{q}^a = \mathbf{A}\mathbf{q}^r$$
$$\begin{bmatrix} \mathbf{q} \\ \mathbf{q}^\imath \\ \mathbf{q}^\jmath \\ \mathbf{q}^\kappa \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \imath\mathbf{I} & \jmath\mathbf{I} & \kappa\mathbf{I} \\ \mathbf{I} & \imath\mathbf{I} & -\jmath\mathbf{I} & -\kappa\mathbf{I} \\ \mathbf{I} & -\imath\mathbf{I} & \jmath\mathbf{I} & -\kappa\mathbf{I} \\ \mathbf{I} & -\imath\mathbf{I} & -\jmath\mathbf{I} & \kappa\mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{q}_a \\ \mathbf{q}_b \\ \mathbf{q}_c \\ \mathbf{q}_d \end{bmatrix} \tag{9}$$

where $\mathbf{I} \in \mathbb{R}^{N\times N}$ is the identity matrix, and $\mathbf{q} = [q_1\ q_2 \cdots q_N]^T \ \in \mathbb{H}^{N\times 1}$; similar description also applies to $\mathbf{q}^\imath$, $\mathbf{q}^\jmath$, $\mathbf{q}^\kappa \in \mathbb{H}^{N\times 1}$, and $\mathbf{q}_a$, $\mathbf{q}_b$, $\mathbf{q}_c$ and $\mathbf{q}_d \ \in \mathbb{R}^{N\times 1}$. The $4N \times 4N$ matrix $\mathbf{A}$ provides an invertible mapping between the augmented quaternion valued signal $\mathbf{q}^a \in \mathbb{H}^{4N\times 1}$ and the quadrivariate 'composite' real valued vector $\mathbf{q}^r = [\mathbf{q}_a^T\ \mathbf{q}_b^T\ \mathbf{q}_c^T\ \mathbf{q}_d^T]^T \in \mathbb{R}^{4N\times 1}$, and its inverse is

$$\mathbf{A}^{-1} = \frac{1}{4}\mathbf{A}^H \tag{10}$$

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 $N = 1$, $\det(\mathbf{A}) = 16$. For any arbitrary N, the determinant of $\mathbf{A}$ therefore becomes

$$\det(\mathbf{A}) = 16^N \tag{11}$$

The basis $\{q, q^\imath, q^\jmath, q^\kappa\}$ in (9) has been selected so as to make the matrix $\mathbf{A}$ is unitary, which facilitates its algebraic manipulation. In the sequel, we will show that due to the relation (8), any other combination of four elements of $\{q,\ q^*,\ q^\imath,\ q^\jmath,\ q^\kappa\}$, for instance the basis $\{q, q^*, q^{\imath *}, q^{\jmath *}\}$ is also valid, but this does not guarantee a unitary $\mathbf{A}$.

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

III. Quaternion Statistics

A. Preliminaries

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

$$\begin{aligned}\mathcal{C}_{\mathbf{qq}} &= E\{\mathbf{q}\mathbf{q}^H\} \\ &= \Re\{\mathcal{C}_{\mathbf{qq}}\} + \imath\Im_\imath\{\mathcal{C}_{\mathbf{qq}}\} + \jmath\Im_\jmath\{\mathcal{C}_{\mathbf{qq}}\} + \kappa\Im_\kappa\{\mathcal{C}_{\mathbf{qq}}\} \end{aligned} \tag{12}$$

and its structure is shown in Table I. 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 I, 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 (7) and (9), 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. They augment the information within the covariance, and are given by

$$\begin{aligned}\mathcal{C}_{\mathbf{q}\imath} &= E\{\mathbf{q}\mathbf{q}^{\imath H}\} \\ &= \Re\{\mathcal{C}_{\mathbf{q}\imath}\} + \imath\Im_\imath\{\mathcal{C}_{\mathbf{q}\imath}\} + \jmath\Im_\jmath\{\mathcal{C}_{\mathbf{q}\imath}\} + \kappa\Im_\kappa\{\mathcal{C}_{\mathbf{q}\imath}\} \end{aligned} \tag{13}$$

$$\begin{aligned}\mathcal{C}_{\mathbf{q}\jmath} &= E\{\mathbf{q}\mathbf{q}^{\jmath H}\} \\ &= \Re\{\mathcal{C}_{\mathbf{q}\jmath}\} + \imath\Im_\imath\{\mathcal{C}_{\mathbf{q}\jmath}\} + \jmath\Im_\jmath\{\mathcal{C}_{\mathbf{q}\jmath}\} + \kappa\Im_\kappa\{\mathcal{C}_{\mathbf{q}\jmath}\} \end{aligned} \tag{14}$$

$$\begin{aligned}\mathcal{C}_{\mathbf{q}\kappa} &= E\{\mathbf{q}\mathbf{q}^{\kappa H}\} \\ &= \Re\{\mathcal{C}_{\mathbf{q}\kappa}\} + \imath\Im_\imath\{\mathcal{C}_{\mathbf{q}\kappa}\} + \jmath\Im_\jmath\{\mathcal{C}_{\mathbf{q}\kappa}\} + \kappa\Im_\kappa\{\mathcal{C}_{\mathbf{q}\kappa}\} \end{aligned} \tag{15}$$

where 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 Table I and Table II. 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 has a skew-symmetric structure, giving rise to its $\imath$-Hermitian

TABLE I

STRUCTURES OF THE QUATERNION-VALUED COVARIANCE MATRICES IN TERMS OF THEIR QUADRIVARIATE REAL-VALUED COUNTERPARTS

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

TABLE II

STRUCTURES OF THE QUATERNION-VALUED COVARIANCE MATRICES IN TERMS OF THEIR QUADRIVARIATE REAL-VALUED COUNTERPARTS

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

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

$$\begin{aligned} \mathcal{C}_{\mathbf{q}^\imath} &= \mathcal{C}_{\mathbf{q}^\imath}^{\imath H} \\ \mathcal{C}_{\mathbf{q}^\jmath} &= \mathcal{C}_{\mathbf{q}^\jmath}^{\jmath H} \\ \mathcal{C}_{\mathbf{q}^\kappa} &= \mathcal{C}_{\mathbf{q}^\kappa}^{\kappa H} \end{aligned} \tag{16}$$

These properties do not arise in the statistics of complex valued random variables [14], [17], and are unique to the quaternion domain.

B. Duality between quaternionic and quadrivariate statistics

Advances in the statistics of complex variables have shown that the covariance matrix alone is not adequate to completely describe the second order statistical properties of general complex-valued random vectors[3] $\mathbf{z} = \mathbf{z}_a + \imath\mathbf{z}_b$. Picinbono showed that the complete description of the second order statistics in $\mathbb{C}$, catering for both proper and improper signals, can be achieved if the real valued bivariate covariance matrices can be computed from their complex valued counterparts (see pp. 118-119 [26]). In Section II, we have 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 comprising the involutions $q^\imath$ and $q^\jmath$ and q^κ. Following on these results, we can obtain a complete second order statistical description in $\mathbb{H}$, 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 as[4]

$$
\begin{aligned}
\mathcal{C}_{\mathbf{q}_a} &= \frac{1}{4}\Re\{\mathcal{C}_{\mathbf{qq}} + \mathcal{C}_{\mathbf{q}^\imath} + \mathcal{C}_{\mathbf{q}^\jmath} + \mathcal{C}_{\mathbf{q}^\kappa}\} & \mathcal{C}_{\mathbf{q}_b} &= \frac{1}{4}\Re\{\mathcal{C}_{\mathbf{qq}} + \mathcal{C}_{\mathbf{q}^\imath} - \mathcal{C}_{\mathbf{q}^\jmath} - \mathcal{C}_{\mathbf{q}^\kappa}\} \\
\mathcal{C}_{\mathbf{q}_c} &= \frac{1}{4}\Re\{\mathcal{C}_{\mathbf{qq}} - \mathcal{C}_{\mathbf{q}^\imath} + \mathcal{C}_{\mathbf{q}^\jmath} - \mathcal{C}_{\mathbf{q}^\kappa}\} & \mathcal{C}_{\mathbf{q}_d} &= \frac{1}{4}\Re\{\mathcal{C}_{\mathbf{qq}} - \mathcal{C}_{\mathbf{q}^\imath} - \mathcal{C}_{\mathbf{q}^\jmath} + \mathcal{C}_{\mathbf{q}^\kappa}\} \\
\mathcal{C}_{\mathbf{q}_b\mathbf{q}_a} &= \frac{1}{4}\Im_\imath\{\mathcal{C}_{\mathbf{qq}} + \mathcal{C}_{\mathbf{q}^\imath} + \mathcal{C}_{\mathbf{q}^\jmath} + \mathcal{C}_{\mathbf{q}^\kappa}\} & \mathcal{C}_{\mathbf{q}_c\mathbf{q}_a} &= \frac{1}{4}\Im_\jmath\{\mathcal{C}_{\mathbf{qq}} + \mathcal{C}_{\mathbf{q}^\imath} + \mathcal{C}_{\mathbf{q}^\jmath} + \mathcal{C}_{\mathbf{q}^\kappa}\} \\
\mathcal{C}_{\mathbf{q}_d\mathbf{q}_a} &= \frac{1}{4}\Im_\kappa\{\mathcal{C}_{\mathbf{qq}} + \mathcal{C}_{\mathbf{q}^\imath} + \mathcal{C}_{\mathbf{q}^\jmath} + \mathcal{C}_{\mathbf{q}^\kappa}\} & \mathcal{C}_{\mathbf{q}_c\mathbf{q}_b} &= \frac{1}{4}\Im_\kappa\{\mathcal{C}_{\mathbf{qq}} + \mathcal{C}_{\mathbf{q}^\imath} - \mathcal{C}_{\mathbf{q}^\jmath} - \mathcal{C}_{\mathbf{q}^\kappa}\} \\
\mathcal{C}_{\mathbf{q}_d\mathbf{q}_b} &= -\frac{1}{4}\Im_\jmath\{\mathcal{C}_{\mathbf{qq}} + \mathcal{C}_{\mathbf{q}^\imath} - \mathcal{C}_{\mathbf{q}^\jmath} - \mathcal{C}_{\mathbf{q}^\kappa}\} & \mathcal{C}_{\mathbf{q}_d\mathbf{q}_c} &= \frac{1}{4}\Im_\imath\{\mathcal{C}_{\mathbf{qq}} - \mathcal{C}_{\mathbf{q}^\imath} + \mathcal{C}_{\mathbf{q}^\jmath} - \mathcal{C}_{\mathbf{q}^\kappa}\}
\end{aligned} \tag{17}
$$

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

$$
\begin{aligned}
\mathcal{C}_{\mathbf{z_a}} &= \frac{1}{2}\Re\{\mathcal{C}_\mathbf{z} + \mathcal{P}_\mathbf{z}\} & \mathcal{C}_{\mathbf{z_b}} &= \frac{1}{2}\Re\{\mathcal{C}_\mathbf{z} - \mathcal{P}_\mathbf{z}\} \\
\mathcal{C}_{\mathbf{z_a}\mathbf{z_b}} &= \frac{1}{2}\Im_\imath\{\mathcal{P}_\mathbf{z} - \mathcal{C}_\mathbf{z}\} & \mathcal{C}_{\mathbf{z_b}\mathbf{z_a}} &= \mathcal{C}^T_{\mathbf{z_a}\mathbf{z_b}}
\end{aligned}
$$

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.

[4]If a different basis, e.g. $\{q, q^*, q^{\imath *}, q^{\jmath *}\}$ is chosen, the full description of the second order statistics is still achieved, as shown in Appendix X-A; this applies to any other combination of quadruples based on $\{q, q^*, q^\imath, q^\jmath, , q^\kappa\}$.

The augmented quaternion-valued covariance matrix of an augmented random vector $\mathbf{q}^a = [\mathbf{q}^T\ \mathbf{q}^{\imath T}\ \mathbf{q}^{\jmath T}\ \mathbf{q}^{\kappa T}]^T$ $\big($see also (9)$\big)$, is therefore given by,

$$\mathcal{C}^a_{\mathbf{q}} = E\{\mathbf{q}^a\mathbf{q}^{aH}\} = \begin{bmatrix} \mathcal{C}_{\mathbf{q}\mathbf{q}} & \mathcal{C}_{\mathbf{q}^\imath} & \mathcal{C}_{\mathbf{q}^\jmath} & \mathcal{C}_{\mathbf{q}^\kappa} \\ \mathcal{C}^H_{\mathbf{q}^\imath} & \mathcal{C}_{\mathbf{q}^\imath\mathbf{q}^\imath} & \mathcal{C}_{\mathbf{q}^\imath\mathbf{q}^\jmath} & \mathcal{C}_{\mathbf{q}^\imath\mathbf{q}^\kappa} \\ \mathcal{C}^H_{\mathbf{q}^\jmath} & \mathcal{C}_{\mathbf{q}^\jmath\mathbf{q}^\imath} & \mathcal{C}_{\mathbf{q}^\jmath\mathbf{q}^\jmath} & \mathcal{C}_{\mathbf{q}^\jmath\mathbf{q}^\kappa} \\ \mathcal{C}^H_{\mathbf{q}^\kappa} & \mathcal{C}_{\mathbf{q}^\kappa\mathbf{q}^\imath} & \mathcal{C}_{\mathbf{q}^\kappa\mathbf{q}^\jmath} & \mathcal{C}_{\mathbf{q}^\kappa\mathbf{q}^\kappa} \end{bmatrix} \tag{18}$$

where the submatrices in (18) are calculated according to

$$\begin{aligned} &\mathcal{C}_{\boldsymbol{\delta}} = E\{\mathbf{q}\boldsymbol{\delta}^H\} && \mathcal{C}_{\boldsymbol{\alpha}\boldsymbol{\beta}} = E\{\boldsymbol{\alpha}\boldsymbol{\beta}^H\} \\ &\boldsymbol{\delta} \in \{\mathbf{q}^\imath, \mathbf{q}^\jmath, \mathbf{q}^\kappa\} && \boldsymbol{\alpha}, \boldsymbol{\beta} \in \{\mathbf{q}, \mathbf{q}^\imath, \mathbf{q}^\jmath, \mathbf{q}^\kappa\} \end{aligned} \tag{19}$$

To verify that the augmented covariance matrix in (18) provides a complete second order statistical description, we need to show that it permits a static invertible one-to-one mapping 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 (9), and since from (10) $\mathbf{q}^r = \mathbf{A}^{-1}\mathbf{q}^a = \frac{1}{4}\mathbf{A}^H\mathbf{q}^a$, the real valued covariance matrix can indeed be expressed in terms of the augmented quaternion valued covariance matrix in (18) as

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

where $\mathbf{A}^{-H} = \big(\mathbf{A}^{-1}\big)^H$. This completes the derivation of the augmented quaternion statistics, suitable for the description of both proper and improper quaternion random processes.

IV. 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 [27]. Based on (9) and (18), it is the augmented random vector $\mathbf{x}^a = [\mathbf{x}^T \ \mathbf{x}^{\imath T} \ \mathbf{x}^{\jmath T} \ \mathbf{x}^{\kappa T}]^T$ that contains all the required second order statistical information. Then, the quaternion widely linear model (QWL) can be constructed as

$$\begin{aligned} y &= \mathbf{w}^{aH}\mathbf{x}^a \\ &= \mathbf{g}^H\mathbf{x} + \mathbf{h}^H\mathbf{x}^\imath + \mathbf{u}^H\mathbf{x}^\jmath + \mathbf{v}^H\mathbf{x}^\kappa \end{aligned} \tag{22}$$

The MSE solution based on the QWL model (22) is then given by

$$\mathbf{w}^a = E\{\mathbf{x}^a\mathbf{x}^{aH}\}^{-1}E\{\mathbf{x}^a d^*\} \tag{23}$$

demonstrating that the QWL solution has the same form as the standard solution, but is based on the augmented covariance matrix $\mathcal{C}^a_{\mathbf{x}_a}$ in (18). On the other hand, the corresponding real-valued quadrivariate model relies on the real-valued covariance matrix in (20) [28]. This correspondence can be used 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

$$\mathcal{C}_R - \lambda\mathbf{I} = \frac{1}{16}\mathbf{A}^H[\mathcal{C}^a_{\mathbf{x}_a} - 4\lambda\mathbf{I}]\mathbf{A} \tag{24}$$

that is, the eigenvalues of the augmented quaternion covariance matrix are four times those of the quadrivariate real-valued correlation matrix. Hence, if the quaternion least mean square (QLMS) algorithm exploits the widely linear model, it will converge four times faster than its multichannel counterpart, for the same learning rate (see also [10]).

V. Second order stationarity

Recall that a real-valued quadrivariate variable is wide-sense stationary if all its four components are wide-sense stationary [29]. Since the four quaternion-valued covariance matrices (12)-(15) provide a full description of the second order statistics, we can now state that a quaternion-valued random process $q(n)$ is wide-sense stationary, provided

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

$$\begin{aligned} \mathcal{C}_{\mathbf{qq}}(n,\tau) &= E\{\mathbf{q}(n)\mathbf{q}^H(n+\tau)\} \\ \mathcal{C}_{\mathbf{q}^\imath}(n,\tau) &= E\{\mathbf{q}(n)\mathbf{q}^{\imath H}(n+\tau)\} \\ \mathcal{C}_{\mathbf{q}^\jmath}(n,\tau) &= E\{\mathbf{q}(n)\mathbf{q}^{\jmath H}(n+\tau)\} \\ \mathcal{C}_{\mathbf{q}^\kappa}(n,\tau) &= E\{\mathbf{q}(n)\mathbf{q}^{\kappa H}(n+\tau)\} \end{aligned}$$

3) The covariance matrix is finite, $\mathcal{C}_{\mathbf{qq}}(n) = E\{\mathbf{q}(n)\mathbf{q}^H(n)\} < \infty \quad \forall\, n$

Observation. *It is sufficient to define stationarity in terms of $\mathcal{C}_{\mathbf{qq}}$, $\mathcal{C}_{\mathbf{q}\imath}$, $\mathcal{C}_{\mathbf{q}\jmath}$, and $\mathcal{C}_{\mathbf{q}\kappa}$, as they provide the complete description of second order statistics*[5] *(17).*

VI. 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 [12]. The two conditions imposed on a complex variable $z = z_a + \imath z_b$ to be proper ($\mathbb{C}$-proper) are therefore

$$\begin{aligned} \sigma^2_{z_a} &= \sigma^2_{z_b} \\ E\{z_a z_b\} &= 0 \end{aligned} \qquad (25)$$

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

By continuity, a quaternion-valued second order circular ($\mathbb{Q}$-proper) variable should satisfy the two conditions in (25) 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 III [18].

The first property, P1, states that all the four components of a $\mathbb{Q}$-proper variable have equal powers. The property P2 implies that all 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.

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

TABLE III

PROPERTIES OF A $\mathbb{Q}$-PROPER RANDOM VARIABLE

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

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 also has to be proper. In addition, any subvector of a proper random vector is also proper. To guarantee the joint $\mathbb{Q}$-properness, each element of the vectors $\mathbf{q}^\vartheta$ and $\mathbf{q}^\upsilon$ should satisfy properties P1 and P2 in Table III, and the elements should be uncorrelated in the sense that their joint $\imath$-$\jmath$-κ-covariance matrices vanish. This is discussed in more detail below.

A. Augmented Statistics and $\mathbb{Q}$-properness

Following on the notion of proper complex variables (as detailed in Section IIIA of [11]), we now extend this definition to quaternion random vectors. Consider a $\mathbb{Q}$-proper random vector $\mathbf{q} = [q_1\ \ q_2 \cdots q_N]^T \ \in \mathbb{H}^{N\times 1}$. Then, $\mathbb{Q}$-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[6],

$$E\{\mathbf{q}\mathbf{q}^{\imath H}\} = \mathbf{0} \qquad E\{\mathbf{q}\mathbf{q}^{\jmath H}\} = \mathbf{0} \qquad E\{\mathbf{q}\mathbf{q}^{\kappa H}\} = \mathbf{0} \tag{26}$$

In other words, a $\mathbb{Q}$-proper signal has a vanishing complementary covariance matrices, specified in Table I and Table II. Also, the invariance of a $\mathbb{Q}$-proper random vector under a linear or affine transformation (shown in Appendix X-B) is similar to that in the complex case (see Lemma 3 [11]). This invariance arises due to the properties in (16) and the condition of vanishing

[6]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 $E\{\mathbf{z}(\mathbf{z}^*)^H\} = E\{\mathbf{z}\mathbf{z}^T\} = \mathbf{0}$.

$\imath-$covariance matrix, $\mathcal{C}_{\mathbf{q}^{\imath}} = \mathbf{0}$, in (13) is equivalent to the conditions

$$\begin{aligned} \mathcal{C}_{\mathbf{q}_b\mathbf{q}_a} &= \mathcal{C}^T_{\mathbf{q}_b\mathbf{q}_a} \qquad \mathcal{C}_{\mathbf{q}_c\mathbf{q}_d} = \mathcal{C}^T_{\mathbf{q}_c\mathbf{q}_d} \\ \mathcal{C}_{\mathbf{q}_c\mathbf{q}_a} &= -\mathcal{C}^T_{\mathbf{q}_c\mathbf{q}_a} \qquad \mathcal{C}_{\mathbf{q}_d\mathbf{q}_b} = -\mathcal{C}^T_{\mathbf{q}_d\mathbf{q}_b} \\ \mathcal{C}_{\mathbf{q}_d\mathbf{q}_a} &= -\mathcal{C}^T_{\mathbf{q}_d\mathbf{q}_a} \qquad \mathcal{C}_{\mathbf{q}_b\mathbf{q}_c} = -\mathcal{C}^T_{\mathbf{q}_b\mathbf{q}_c} \end{aligned} \tag{27}$$

Similarly, the vanishing $\jmath-$covariance matrix, $\mathcal{C}_{\mathbf{q}^{\jmath}} = \mathbf{0}$, in (14) implies

$$\begin{aligned} \mathcal{C}_{\mathbf{q}_b\mathbf{q}_a} &= -\mathcal{C}^T_{\mathbf{q}_b\mathbf{q}_a} \qquad \mathcal{C}_{\mathbf{q}_c\mathbf{q}_d} = -\mathcal{C}^T_{\mathbf{q}_c\mathbf{q}_d} \\ \mathcal{C}_{\mathbf{q}_c\mathbf{q}_a} &= \mathcal{C}^T_{\mathbf{q}_c\mathbf{q}_a} \qquad \mathcal{C}_{\mathbf{q}_d\mathbf{q}_b} = \mathcal{C}^T_{\mathbf{q}_d\mathbf{q}_b} \\ \mathcal{C}_{\mathbf{q}_d\mathbf{q}_a} &= -\mathcal{C}^T_{\mathbf{q}_d\mathbf{q}_a} \qquad \mathcal{C}_{\mathbf{q}_b\mathbf{q}_c} = -\mathcal{C}^T_{\mathbf{q}_b\mathbf{q}_c} \end{aligned} \tag{28}$$

whereas, the vanishing the $\kappa-$covariance matrix, $\mathcal{C}_{\mathbf{q}^{\kappa}} = \mathbf{0}$, in (15) yields

$$\begin{aligned} \mathcal{C}_{\mathbf{q}_b\mathbf{q}_a} &= -\mathcal{C}^T_{\mathbf{q}_b\mathbf{q}_a} \qquad \mathcal{C}_{\mathbf{q}_c\mathbf{q}_d} = -\mathcal{C}^T_{\mathbf{q}_c\mathbf{q}_d} \\ \mathcal{C}_{\mathbf{q}_c\mathbf{q}_a} &= -\mathcal{C}^T_{\mathbf{q}_c\mathbf{q}_a} \qquad \mathcal{C}_{\mathbf{q}_d\mathbf{q}_b} = -\mathcal{C}^T_{\mathbf{q}_d\mathbf{q}_b} \\ \mathcal{C}_{\mathbf{q}_d\mathbf{q}_a} &= \mathcal{C}^T_{\mathbf{q}_d\mathbf{q}_a} \qquad \mathcal{C}_{\mathbf{q}_b\mathbf{q}_c} = \mathcal{C}^T_{\mathbf{q}_b\mathbf{q}_c} \end{aligned} \tag{29}$$

Observe that $\mathcal{C}_{\mathbf{q}_b\mathbf{q}_a} = -\mathcal{C}^T_{\mathbf{q}_b\mathbf{q}_a}$ for (28)-(29), whereas $\mathcal{C}_{\mathbf{q}_b\mathbf{q}_a} = \mathcal{C}^T_{\mathbf{q}_b\mathbf{q}_a}$ for (27), meaning 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. In other words, the conditions (27)-(29) mean 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, means that all the four individual components of each quaternion variable q_ℓ are uncorrelated $\big($property P2 in Table III$\big)$. This also means that the components of q_ℓ and q_ϱ are uncorrelated for $\ell \neq \varrho$ (in contrast to the complex case [11], [12]). We can therefore conclude that 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 III, that the covariance matrices (12-(15)) 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 $\mathcal{C}^a_{\mathbf{q}}$ of a $\mathbb{Q}$-proper random vector, given by viz.

$$\mathcal{C}^a_{\mathbf{q}} = \begin{bmatrix} \mathcal{C}_{\mathbf{qq}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathcal{C}_{\mathbf{q}^\imath\mathbf{q}^\imath} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathcal{C}_{\mathbf{q}^\jmath\mathbf{q}^\jmath} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathcal{C}_{\mathbf{q}^\kappa\mathbf{q}^\kappa} \end{bmatrix} = 4\sigma^2\mathbf{I} \tag{30}$$

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

VII. 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 [13]; this was further elaborated by Picinbono [30]. In the same spirit, we address probability distributions of both proper and improper processes in $\mathbb{H}$, and propose a generic Gaussian distribution for multivariate quaternion valued variables.

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

$$p(\mathbf{q}_a, \mathbf{q}_b, \mathbf{q}_c, \mathbf{q}_d) = \frac{1}{(2\pi)^{2N}\det(\mathcal{C}_R)^{\frac{1}{2}}} \exp\{-\frac{1}{2} f(\mathbf{q}_a, \mathbf{q}_b, \mathbf{q}_c, \mathbf{q}_d)\} \tag{31}$$

where

$$f(\mathbf{q}_a, \mathbf{q}_b, \mathbf{q}_c, \mathbf{q}_d) = \mathbf{q}^{rT}\mathcal{C}_R^{-1}\mathbf{q}^r = \mathbf{q}^{rH}\mathcal{C}_R^{-1}\mathbf{q}^r \tag{32}$$

It is assumed that $\mathbf{q}_a$, $\mathbf{q}_b$,$\mathbf{q}_c$ and $\mathbf{q}_d \in \mathbb{R}^{N\times 1}$ have zero mean, but this does not restrict the generality of the results. To make the Gaussian distribution (31) cater for both $\mathbb{Q}$-proper and $\mathbb{Q}$-improper signals, we need to express it in terms of the augmented quaternion valued vector $\mathbf{q}^a$ (9). To this end, the determinant of the quadrivariate covariance $\mathcal{C}_R$ and the quadratic function (32) need to be further investigated.

To examine the duality between the real-valued quadrivariate matrix $\mathcal{C}_R$ (20), and the augmented quaternion-valued covariance $\mathcal{C}_{\mathbf{q}}^{a}$ from (21), we shall first express the determinant of $\mathcal{C}_R$ as a function of $\mathcal{C}_{\mathbf{q}}^{a}$, that is

$$\begin{aligned} \det(\mathcal{C}_R) &= \det(\mathbf{A}^{-1}\mathcal{C}_{\mathbf{q}}^{a}\mathbf{A}^{-H}) && (33) \\ &= \det(\mathbf{A}^{-1})\det(\mathcal{C}_{\mathbf{q}}^{a})\det(\mathbf{A}^{-H}) && (34) \end{aligned}$$

where $\mathbf{A}$ is given in (9). From (11), $\det(\mathbf{A}) = 16^N$ and since $\det(A^{-1}) = \det(A)^{-1}$, the above expression can be further simplified to

$$\det(\mathcal{C}_R) = \left(\frac{1}{16}\right)^{2N} \det(\mathcal{C}_{\mathbf{q}}^{a}) \tag{35}$$

The quadratic function (32) can be also expressed as a function of the augmented quaternion-valued random vector $\mathbf{q}^a$, given by

$$\begin{aligned} f(\mathbf{q}_a, \mathbf{q}_b, \mathbf{q}_c, \mathbf{q}_d) &= \mathbf{q}^{rH}\mathcal{C}_R^{-1}\mathbf{q}^r \\ &= \Big(\mathbf{q}^{aH}\mathbf{A}^{-H}\Big)\Big(A^H\mathcal{C}_{\mathbf{q}}^{a^{-1}}\mathbf{A}\Big)\Big(\mathbf{A}^{-1}\mathbf{q}^a\Big) \\ &= \mathbf{q}^{aH}\mathcal{C}_{\mathbf{q}}^{a^{-1}}\mathbf{q}^a = f(\mathbf{q}, \mathbf{q}^*, \mathbf{q}^{\imath *}, \mathbf{q}^{\jmath *}) && (36) \end{aligned}$$

By substituting (35) and (36) into (31), we can express the Gaussian probability density function for an augmented multivariate quaternion-valued random vector $\mathbf{q}^a$ as

$$p(\mathbf{q}^a) = p(\mathbf{q}, \mathbf{q}^*, \mathbf{q}^{\imath *}, \mathbf{q}^{\jmath *}) = \frac{1}{(\pi^2/4)^N \det(\mathcal{C}_{\mathbf{q}}^{a})^{\frac{1}{2}}} \exp\{-\frac{1}{2}\mathbf{q}^{aH}\mathcal{C}_{\mathbf{q}}^{a^{-1}}\mathbf{q}^a\} \tag{37}$$

For a $\mathbb{Q}$-proper vector, it can be shown $\big($using (30)$\big)$ that the Gaussian distribution (37) simplifies to

$$p(\mathbf{q}, \mathbf{q}^*, \mathbf{q}^{\imath *}, \mathbf{q}^{\jmath *}) = \frac{1}{(2\pi\sigma^2)^{2N}} \exp\left\{-\frac{1}{2\sigma^2}\mathbf{q}^H\mathbf{q}\right\} \tag{38}$$

that is, the argument in the exponential is a function of only $|\mathbf{q}|$, thus highlighting the correspondence with the real and proper complex Gaussian distributions [14].

VIII. A NOTE ON INFORMATION THEORETIC MEASURES

A. Entropy for Quaternion-valued Random Vectors

Based on Section VII and the results in [11], we can now generalise the maximum entropy principle to the quaternion-valued multivariate case [11], [31]. The 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 X-C)

$$H(\mathbf{q}) = \log\left[(\pi e/2)^{2N} \det(\mathcal{C}^a_{\mathbf{q}})^{\frac{1}{2}}\right] \tag{39}$$

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

$$H(\mathbf{q}) \leq 2N \log\left[(2\pi e \sigma^2)\right] = H_{\text{proper}} \tag{40}$$

The equality (40) holds, if and only if, $\mathbf{q}$ is a centered $\mathbb{Q}$-proper Gaussian random vector (as shown in Appendix X-D). It is straightforward to show that the entropy of a quaternion random vector with an arbitrary probability density function $p_A(\mathbf{q})$ cannot be greater than that of the Gaussian distribution[7] (40), thus confirming that a $\mathbb{Q}$-proper Gaussian process attains the upper entropy limit, as shown in Appendix X-D. In addition, the difference in entropy is due to the improperness of a quaternion-valued Gaussian random vector can be quantified by the difference between (40) and (39).

B. Beyond Mutual Information - Interaction Information

Another important information theoretic measure is mutual information (MI). Standard MI considers only two variables, and we next provide its generalisation to higher dimensions using the so-called 'interaction information' $\mathcal{I}$ [33]. Unlike mutual information, 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 $\mathcal{I} < 0$ [33]. 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

$$\begin{aligned} \mathcal{I}(\mathbf{q};\mathbf{q}^{\imath};\mathbf{q}^{\jmath};\mathbf{q}^{\kappa}) \quad = \quad & \log\left[\frac{(8\sigma^4)^N}{\det(\mathcal{C}^a_{\mathbf{q}})^{\frac{1}{2}}}\right] \\ & + H(\mathbf{q}_a,\mathbf{q}_b,\mathbf{q}_c) + H(\mathbf{q}_a,\mathbf{q}_b,\mathbf{q}_d) + H(\mathbf{q}_a,\mathbf{q}_c,\mathbf{q}_d) + H(\mathbf{q}_c,\mathbf{q}_b,\mathbf{q}_d) \\ & - H(\mathbf{q}_a,\mathbf{q}_b) - H(\mathbf{q}_a,\mathbf{q}_c) - H(\mathbf{q}_a,\mathbf{q}_d) - H(\mathbf{q}_b,\mathbf{q}_c) - H(\mathbf{q}_b,\mathbf{q}_d) - H(\mathbf{q}_c,\mathbf{q}_d) \end{aligned} \tag{41}$$

which clearly attains the value of $\mathcal{I} = 0$, for $\mathbb{Q}$-proper signals. The derivation is included in Appendix X-E.

[7] The proof given on see p. 336 of [32].

IX. Conclusion

Second order statistics and information theoretic measures for general quaternion-valued random variables and processes have been revisited. To make use of complete information within quaternion-valued second order statistics, complementary statistical measures 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 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 findings.

X. Appendix

A. The complete description of second order statistics with an alternative basis $\{q, q^*, q^{\imath *}, q^{\jmath *}\}$

We can express the componentwise 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 $\mathbf{q}$ in terms of the quaternion-valued covariance and pseudocovariance matrices $\mathcal{P}_\mathbf{q} = E\{\mathbf{q}\mathbf{q}^T\}$, $\mathcal{P}^\imath_\mathbf{q} = E\{\mathbf{q}\mathbf{q}^{\imath T}\}$, $\mathcal{P}^\jmath_\mathbf{q} = E\{\mathbf{q}\mathbf{q}^{\jmath T}\}$ as

$$\begin{aligned}
\mathcal{C}_{\mathbf{q}_a} &= \frac{1}{2}\Re\{\mathcal{C}_\mathbf{q} + \mathcal{P}_\mathbf{q}\} & \mathcal{C}_{\mathbf{q}_b} &= \frac{1}{2}\Re\{\mathcal{C}_\mathbf{q} - \mathcal{P}^\imath_\mathbf{q}\} \\
\mathcal{C}_{\mathbf{q}_c} &= \frac{1}{2}\Re\{\mathcal{C}_\mathbf{q} - \mathcal{P}^\jmath_\mathbf{q}\} & \mathcal{C}_{\mathbf{q}_d} &= \Re\{\mathcal{C}_\mathbf{q}\} - \big(\mathcal{C}_{\mathbf{q}_a} + \mathcal{C}_{\mathbf{q}_b} + \mathcal{C}_{\mathbf{q}_c}\big) \\
\mathcal{C}_{\mathbf{q}_b\mathbf{q}_a} &= \frac{1}{2}\Im_i\{\mathcal{C}_\mathbf{q} + \mathcal{P}_\mathbf{q}\} & \mathcal{C}_{\mathbf{q}_c\mathbf{q}_a} &= \frac{1}{2}\Im_j\{\mathcal{C}_\mathbf{q} + \mathcal{P}_\mathbf{q}\} \\
\mathcal{C}_{\mathbf{q}_d\mathbf{q}_a} &= \frac{1}{2}\Im_k\{\mathcal{C}_\mathbf{q} + \mathcal{P}_\mathbf{q}\} & \mathcal{C}_{\mathbf{q}_c\mathbf{q}_b} &= \frac{1}{2}\Im_k\{\mathcal{C}_\mathbf{q} - \mathcal{P}^\imath_\mathbf{q}\} \\
\mathcal{C}_{\mathbf{q}_d\mathbf{q}_c} &= \frac{1}{2}\Im_i\{\mathcal{C}_\mathbf{q} - \mathcal{P}^\jmath_\mathbf{q}\} & \mathcal{C}_{\mathbf{q}_d\mathbf{q}_b} &= -\frac{1}{2}\Im_j\{\mathcal{C}_\mathbf{q} - \mathcal{P}^\kappa_\mathbf{q}\}
\end{aligned} \tag{42}$$

This illustrates the validity of the above basis in augmented quaternion valued statistics.

B. Invariance of $\mathbb{Q}$*-proper random vectors under an affine or linear transformation*

Consider an affine process $\mathbf{y} = \mathbf{A}\mathbf{q} + \mathbf{b}$, where $\mathbf{q}$ is a $\mathbb{Q}-$proper random vector $\in \mathbb{H}^N$, $\mathbf{A} \in \mathbb{H}^{M\times N}$, and $\mathbf{b} \in \mathbb{H}^M$ are constant. Based on the proof of Lemma 3 of [11] and the special properties of complementary covariance matrices in (16), $\mathbf{y}$ is also a $\mathbb{Q}-$proper random vector, as shown by

$$\begin{aligned} \mathcal{C}_{\mathbf{y}^\imath} &= E\{\mathbf{y}\mathbf{y}^{\imath H}\} = \mathbf{A}\mathcal{C}_{\mathbf{q}^\imath}\mathbf{A}^{\imath H} = \mathbf{0} \\ \mathcal{C}_{\mathbf{y}^\jmath} &= E\{\mathbf{y}\mathbf{y}^{\jmath H}\} = \mathbf{A}\mathcal{C}_{\mathbf{q}^\jmath}\mathbf{A}^{\jmath H} = \mathbf{0} \\ \mathcal{C}_{\mathbf{y}^\kappa} &= E\{\mathbf{y}\mathbf{y}^{\kappa H\}} = \mathbf{A}\mathcal{C}_{\mathbf{q}^\kappa}\mathbf{A}^{\kappa H} = \mathbf{0} \end{aligned}$$

C. 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 (37). For convenience (with a slight abuse of notation), we denote $\int\int\int\int\cdots\int\int\int\int$ by $\oint$ 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 $d\mathbf{q}$

$$\begin{aligned} \oint p_A(\mathbf{q})\log\big[\frac{1}{p(\mathbf{q})}\big]d\mathbf{q} &= \oint p_A(\mathbf{q})\log\big[(\pi^2/4)^N\det(\mathcal{C}^a_{\mathbf{q}})^{\frac{1}{2}}\exp\{\frac{1}{2}\mathbf{q}^{a^H}\mathcal{C}_{\mathbf{q}}^{a^{-1}}\mathbf{q}^a\}\big]d\mathbf{q} \\ &\approx \oint p_A(\mathbf{q})\log\big[(\pi^2/4)^N\det(\mathcal{C}^a_{\mathbf{q}})^{\frac{1}{2}}\exp\{2N\}\big]d\mathbf{q} \\ &\approx \log\big[(\pi^2e^2/4)^N\det(\mathcal{C}^a_{\mathbf{q}})^{\frac{1}{2}}\big]\oint p_A(\mathbf{q})d\mathbf{q} \\ &\approx \log\big[(\pi e/2)^{2N}\det(\mathcal{C}^a_{\mathbf{q}})^{\frac{1}{2}}\big] \end{aligned} \tag{43}$$

For a $\mathbb{Q}$-proper Gaussian random vector, the augmented covariance matrix has the special structure (30), its determinant is $\det(\mathcal{C}^a_{\mathbf{q}}) = (4\sigma^2)^{4N}$, and the expression (43) can be further simplified into

$$H_{\text{proper}} = 2N\log\big[(2\pi e\sigma^2)\big] \tag{44}$$

D. 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 [31])

$$H(\mathbf{q}^r_s) \le \frac{1}{2}\log\big[(2\pi e)^4\det(\mathcal{C}_R)\big] \tag{45}$$

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

$$\mathcal{C}_R = \begin{bmatrix} \mathbf{\Gamma} & \mathbf{B} \\ \mathbf{B}^T & \mathbf{C} \end{bmatrix} \tag{46}$$

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

$$\begin{aligned} \det(\mathcal{C}_R) &= \det(\mathbf{\Gamma})\det(\mathbf{C} - \mathbf{B}^T\mathbf{\Gamma}^{-1}\mathbf{B}) \\ &\leq \det(\mathbf{\Gamma})\det(\mathbf{C}) \end{aligned} \tag{47}$$

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

$$E\{q_a q_c\} = E\{q_a q_d\} = E\{q_b q_c\} = E\{q_b q_d\} = 0 \tag{48}$$

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

$$\begin{aligned} \det(\mathcal{C}_R) &\leq \left[E\{q_a^2\}E\{q_b^2\} - E\{q_a q_b\}^2\right]\left[E\{q_c^2\}E\{q_d^2\} - E\{q_c q_d\}^2\right] \\ &\leq E\{q_a^2\}E\{q_b^2\}E\{q_c^2\}E\{q_d^2\} \underbrace{+E\{q_a q_b\}^2 E\{q_c q_d\}^2}_{\phi} + \\ &\quad \underbrace{-E\{q_a q_b\}^2 E\{q_c^2\}E\{q_d^2\}}_{\chi} \underbrace{-E\{q_c q_d\}^2 E\{q_a^2\}E\{q_b^2\}}_{\tau} \end{aligned} \tag{49}$$

By examining (49), and factorising ϕ and χ as

$$\phi + \chi = E\{q_a q_b\}^2 \left[E\{q_c q_d\}^2 - E\{q_c^2\}E\{q_d^2\}\right] \leq 0 \tag{50}$$

the maximum value of $\phi + \chi = 0$ indicates that either

$$E\{q_a q_b\}^2 = 0 \tag{51}$$

or

$$E\{q_c q_d\}^2 = 0 \tag{52}$$

The same statement can be made for $\phi + \tau \leq 0$. Therefore, equations (48), and (51)-(52) satisfy property P2 of a $\mathbb{Q}$-proper variable (see Table III), and the determinant of $\mathcal{C}_R$ is upper bounded by

$$\det(\mathcal{C}_R) \leq E\{q_a^2\}E\{q_b^2\}E\{q_c^2\}E\{q_d^2\} \tag{53}$$

Using constrained equality based optimisation (Lagrange multipliers), we show below that inequality (53) is maximised when condition P1 of $\mathbb{Q}$-properness in Table III is satisfied, yielding

$$\det(\mathcal{C}_R) \leq \left(\frac{E\{|q|^2\}}{4}\right)^4 \tag{54}$$

This optimisation problem can be posed as

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

and can be solved using Lagrange multipliers as

$$f\Big(E\{q_a^2\},E\{q_b^2\},E\{q_c^2\},E\{q_d^2\},\lambda\Big) = E\{q_a^2\}E\{q_b^2\}E\{q_c^2\}E\{q_d^2\}+ \lambda\Big(E\{q_a^2\}+E\{q_b^2\}+E\{q_c^2\}+E\{q_d^2\}-E\{|q|^2\}\Big) \tag{55}$$

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

$$\frac{\partial f}{\partial E\{q_a^2\}} = E\{q_b^2\}E\{q_c^2\}E\{q_d^2\}+\lambda = 0 \tag{56}$$
$$\frac{\partial f}{\partial E\{q_b^2\}} = E\{q_a^2\}E\{q_c^2\}E\{q_d^2\}+\lambda = 0 \tag{57}$$
$$\frac{\partial f}{\partial E\{q_c^2\}} = E\{q_a^2\}E\{q_b^2\}E\{q_d^2\}+\lambda = 0 \tag{58}$$
$$\frac{\partial f}{\partial E\{q_d^2\}} = E\{q_a^2\}E\{q_b^2\}E\{q_c^2\}+\lambda = 0 \tag{59}$$
$$\frac{\partial f}{\partial \lambda} = E\{q_a^2\}+E\{q_b^2\}+E\{q_c^2\}+E\{q_d^2\}-E\{|q|^2\} = 0 \tag{60}$$

Solving the equations (56)-(59) leads to

$$E\{q_a^2\} = E\{q_b^2\} = E\{q_c^2\} = E\{q_d^2\} \tag{61}$$

which when replaced in (60) yields the solution

$$E\{q_a^2\} = E\{q_b^2\} = E\{q_c^2\} = E\{q_d^2\} = \frac{E\{|q|^2\}}{4} \tag{62}$$

Since the function $\log(\cdot)$ is monotically increasing, we can substitute the maximum value of $\det(\mathcal{C}_R)$ from (54) into (45), to obtain the upper entropy bound in the form

$$\begin{aligned} H(\mathbf{q}_s^r) &\leq \log\left[\frac{\big(\pi^2 e^2 E\{|q|^2\}^2}{4}\right] \\ &\leq \log\Big[4\pi^2 e^2 \sigma^4\Big] \end{aligned} \tag{63}$$

This upper bound is equivalent to the entropy of a $\mathbb{Q}$-proper Gaussian quaternion random variable (40) 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 (37) for quaternion random variables.

E. Interaction Information $\mathcal{I}(\mathbf{q};\mathbf{q}^{\imath};\mathbf{q}^{\jmath};\mathbf{q}^{\kappa})$

Prior to the formulation of $\mathcal{I}(\mathbf{q};\mathbf{q}^{\imath};\mathbf{q}^{\jmath};\mathbf{q}^{\kappa})$, note that the interaction information $\mathcal{I}$ of $\mathbf{q}^a = [\mathbf{q}^T \ \mathbf{q}^{\imath T} \ \mathbf{q}^{\jmath T} \ \mathbf{q}^{\kappa T}]^T \ \in \mathbb{H}^{4N\times 1}$ is equivalent to that of $\mathbf{q}^r = [\mathbf{q}_a^T \ \mathbf{q}_b^T \ \mathbf{q}_c^T \ \mathbf{q}_d^T]^T \ \in \mathbb{R}^{4N\times 1}$, due to their deterministic relationship

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

The matrix $\mathbf{A}$ does not contribute to the interaction information of $\mathbf{q}^a$, and therefore,

$$\begin{aligned}
\mathcal{I}(\mathbf{q};\mathbf{q}^{\imath};\mathbf{q}^{\jmath};\mathbf{q}^{\kappa}) &= \mathcal{I}(\mathbf{q}_a;\mathbf{q}_b;\mathbf{q}_c;\mathbf{q}_d) \\
&= \underbrace{H(\mathbf{q}_a)+H(\mathbf{q}_b)+H(\mathbf{q}_c)+H(\mathbf{q}_d)-H(\mathbf{q}_a,\mathbf{q}_b,\mathbf{q}_c,\mathbf{q}_d)}_{H_{\text{proper}}-H(\mathbf{q})} \\
&\quad +H(\mathbf{q}_a,\mathbf{q}_b,\mathbf{q}_c)+H(\mathbf{q}_a,\mathbf{q}_b,\mathbf{q}_d)+H(\mathbf{q}_a,\mathbf{q}_c,\mathbf{q}_d)+H(\mathbf{q}_c,\mathbf{q}_b,\mathbf{q}_d) \\
&\quad -H(\mathbf{q}_a,\mathbf{q}_b)-H(\mathbf{q}_a,\mathbf{q}_c)-H(\mathbf{q}_a,\mathbf{q}_d)-H(\mathbf{q}_b,\mathbf{q}_c)-H(\mathbf{q}_b,\mathbf{q}_d)-H(\mathbf{q}_c,\mathbf{q}_d) \\
&= \log\left[\frac{(8\sigma^4)^N}{\det(\mathcal{C}^a_{\mathbf{q}})^{\frac{1}{2}}}\right] \\
&\quad +H(\mathbf{q}_a,\mathbf{q}_b,\mathbf{q}_c)+H(\mathbf{q}_a,\mathbf{q}_b,\mathbf{q}_d)+H(\mathbf{q}_a,\mathbf{q}_c,\mathbf{q}_d)+H(\mathbf{q}_c,\mathbf{q}_b,\mathbf{q}_d) \\
&\quad -H(\mathbf{q}_a,\mathbf{q}_b)-H(\mathbf{q}_a,\mathbf{q}_c)-H(\mathbf{q}_a,\mathbf{q}_d)-H(\mathbf{q}_b,\mathbf{q}_c)-H(\mathbf{q}_b,\mathbf{q}_d)-H(\mathbf{q}_c,\mathbf{q}_d)
\end{aligned} \tag{64}$$

REFERENCES

[1] T. K. Tsui, X. P. Zhang, and D. Androutsos, “Color Image Watermarking Using Multidimensional Fourier Transforms,” *IEEE Transactions on Information Forensics and Security*, vol. 3, pp. 16–28, 2008.

[2] X. Yun, I. E. Bachmann, and B. McGhee, “A simplified quaternion-based algorithm for orientation estimation from earth gravity and magnetic field measurements,” *IEEE Transactions on Instrumentation and Measurement*, vol. 57, no. 3, pp. 638–650, 2008.

[3] B. A. Sethuraman, B. S. Rajan, and V. Shashidhar, “Full-diversity, high-rate space-time block codes from division algebras,” *IEEE Transactions on Information Theory*, vol. 49, pp. 2596–2616, 2003.

[4] J. C. K. Chou, “Quaternion kinematic and dynamic differential equations,” *IEEE Transactions on Robotics and Automation*, vol. 8, no. 1, pp. 53 – 64, 1992.

[5] L. Fortuna, G. Muscato, and M. G. Xibilia, “A comparison between HMLP and HRBF for attitude control,” *IEEE Transactions on Neural Networks*, vol. 12, no. 2, pp. 318–328, 2001.

[6] G. Antonelli, F. Caccavale, S. Chiaverini, and L. Villani, “Tracking control for underwater vehicle-manipulator systems with velocity estimation,” *IEEE Journal of Oceanic Engineering*, vol. 25, pp. 399–413, Jul 2000.

[7] D. Choukkroun, I. Y. Bar-Itzhack, and Y. Oshman, “Novel Quaternion Kalman Filter,” *IEEE Transactions on Aerospace and Electronics Systems*, vol. 42, no. 1, pp. 174–190, 2006.

[8] S. Miron, N. L. Bihan, and J. Mars, “Quaternion-MUSIC for vector-sensor array processing,” *IEEE Transactions on Signal Processing*, vol. 54, no. 4, pp. 1218– 1229, 2006.

[9] N. L. Bihan and J. Mars, “Singular Value Decomposition of Quaternion Matrices: A New Tool for Vector-Sensor Signal Processing,” *Signal Processing*, vol. 84, no. 7, pp. 1177 – 1199, 2004.

[10] C. Cheong-Took and D. P. Mandic, “The Quaternion LMS Algorithm for Adaptive Filtering of Hypercomplex Real World Processes,” *IEEE Transactions on Signal Processing*, vol. 57, no. 4, pp. 1316–1327, 2009.

[11] F. Neeser and J. Massey, “Proper Complex Random Processes With Applications to Information Theory,” *IEEE Transactions on Information Theory*, vol. 39, no. 4, pp. 1293–1302, 1993.

[12] B. Picinbono, “On Circularity,” *IEEE Transactions on Signal Processing*, vol. 42, no. 12, pp. 3473–3482, 1994.

[13] A. V. D. Bos, “The Multivariate Complex Normal Distribution - A Generalization,” *IEEE Transactions on Information Theory*, vol. 41, no. 2, pp. 537 – 539, 1995.

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

[15] M. Valkama and V. Koivunen, “Circular and Non-Circular Complex random Signals with Applications in Wireless Communications,” *IEEE International Conference on Acoustics, Speech, and Signal Processing (Tutorial)*, April 2007.

[16] J. Navarro-Moreno, “ARMA Prediction of Widely Linear Systems by Using the Innovations Algorithm,” *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 3061–3068, 2008.

[17] J. Eriksson and V. Koivunen, “Complex Random Vectors and ICA Models: Identifiability, Uniqueness, and Separability,” *IEEE Transactions on Information Theory*, vol. 52, no. 3, pp. 1017–1029, 2006.

[18] N. N. Vakhania, “Random vectors with values in quaternion Hilbert spaces,” *Theory of Probability and its Applications*, vol. 43, no. 1, pp. 99–115, 1999.

[19] P. O. Amblard and N. LeBihan, “On properness of quaternion valued random variables,” *in Proceedings of International Conference on Mathematics (IMA) in Signal Processing, Cirencester*, pp. 23–26, 2004.

[20] C. E. Moxey and S. J. Sangwine, “Hypercomplex Correlation Techniques for Vector Images,” *IEEE Transactions on Signal Processing*, vol. 51, no. 7, pp. 1941 – 1953, 2003.

[21] P. O. Amblard, M. Gaeta, and J. L. Lacoume, “Statistics for complex variables and signalsPart I: variables,” *Signal Processing (Elsevier)*, vol. 53, pp. 1–13, 1996.

[22] P. O. Amblard, M. Gaeta, and J. L. Lacoume, “Statistics for complex variables and signalsPart II: signals,” *Signal Processing (Elsevier)*, vol. 53, pp. 15–25, 1996.

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

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

[25] T. A. Ell and S. J. Sangwine, “Quaternion Involutions and Anti-Involutions,” *Computers & Mathematics with Applications*, vol. 53, no. 1, pp. 137–143, 2007.

[26] B. Picinbono, *Random Signals and Systems*. Prentice Hall, 1993.

[27] B. Picinbono and P. Chevalier, “Widely linear estimation with complex data,” *IEEE Transactions on Signal Processing*, vol. 43, no. 8, pp. 2030–2033, 1995.

[28] Y. Huang and J. Benesty, *Audio Signal Processing for Next Generation Multimedia Communication Systems*. Kluwer Academic Publishers, 2004.

[29] W. A. Gardner, *Introduction to Random Processes with Applications to Signals and Systems*. New York: McGraw-Hill (2nd Ed.), 1990.

[30] B. Picinbono, "Second-Order Complex Vectors and Normal Distributions," *IEEE Transactions on Signal Processing*, vol. 44, no. 10, pp. 2637–2640, 1996.

[31] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley: New York, 1991.

[32] R. G. Gallager, *Information Theory and Reliable Communications*. Wiley: New York, 1968.

[33] W. J. McGill, "Multivariate information transmission," *IEEE Transactions Information Theory*, vol. 4, no. 4, pp. 93–111, 1954.

[34] R. A. Horn, *Matrix Analysis*. Cambridge University Press, 1990.