Proceedings of the International Symposium on Evolving Intelligent Systems, Plamen Angelov, Dimitar Filev, and Nikola Kasabov (Eds.), at the AISB 2010 convention, 29 March – 1 April 2010, De Montfort University, Leicester, UK

New Developments in Statistical Signal Processing of Quaternion Random Variables with Applications in Wind Forecasting

C. Cheong Took and **D. P. Mandic**[1]

Abstract. Quaternions have found applications in various machine learning algorithms, however these algorithms usually do not exploit the complete available second order statistics of quaternions. To this end, we present the so-called quaternion 'augmented' statistics to show how to make use of the complete second order statistical information available within the quaternion domain $\mathbb{H}$. Next, as a case study, the widely linear model, which operates on the quaternion augmented statistics, is employed to enhance the performance of the Quaternion Least Mean Square algorithm. Simulations on time series prediction of both chaotic and non-stationary real world data support the approach.

1 Introduction

Recently, the resurgence in quaternion signal processing has arisen due to the advantages of quaternion algebra over real-valued quadrivariate vector algebra in the modelling of such data. In the machine learning community, quaternions have been employed in robotics, neural networks for chaotic time series prediction [2], Kalman filtering for aeronautic applications [6] and the least-mean-square estimation for the adaptive filtering of hypercomplex processes [5].

As quaternions are considered as the four-dimensional extension of complex numbers, we 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. Pioneering works in augmented complex statistics have been undertaken by Neeser and Massey, who formulated the concept of properness (rotation invariant probability distribution) into complex-valued statistics. These researchers showed that the covariance $E\{\mathbf{z}\mathbf{z}^H\}$ of a complex random vector $\mathbf{z}$ alone is not sufficient to have a complete second order statistical description [9] and thus the pseudocovariance $E\{\mathbf{z}\mathbf{z}^T\}$ should also be considered to cater for improper signals. These theoretical findings have given way to a panoply of novel algorithms in adaptive signal processing, autoregressive moving average (ARMA) modelling, and independent component analysis [8].

Existing techniques typically consider only the information contained in quaternion-valued covariance [2,6] and therefore, they are bound not to maximise the available statistical information. This is because quaternion-valued statistics are still their infancy [1,11]. To shed light on quaternion statistics, Vakhania extended the concept of 'properness' to the quaternion domain, however, his definition of $\mathbb{Q}$-properness was limited to the invariance of the probability density function (pdf) under rotations of around angle of $\pi/2$ [11]. Amblard and Le Bihan improved the conditions of $\mathbb{Q}$-properness to an arbitrary axis and angle of rotation φ, that is [1] $q \triangleq e^{\nu\varphi} q \quad \forall\, \varphi$, for any pure unit quaternion ν (whose real part vanishes); symbol $\triangleq$ means equality in terms of pdf. These results provide an initial insight into $\mathbb{Q}$-properness, however they are applicable to only single quaternion variables and furthermore, it is not straightforward to apply them to quaternion-valued processes.

To that end, this work presents a generic framework for second order statistical analysis of the generality of quaternion-valued random variables and vectors, both second order circular and noncircular [4]. It is shown that in order to take advantage of the complete second order information, it is crucial to consider complementary covariance matrices, thus accounting for a possible improperness of quaternion processes. The advantages of the technique presented herein are likely to be similar to the benefits of the augmented statistics for noncircular complex-valued data [8]. Our analysis illustrates that the basis for augmented quaternion statistics should include quaternion involutions, and that all the necessary second order statistical information are within the so introduced augmented covariance matrix. As a proof of concept of the augmented quaternion statistics, we formulate the widely linear model in $\mathbb{H}$ to incorporate such statistics and illustrate its advantages, such as enhanced performance and improved convergence and steady state properties of the learning algorithms [3]. The theoretical findings presented herein are likely to benefit the next generation evolutionary learning algorithms.

2 Properties of Quaternion Random Vectors

2.1 Quaternion Algebra

The quaternion domain can be considered as a non-commutative four-dimensional extension of complex numbers. A quaternion variable $q \in \mathbb{H}$ consists of a real part $\Re\{\cdot\}$ (denoted by subscript a) and a vector-part $\Im\{\cdot\}$ of three imaginary components (denoted by subscripts b, c, and d), and can be expressed as:

$$\begin{aligned} q &= \Re\{q\} + \imath\Im_i\{q\} + \jmath\Im_j\{q\} + \kappa\Im_k\{q\} \\ &= q_a + \imath q_b + \jmath q_c + \kappa q_d \quad \in \mathbb{H} \end{aligned} \tag{1}$$

The quaternion conjugate is defined as

$$\begin{aligned} q^* &= \Re\{q\} - \Im\{q\} \\ &= q_a - \imath q_b - \jmath q_c - \kappa q_d \end{aligned} \tag{2}$$

[1] Department of Electrical and Electronic Engineering, Imperial College London, SW7 2AZ, U.K.; E-mail: {c.cheong-took, d.mandic}@imperial.ac.uk

The orthogonal unit vectors $\imath$, $\jmath$, κ also considered as imaginary numbers, obey the following product rules

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

Another important algebraic manipulation is the quaternion product, given as

$$q_1 q_2 = \Re\{q_1 q_2\} + \Im\{q_1 q_2\} \tag{4}$$

where

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

where the symbol "$\times$" denotes the cross product; Notice the quaternion product inherits the non-commutativity property of the cross product.

2.2 Quaternion Involutions and the Augmented Basis Vector

It is well-known that the real and imaginary part of a complex number $z = z_a + \imath z_b$ to be extracted as $z_a = \frac{1}{2}(z + z^*)$ and $z_b = \frac{1}{2\imath}(z - z^*)$. We therefore need to employ both z and z^* to describe the elements of the corresponding signal in $\mathbb{R}^2$ and they are used as a basis for the augmented complex statistics, where the 'augmented' basis vector is given by $[z\ z^*]^T$. On the other hand, the quaternion domain does not allow such convenient operation and the correspondence between the elements of a quaternion valued variable in $\mathbb{H}$ and those of a quadrivariate vector in $\mathbb{R}^4$ is not clear. To circumvent this issue, we can utilise the three perpendicular quaternion involutions (self-inverse mappings), given by

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

Based on (5), we can compute the four components of the quaternion variable q as [7]

$$\begin{aligned} q_a &= \frac{1}{2}(q + q^*) \qquad & q_b &= \frac{1}{2\imath}(q - q^{\imath *}) \\ q_c &= \frac{1}{2\jmath}(q - q^{\jmath *}) \qquad & q_d &= \frac{1}{2\kappa}(q - q^{\kappa *}) \end{aligned} \tag{6}$$

Observe that the quaternion conjugate operation $(\cdot)^*$ is also an involution, and can be expressed as

$$q^* = \frac{1}{2}(q^\imath + q^\jmath + q^\kappa - q) \tag{7}$$

Prior to the introduction of quaternion 'augmented' statistics, we first demonstrate the duality between the augmented quaternion vector $\mathbf{q}^a \in \mathbb{H}^{4N\times 1}$ and the quadrivariate real valued vectors in $\mathbb{R}^{4N\times 1}$. To ensure that the augmented statistics in $\mathbb{H}$ are suitable for the description of both second order circular and noncircular signals, following on (see pp. 118-119 [10]), we provide a one-to-one correspondence between the components of a quaternion variable and its quadrivariate real counterpart. For convenience, we shall utilise a combination[2] of $\{q, q^\imath, q^\jmath, q^\kappa\}$, 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

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

$$\begin{bmatrix} \mathbf{q} \\ \mathbf{q}^\imath \\ \mathbf{q}^\jmath \\ \mathbf{q}^\kappa \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \imath\mathbf{I} & \jmath\mathbf{I} & \kappa\mathbf{I} \\ \mathbf{I} & \imath\mathbf{I} & -\jmath\mathbf{I} & -\kappa\mathbf{I} \\ \mathbf{I} & -\imath\mathbf{I} & \jmath\mathbf{I} & -\kappa\mathbf{I} \\ \mathbf{I} & -\imath\mathbf{I} & -\jmath\mathbf{I} & \kappa\mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{q}_a \\ \mathbf{q}_b \\ \mathbf{q}_c \\ \mathbf{q}_d \end{bmatrix} \tag{8}$$

where $\mathbf{I} \in \mathbb{R}^{N\times N}$ is the identity matrix, and $\mathbf{q} = [q_1\ q_2 \cdots q_N]^T \in \mathbb{H}^{N\times 1}$; similarly for $\mathbf{q}^\imath$, $\mathbf{q}^\jmath$, $\mathbf{q}^\kappa \in \mathbb{H}^{N\times 1}$, and $\mathbf{q}_a$, $\mathbf{q}_b$, $\mathbf{q}_c$ and $\mathbf{q}_d \in \mathbb{R}^{N\times 1}$. The invertible $4N \times 4N$ matrix $\mathbf{A}$ provides a 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 given by

$$\mathbf{A}^{-1} = \frac{1}{4}\mathbf{A}^H \tag{9}$$

and in the sequel, $\mathbf{q}^r = \frac{1}{4}\mathbf{A}^H\mathbf{q}^a$. The determinant of $\mathbf{A}$ can be readily 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 $\mathbf{A}$ thus becomes

$$\det(\mathbf{A}) = 16^N \tag{10}$$

The basis $\{q, q^\imath, q^\jmath, q^\kappa\}$ in (8) has been chosen to make the matrix $\mathbf{A}$ unitary, thus facilitating its algebraic manipulation.

3 Quaternion Statistics

3.1 Preliminaries

Given that the mapping between quadrivariate real-valued vectors and their quaternion counterparts has now been established, we proceed on presenting the quaternion augmented statistics. In the simplest case, the elementary covariance matrix $\mathcal{C}_{\mathbf{qq}}$ of a quaternion random vector $\mathbf{q} = [q_1 \cdots q_N]^T$ is given by

$$\mathcal{C}_{\mathbf{qq}} = E\{\mathbf{q}\mathbf{q}^H\} \tag{11}$$

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 $\mathbf{q}_d \in \mathbb{R}^{N\times 1}$.

Based on (6) and (8), the real-valued componentwise correlation matrices of the components $\mathbf{q}_a$, $\mathbf{q}_b$, $\mathbf{q}_c$ and $\mathbf{q}_d$ cannot be calculated from the quaternion-valued covariance matrix $\mathcal{C}_{\mathbf{q}}$ alone. Hence, second order information within the quaternion-valued vector $\mathbf{q}$ cannot be described fully by the covariance matrix, and we thus have to introduce complementary correlation matrices: the $\imath$-covariance $\mathcal{C}_{\mathbf{q}\imath}$, the $\jmath$-covariance $\mathcal{C}_{\mathbf{q}\jmath}$, and the κ-covariance $\mathcal{C}_{\mathbf{q}\kappa}$. They complement the information within the covariance, and are given by

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

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 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, meaning that

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

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

These properties are unique to the quaternion domain $\mathbb{H}$, since they do not arise in the statistics of complex valued random variables [8].

3.2 Duality between quaternionic and quadrivariate statistics

In order to recover the real-valued covariance matrices of the real part $\mathbf{z}_a$ and the imaginary part $\mathbf{z}_b$ from the complex domain, both the complex-valued 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 employed as follows

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

Similarly, we can capture the 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 $\mathbf{q}$ are functions of the quaternion-valued covariance and the complementary covariance matrices as

$$\begin{aligned}
\mathcal{C}_{\mathbf{q}_a} &= \frac{1}{4}\Re\{\mathcal{C}_{\mathbf{qq}} + \mathcal{C}_{\mathbf{q}^\imath} + \mathcal{C}_{\mathbf{q}^\jmath} + \mathcal{C}_{\mathbf{q}^\kappa}\} \\
\mathcal{C}_{\mathbf{q}_b} &= \frac{1}{4}\Re\{\mathcal{C}_{\mathbf{qq}} + \mathcal{C}_{\mathbf{q}^\imath} - \mathcal{C}_{\mathbf{q}^\jmath} - \mathcal{C}_{\mathbf{q}^\kappa}\} \\
\mathcal{C}_{\mathbf{q}_c} &= \frac{1}{4}\Re\{\mathcal{C}_{\mathbf{qq}} - \mathcal{C}_{\mathbf{q}^\imath} + \mathcal{C}_{\mathbf{q}^\jmath} - \mathcal{C}_{\mathbf{q}^\kappa}\} \\
\mathcal{C}_{\mathbf{q}_d} &= \frac{1}{4}\Re\{\mathcal{C}_{\mathbf{qq}} - \mathcal{C}_{\mathbf{q}^\imath} - \mathcal{C}_{\mathbf{q}^\jmath} + \mathcal{C}_{\mathbf{q}^\kappa}\} \\
\mathcal{C}_{\mathbf{q}_b\mathbf{q}_a} &= \frac{1}{4}\Im_\imath\{\mathcal{C}_{\mathbf{qq}} + \mathcal{C}_{\mathbf{q}^\imath} + \mathcal{C}_{\mathbf{q}^\jmath} + \mathcal{C}_{\mathbf{q}^\kappa}\} \\
\mathcal{C}_{\mathbf{q}_c\mathbf{q}_a} &= \frac{1}{4}\Im_\jmath\{\mathcal{C}_{\mathbf{qq}} + \mathcal{C}_{\mathbf{q}^\imath} + \mathcal{C}_{\mathbf{q}^\jmath} + \mathcal{C}_{\mathbf{q}^\kappa}\} \\
\mathcal{C}_{\mathbf{q}_d\mathbf{q}_a} &= \frac{1}{4}\Im_\kappa\{\mathcal{C}_{\mathbf{qq}} + \mathcal{C}_{\mathbf{q}^\imath} + \mathcal{C}_{\mathbf{q}^\jmath} + \mathcal{C}_{\mathbf{q}^\kappa}\} \\
\mathcal{C}_{\mathbf{q}_c\mathbf{q}_b} &= \frac{1}{4}\Im_\kappa\{\mathcal{C}_{\mathbf{qq}} + \mathcal{C}_{\mathbf{q}^\imath} - \mathcal{C}_{\mathbf{q}^\jmath} - \mathcal{C}_{\mathbf{q}^\kappa}\} \\
\mathcal{C}_{\mathbf{q}_d\mathbf{q}_b} &= -\frac{1}{4}\Im_\jmath\{\mathcal{C}_{\mathbf{qq}} + \mathcal{C}_{\mathbf{q}^\imath} - \mathcal{C}_{\mathbf{q}^\jmath} - \mathcal{C}_{\mathbf{q}^\kappa}\} \\
\mathcal{C}_{\mathbf{q}_d\mathbf{q}_c} &= \frac{1}{4}\Im_\imath\{\mathcal{C}_{\mathbf{qq}} - \mathcal{C}_{\mathbf{q}^\imath} + \mathcal{C}_{\mathbf{q}^\jmath} - \mathcal{C}_{\mathbf{q}^\kappa}\} \qquad (14)
\end{aligned}$$

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 (8)), is therefore defined as

$$\mathcal{C}^a_{\mathbf{q}} = E\{\mathbf{q}^a\mathbf{q}^{aH}\} = \begin{bmatrix} \mathcal{C}_{\mathbf{qq}} & \mathcal{C}_{\mathbf{q}^\imath} & \mathcal{C}_{\mathbf{q}^\jmath} & \mathcal{C}_{\mathbf{q}^\kappa} \\ \mathcal{C}^H_{\mathbf{q}^\imath} & \mathcal{C}_{\mathbf{q}^\imath\mathbf{q}^\imath} & \mathcal{C}_{\mathbf{q}^\imath\mathbf{q}^\jmath} & \mathcal{C}_{\mathbf{q}^\imath\mathbf{q}^\kappa} \\ \mathcal{C}^H_{\mathbf{q}^\jmath} & \mathcal{C}_{\mathbf{q}^\jmath\mathbf{q}^\imath} & \mathcal{C}_{\mathbf{q}^\jmath\mathbf{q}^\jmath} & \mathcal{C}_{\mathbf{q}^\jmath\mathbf{q}^\kappa} \\ \mathcal{C}^H_{\mathbf{q}^\kappa} & \mathcal{C}_{\mathbf{q}^\kappa\mathbf{q}^\imath} & \mathcal{C}_{\mathbf{q}^\kappa\mathbf{q}^\jmath} & \mathcal{C}_{\mathbf{q}^\kappa\mathbf{q}^\kappa} \end{bmatrix} \qquad (15)$$

where the submatrices in (15) are computed according to

$$\mathcal{C}_{\boldsymbol{\delta}} = E\{\mathbf{q}\boldsymbol{\delta}^H\} \qquad \mathcal{C}_{\boldsymbol{\alpha\beta}} = E\{\boldsymbol{\alpha}\boldsymbol{\beta}^H\}$$
$$\boldsymbol{\delta} \in \{\mathbf{q}^\imath, \mathbf{q}^\jmath, \mathbf{q}^\kappa\} \qquad \boldsymbol{\alpha}, \boldsymbol{\beta} \in \{\mathbf{q}, \mathbf{q}^\imath, \mathbf{q}^\jmath, \mathbf{q}^\kappa\}$$

To assess whether the augmented covariance matrix in (15) provides a complete second order statistical description, we need to demonstrate that it allows a static invertible one-to-one mapping with the corresponding real valued quadrivariate covariance matrix $\mathcal{C}_R$, given by

$$\mathcal{C}_R = E\{\mathbf{q}^r\mathbf{q}^{rT}\} = \begin{bmatrix} \mathcal{C}_{\mathbf{q}_a} & \mathcal{C}_{\mathbf{q}_a\mathbf{q}_b} & \mathcal{C}_{\mathbf{q}_a\mathbf{q}_c} & \mathcal{C}_{\mathbf{q}_a\mathbf{q}_d} \\ \mathcal{C}_{\mathbf{q}_b\mathbf{q}_a} & \mathcal{C}_{\mathbf{q}_b} & \mathcal{C}_{\mathbf{q}_b\mathbf{q}_c} & \mathcal{C}_{\mathbf{q}_b\mathbf{q}_d} \\ \mathcal{C}_{\mathbf{q}_c\mathbf{q}_a} & \mathcal{C}_{\mathbf{q}_c\mathbf{q}_b} & \mathcal{C}_{\mathbf{q}_c} & \mathcal{C}_{\mathbf{q}_c\mathbf{q}_d} \\ \mathcal{C}_{\mathbf{q}_d\mathbf{q}_a} & \mathcal{C}_{\mathbf{q}_d\mathbf{q}_b} & \mathcal{C}_{\mathbf{q}_d\mathbf{q}_c} & \mathcal{C}_{\mathbf{q}_d} \end{bmatrix} \qquad (16)$$

Based on the mapping between the augmented quaternion-valued vector $\mathbf{q}^a$ and the corresponding real valued 'composite' vector $\mathbf{q}^r$ in (8), and that $\mathbf{q}^r = \mathbf{A}^{-1}\mathbf{q}^a = \frac{1}{4}\mathbf{A}^H\mathbf{q}^a$, the real valued covariance matrix can be calculated as

$$\begin{aligned}
\mathcal{C}_R &= \mathbf{A}^{-1}\mathcal{C}^a_{\mathbf{q}}\mathbf{A}^{-H} \\
&= \frac{1}{16}\mathbf{A}^H\mathcal{C}^a_{\mathbf{q}}\mathbf{A} \qquad (17)
\end{aligned}$$

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

3.3 Case Study: 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 [8]. Based on (8) and (15), 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

$$y = \mathbf{w}^{aH}\mathbf{x}^a = \mathbf{g}^H\mathbf{x} + \mathbf{h}^H\mathbf{x}^\imath + \mathbf{u}^H\mathbf{x}^\jmath + \mathbf{v}^H\mathbf{x}^\kappa \qquad (18)$$

Based on this model, it can be shown that the Quaternion LMS (QLMS) algorithm [5] becomes the widely linear Quaternion LMS (WL-QLMS), whose update is given by [3]

$$\mathbf{w}^a(n+1) = \mathbf{w}^a(n) + \mu[2\mathbf{x}^a e^*(n) - e(n)\mathbf{x}^{a*}] \qquad (19)$$

We now proceed on illustrating the advantages of considering the augmented statistics in the next simulation section.

4 Simulations

Three sets of simulations were conducted in a chaotic time-series prediction setting in order to demonstrate the benefits of considering the augmented statistics within quaternion-valued algorithms. As an example, the widely linear Quaternion LMS (19) which operates based on augmented statistics was assessed against the quaternion LMS [5] to predict the one-step ahead chaotic signals and nonstationary wind signals. algorithms. Benchmark dataset used were the three-dimensional (3D) Lorenz attractor and the chaotic four dimensional (4D) Saito's signal [2], and a real-world 3D wind field fused with air temperature to form a four dimensional quaternion signal. The performance index was the prediction gain $Rp = 10\log\frac{\sigma^2_{\mathbf{x}}}{\sigma^2_e}$, where $\sigma^2_{\mathbf{x}}$ and σ^2_e denote respectively the estimated variances of the input and error.

4.1 Simulation 1: Three-dimensional Chaotic Lorenz Attractor

The Lorenz attractor is a three-dimensional nonlinear system used originally to model atmospheric turbulence, but also to model lasers, dynamos, and the motion of waterwheel (For more details, see p. 95 [2]). The Lorenz attractor, shown in the left plot of Fig. 2, can be regarded as as a pure quaternion, and is used as a benchmark signal to test the performance of the QLMS algorithms. Fig. 1 shows the estimates of the Quaternion LMS and its widely linear counterpart, and the prediction gains measured were 30.45 dB and 32.70 dB respectively, confirming the advantage of the augmented statistics.

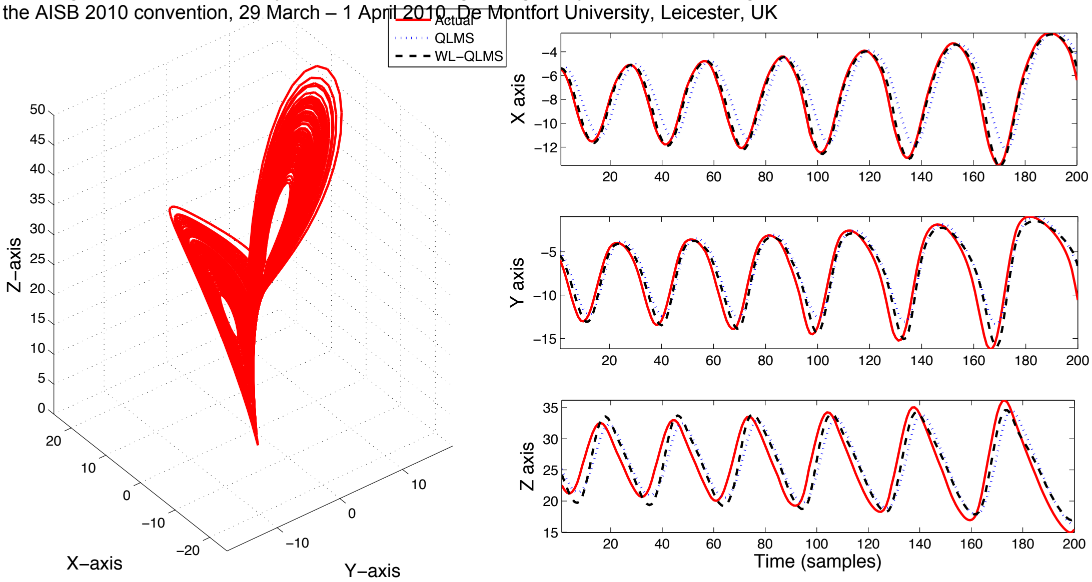

Figure 1. Prediction estimates of QLMS (dotted lines) and WL-QLMS (dashed lines) algorithms on the 3D Lorenz chaotic signal.

4.2 Simulation 2: Four-Dimensional Chaotic Saito's Process

A benchmark chaotic system (shown in left plot of Fig. 2) considered was the 4D Saito's circuit. For more details, see pp. 87-88 [2]. The right plots of Fig. 2 indicates that the widely linear Quaternion LMS (Rp =27.48 dB) outperformed the Quaternion LMS (Rp = 24.74 dB).

4.3 Simulation 3: Wind Forecasting

To illustrate the advantage of the augmented statistics on a nonstationary real-world data, a 3D wind wind field (North-South, East-West and Vertical wind speed directions) was used[3], together with air temperature measurements were considered in a quaternion-valued wind forecasting setting. It is clear from Fig. 3 that the augmented statistics provided enhanced performance, due to the superiority of the widely linear Quaternion LMS over Quaternion LMS. The prediction gains ($Rp_{\text{WL-QLMS}} = 37.36$ dB and $Rp_{\text{QLMS}} = 32.10$ dB) confirmed the observation.

5 Conclusion

Second order statistics 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 covariances, the $\imath$-covariance, the $\jmath$-covariance, and the κ-covariance matrices have been employed. As a case study, we have demonstrated the superiority of the WL-QLMS over QLMS on both synthetic and real world signals, thus supporting the analysis.

[3] The wind data were recorded by Prof. Kazuyuki Aihara and his team at the University of Tokyo, in an urban environment. The wind data was initially sampled at 50 Hz, but re-sampled at 5 Hz for simulation purposes.

REFERENCES

[1] 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*, 23–26, (2004).

[2] P. Arena, L. Fortuna, G. Muscato, and M. G. Xibilia, *Neural Networks in Multidimensional Domains*, Lecture Notes in Control and Information Sciences (Springer Verlag), vol. 234, 1998.

[3] C. Cheong-Took and D. Mandic, 'A Quaternion Widely Linear Adaptive Filter', *Submitted*, (Dec. 2009).

[4] C. Cheong-Took and D. P. Mandic, 'Second-Order Statistics of Quaternion Random Processes', *Submitted*, (Nov. 2009).

[5] C. Cheong-Took and D. P. Mandic, 'The Quaternion LMS Algorithm for Adaptive Filtering of Hypercomplex Real World Processes', *IEEE Transactions on Signal Processing*, **57**(4), 1316–1327, (2009).

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

[7] T. A. Ell and S. J. Sangwine, 'Quaternion Involutions and Anti-Involutions', *Computers & Mathematics with Applications*, **53**(1), 137–143, (2007).

[8] D. P. Mandic and Vanessa Su Lee Goh, *Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models*, John Wiley and Sons Ltd, 2009.

[9] F. Neeser and J. Massey, 'Proper Complex Random Processes With Applications to Information Theory', *IEEE Transactions on Information Theory*, **39**(4), 1293–1302, (1993).

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

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

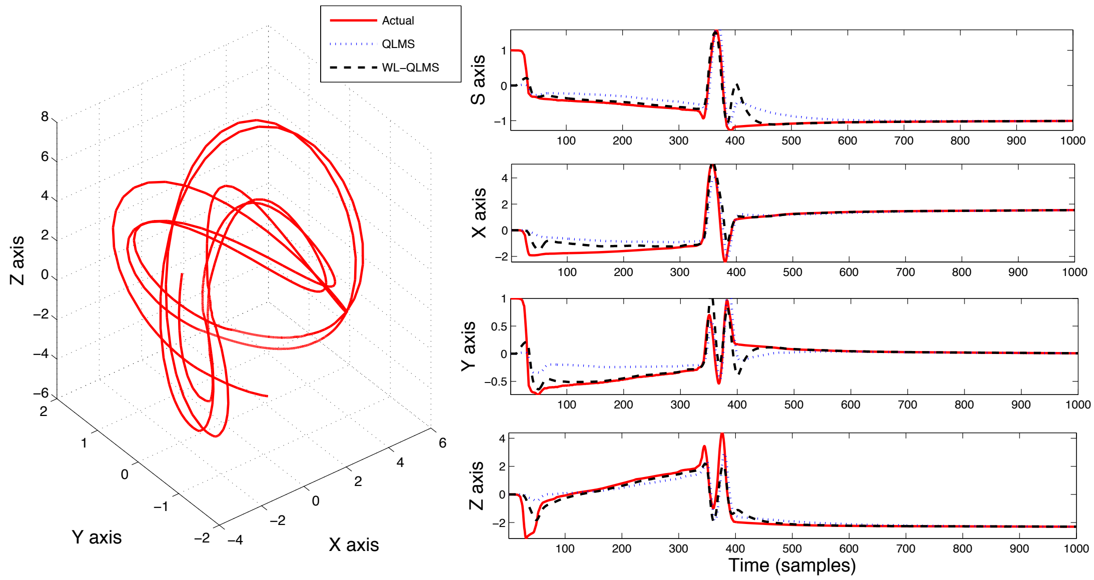

Figure 2. Prediction estimates of QLMS (dotted lines) and WL-QLMS (dashed lines) algorithms on the 4D Saito's signal.

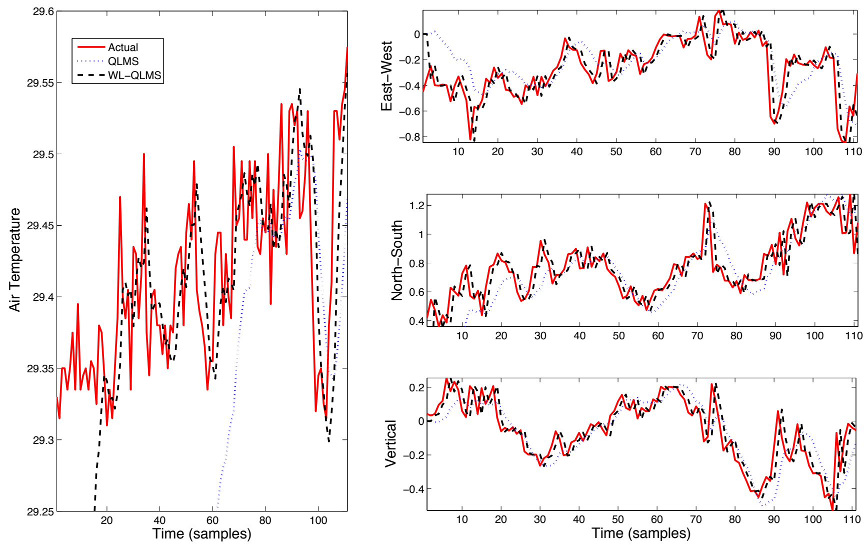

Figure 3. Prediction estimates of QLMS (dotted lines) and WL-QLMS (dashed lines) algorithms on 3D wind field fused with air temperature.

Proceedings of the International Symposium on Evolving Intelligent Systems, Plamen Angelov, Dimitar Filev, and Nikola Kasabov (Eds.), at the AISB 2010 convention, 29 March – 1 April 2010, De Montfort University, Leicester, UK

Application of ANN-GA Hybrid to run a conveyor control system

Paul Morley*
Department of Design & Innovation
The Open University
Milton Keynes, MK7 6AA, England
pm4958@tutor.open.ac.uk

Jeff Johnson
Department of Design & Innovation
The Open University
Milton Keynes, MK7 6AA, England

ABSTRACT
This paper describes the use of a genetic algorithm to find the link weights of an Artificial Neural Network used to control an industrial conveyor system, and further proposes a control strategy to allow continual adaptation of the system to changing circumstances thus maintaining the optimization of the system. Results indicate feasibility of this approach, and that performance of the ANN can be as good as conventional systematic controls

Key words
artificial neural networks, evolutionary computation, machine learning, genetic algorithm, self-adapting controls

1. INTRODUCTION
Artificial neural networks (ANNs), and genetic algorithms (GAs), are both well known and widely used in artificial intelligence fields, including their use in controls applications.

ANNs are useful for pattern recognition, and generalizing patterns in complex multi-dimensional space. GAs are useful for searching widely in large solution spaces.

The problem with ANNs, is finding the weights of the links. Various methods have been tried, the most widely known being back-propagation, which is essentially an error minimization (hill climbing) algorithm. However, there is a key limitation of this approach: it requires a set of training data which is already classified (in order to present an input pattern with a required output).
In an industrial context, conventional control systems are often set-up once achieving a reasonable level of performance (not necessarily optimal), and then left forever. However circumstances change, and in the context described here the mix of work coming through the system will drift causing a gradual fall-off in performance. Therefore it was considered that an learning system which could constantly adapt to the circumstances would be beneficial.

This paper explores the use of a hybrid system for controlling an industrial conveyor system. A single ANN is used as a black-box decision maker about the next operation to be executed, and a GA is used to find the link weights of a population of ANNs.

2. NEURAL NETWORKS
Neural Networks were first developed by McCulloch & Pitts, inspired by biological neural systems. The idea that of computing using a large number of relatively simple units working in parallel as opposed to the conventional paradigm of sequential computing is intuitively attractive.

There is also proof that in theory at least, a 2 layer network is able to implement any measurable function to an arbitrary degree of accuracy. This important result was first established by Kolmogorov (the Kolmogorov existence theorem is actually not very useful in practice, because of the unknown nature of the functions used by each unit) and then in a manner more directly relevant to ANN development by Hornik et al. [1]. Although these results show that a 2 layer (i.e. 1 hidden layer) network could in theory implement any function, in practice it may be the case that a more efficient implementation could be found by using more layers.

There are many forms of ANNs. In this experiment the a fairly simple feed-forward network was used with 3 processing layers. This network comprises layers of individual units fully interconnected to the next layer, each unit implementing a non-linear sigmoid function of a weighted sum of all the inputs

Traditionally the weights associated with the links between each neuron are found by a back-propagation algorithm, which is essentially an error minimization (hill climbing) algorithm. Whilst this is undoubtedly effective, it assumes that there is a Training Set of pre-classified data samples which can be used to train the network (set the weights) and then a separate testing set. This is known as Supervised learning. [2]