A Quaternion Widely Linear Adaptive Filter

Clive Cheong Took, *Member, IEEE,* and Danilo P. Mandic, *Senior Member, IEEE*

***Abstract*—A quaternion widely linear (QWL) model for quaternion valued mean-squared-error (MSE) estimation is proposed. The augmented statistics are first introduced into the field of quaternions, and it is demonstrated that this allows for capturing the complete second order statistics available. The QWL model is next incorporated into the Quaternion Least Mean Square (QLMS) algorithm to yield the widely linear QLMS (WL-QLMS). This allows for a unified approach to adaptive filtering of both $\mathbb{Q}$-proper and $\mathbb{Q}$-improper signals, leading to improved accuracies compared to the QLMS class of algorithms. Simulations on both benchmark and real world data support the analysis.**

***Index Terms*—Quaternion adaptive filtering, $\mathbb{Q}$-properness, quaternion second order noncircularity, Wiener model, quadrivariate processes, quaternion LMS (QLMS), widely linear model, widely linear QLMS.**

I. Introduction

Standard techniques employed in multichannel statistical signal processing typically do not fully cater for the 'coupled' nature of the available information within the channels. Thus, most practical approaches operate based on channelwise processing, which is not optimal for general multivariate signals (where data channels are typically correlated). On the other hand, the quaternion domain $\mathbb{H}$ allows for the direct modelling of three- and four-dimensional signals, and its algebra naturally accounts for the coupling between the signal components.

The use of quaternions is rapidly gaining in popularity, as for instance, many multivariate problems based on vector sensors (motion body sensors, seismics, wind modelling) can be cast into the quaternion domain. The recent resurgence of quaternion valued signal processing stems from the potential advantages that special properties of quaternion algebra offer over real valued vector algebra in multivariate modelling. Applications of quaternions include those in vector sensing [1], machine learning [2], and adaptive filters [3].

Recent advances in complex valued signal processing have been based on the widely linear model proposed by Picinbono [4]. This model, together with the corresponding augmented complex statistics, has been successfully used to design enhanced algorithms in communications [5] [6] and adaptive filters [7]. These studies have shown that widely linear modelling and the associated augmented statistics offer theoretical and practical advantages over the standard complex models, and are applicable to the generality of complex signals, both circular and noncircular.

Models suitable for the processing of signals with rotation dependent distribution (noncircular) are lacking in the quaternion domain, and their development has recently attracted significant research effort [3]. Current second order algorithms operate based on only the quaternion valued covariance [1]–[3] and thus do not fully exploit the available statistical information. Advances in this direction include the work by Vakhania, who defined the concept of $\mathbb{Q}$-properness as the invariance of the distribution of a quaternion valued variable under some specific rotations around the angle of $\pi/2$ [8]. Amblard and Le Bihan relaxed the conditions of $\mathbb{Q}$-properness to an arbitrary axis and angle of rotation φ, that is, $q \triangleq e^{\nu\varphi} q$ [9] for any pure unit quaternion ν (whose real part vanishes); where symbol $\triangleq$ denotes equality in terms of probability density function (pdf).

Although these results provide an initial insight into the processing of general quaternionic signals, they are not straightforward to apply in the context of adaptive filtering applications. To this end, we first propose the quaternion widely linear model, specifically designed for the unified modelling of the generality of quaternion signals, both $\mathbb{Q}$-proper and $\mathbb{Q}$-improper. The benefits of such an approach are shown to be analogous to the benefits that the augmented statistics provides for complex valued data [7]. Next, the QWL model is incorporated into the quaternion LMS [3] to yield the widely linear QLMS (WL-QLMS), and its theoretical and practical advantages are demonstrated through analysis and simulations.

 Manuscript received December 4, 2009; revised April 9, 2010.

C. Cheong Took and D. P. Mandic are with the Dept. of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, U.K. (Email: {c.cheong-took, d.mandic }@imperial.ac.uk). Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

II. Properties of Quaternion Random Vectors

A. Quaternion Algebra

The quaternion domain, a non-commutative extension of the complex domain, provides a natural framework for the processing of three- and four-dimensional signals. A quaternion variable $q \in \mathbb{H}$ comprises a real part $\Re\{\cdot\}$ and a vector-part, also known as a pure quaternion $\Im\{\cdot\}$, consisting of three imaginary components, and can be expressed as:

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

The relationship between the orthogonal unit vectors, $\imath,\ \jmath,\ \kappa$ are given by

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

Given $q_1, q_2 \in \mathbb{H}$, the noncommutative quaternion product is computed as

$$\begin{aligned} q_1 q_2 &= \Re\{q_1 q_2\} + \Im\{q_1 q_2\} \\ \text{where} \quad \Re\{q_1 q_2\} &= q_{1,a}q_{2,a} + q_{1,b}q_{2,b} + q_{1,c}q_{2,c} + q_{1,d}q_{2,d} \\ \Im\{q_1 q_2\} &= q_{1,a}\Im\{q_2\} + q_{2,a}\Im\{q_1\} + \Im\{q_1\} \times \Im\{q_2\} \end{aligned} \tag{3}$$

where the symbol "×" denotes the vector product.

B. Quaternion Involutions

Given a complex number $z = z_a + \imath z_b$, its real and imaginary part can be extracted as $z_a = \frac{1}{2}(z + z^*)$ and $z_b = \frac{1}{2\imath}(z - z^*)$. However, such convenient manipulation is not possible in the quaternion domain. To circumvent this problem, the three perpendicular quaternion involutions (self-inverse mappings) given by

$$\begin{aligned} q^\imath &= -\imath q \imath = q_a + \imath q_b - \jmath q_c - \kappa q_d \\ q^\jmath &= -\jmath q \jmath = q_a - \imath q_b + \jmath q_c - \kappa q_d \\ q^\kappa &= -\kappa q \kappa = q_a - \imath q_b - \jmath q_c + \kappa q_d \end{aligned} \tag{4}$$

can be employed, and the four components of the quaternion variable q can now be computed as [10]

$$q_a = \frac{1}{2}(q + q^*) \qquad q_b = \frac{1}{2\imath}(q - q^{\imath *})$$
$$q_c = \frac{1}{2\jmath}(q - q^{\jmath *}) \qquad q_d = \frac{1}{2\kappa}(q - q^{\kappa *}) \tag{5}$$

The quaternion conjugate operation $(\cdot)^*$ is also an involution and can be expressed as a linear function of the three perpendicular involutions, that is

$$q^* = \frac{1}{2}(q^\imath + q^\jmath + q^\kappa - q) \tag{6}$$

III. Quaternion Statistics

Picinbono demonstrated that the complete description of the second order statistics in $\mathbb{C}$ can be achieved, provided the real valued bivariate covariance matrices can be calculated from their complex valued counterparts (see pp. 118-119 [11]). Following on this result, we next show that the complete second order statistical description in $\mathbb{H}$ is obtained when the real valued quadrivariate covariances are expressed in terms of their quaternion valued counterparts, as shown in Appendix IX-A. However, unlike the complex domain $\mathbb{C}$, where for this purpose it is sufficient to combine the complex vector $\mathbf{z}$ and its conjugate $\mathbf{z}^*$ into the augmented complex vector $\mathbf{z}^a = [\mathbf{z}^T \mathbf{z}^H]^T$, in the quaternion domain, we also need to consider the involutions (4). We can therefore build an augmented quaternion vector, comprising of any four of the five quantities $\{q, q^*, q^\imath, q^\jmath, q^\kappa\}$ or their conjugates. One convenient augmented basis is $q^a = \{q, q^*, q^{\imath *}, q^{\jmath *}\}$, and will be used in this work. Then, the augmented vector $\mathbf{q}^a = [\mathbf{q}^T \ \mathbf{q}^H \ \mathbf{q}^{\imath H} \ \mathbf{q}^{\jmath H}]^T$ contains all the necessary second order statistical information and its augmented covariance matrix is given by

$$\mathcal{C}^a_{\mathbf{q}} = E\{\mathbf{q}^a \mathbf{q}^{aH}\} = \begin{bmatrix} \mathcal{C}_{\mathbf{q}} & \mathcal{P}_{\mathbf{q}} & \mathcal{P}^\imath_{\mathbf{q}} & \mathcal{P}^\jmath_{\mathbf{q}} \\ \mathcal{P}^H_{\mathbf{q}} & \tilde{\mathcal{C}}_{\mathbf{q}} & \tilde{\mathcal{C}}_{\mathbf{q}\mathbf{q}^\imath} & \tilde{\mathcal{C}}_{\mathbf{q}\mathbf{q}^\jmath} \\ \mathcal{P}^{\imath H}_{\mathbf{q}} & \tilde{\mathcal{C}}^H_{\mathbf{q}\mathbf{q}^\imath} & \tilde{\mathcal{C}}_{\mathbf{q}^\imath} & \tilde{\mathcal{C}}_{\mathbf{q}^\imath\mathbf{q}^\jmath} \\ \mathcal{P}^{\jmath H}_{\mathbf{q}} & \tilde{\mathcal{C}}^H_{\mathbf{q}\mathbf{q}^\jmath} & \tilde{\mathcal{C}}^H_{\mathbf{q}^\imath\mathbf{q}^\jmath} & \tilde{\mathcal{C}}_{\mathbf{q}^\jmath} \end{bmatrix} \tag{7}$$

The submatrices in (7) are calculated according to

$$\mathcal{C}_{\boldsymbol{\alpha}} = E\{\boldsymbol{\alpha}\boldsymbol{\alpha}^H\} \qquad \tilde{\mathcal{C}}_{\boldsymbol{\alpha}} = E\{\boldsymbol{\alpha}^*\boldsymbol{\alpha}^T\}$$
$$\tilde{\mathcal{C}}_{\boldsymbol{\alpha}\boldsymbol{\beta}} = E\{\boldsymbol{\alpha}^*\boldsymbol{\beta}^T\} \qquad \mathcal{P}_{\boldsymbol{\alpha}} = E\{\boldsymbol{\alpha}\boldsymbol{\alpha}^T\} \quad \boldsymbol{\alpha}, \boldsymbol{\beta} \in \{\mathbf{q}, \mathbf{q}^\imath, \mathbf{q}^\jmath\}$$

We refer to $\tilde{\mathcal{C}}_{\boldsymbol{\alpha}}$ as the quasicovariance and $\tilde{\mathcal{C}}_{\boldsymbol{\alpha}\boldsymbol{\beta}}$ as the cross-quasicovariance matrices. Unlike in the complex case [4], the non-commutativity of the quaternion product results in $\mathcal{C}_{\boldsymbol{\alpha}} \neq \tilde{\mathcal{C}}^*_{\boldsymbol{\alpha}}$.

A. Circularity in $\mathbb{H}$ and $\mathbb{Q}$-properness

For a quaternion valued variable to be second order circular (or $\mathbb{Q}$-proper), its probability distribution should be rotation-invariant with respect to 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, a $\mathbb{Q}$-proper quaternion random variable should satisfy the following properties [8]

$$\begin{aligned} &\text{P1:} && E\{q_\delta^2\} = E\{q_\epsilon^2\} = \sigma^2 \quad \forall\ \delta, \epsilon = a, b, c, d \\ &\text{P2:} && E\{q_\delta q_\epsilon\} = 0 \quad \forall\ \delta, \epsilon = a, b, c, d \text{ and } \delta \neq \epsilon \\ &\text{P3:} && E\{qq\} = -2E\{q_\delta^2\} = -2\sigma^2 \quad \forall\ \delta = a, b, c, d \\ &\text{P4:} && E\{|q|^2\} = 4E\{q_\delta^2\} = 4\sigma^2 \quad \forall\ \delta = a, b, c, d \end{aligned} \tag{8}$$

The first property, P1, states that all the four-signal components of a quaternion valued variable have equal variances. The property P2 implies that the components of q are not correlated. Property P3 suggests that the pseudocovariance matrix of $\mathbb{Q}$-proper signals does not vanish, in contrast to the complex case. Finally, the fourth property states that the power of a quaternion variable is a sum of the powers of the signal components. Observe that properties P1 and P2 imply properties P3 and P4.

B. Augmented Statistics of $\mathbb{Q}$-proper variables

Notice that $\mathbb{Q}$-properness also implies that the quaternion vector $\mathbf{q}$ is uncorrelated with the vector involutions $\mathbf{q}^\imath$, $\mathbf{q}^\jmath$, $\mathbf{q}^\kappa$, that is[1],

$$E\{\mathbf{q}\mathbf{q}^{\imath H}\} = \mathbf{0} \qquad E\{\mathbf{q}\mathbf{q}^{\jmath H}\} = \mathbf{0} \qquad E\{\mathbf{q}\mathbf{q}^{\kappa H}\} = \mathbf{0} \tag{9}$$

This simplifies the structure of the augmented covariance matrix $\mathcal{C}^a_{\mathbf{q}}$ of a $\mathbb{Q}$-proper random vector, as

$$\mathcal{C}^a_{\mathbf{q}} = \begin{bmatrix} \mathcal{C}_{\mathbf{q}} & \mathcal{P}_{\mathbf{q}} & \mathcal{P}^\imath_{\mathbf{q}} & \mathcal{P}^\jmath_{\mathbf{q}} \\ \mathcal{P}_{\mathbf{q}} & \mathcal{C}_{\mathbf{q}} & \mathbf{0} & \mathbf{0} \\ \mathcal{P}^\imath_{\mathbf{q}} & \mathbf{0} & \mathcal{C}_{\mathbf{q}} & \mathbf{0} \\ \mathcal{P}^\jmath_{\mathbf{q}} & \mathbf{0} & \mathbf{0} & \mathcal{C}_{\mathbf{q}} \end{bmatrix} = 2\sigma^2 \begin{bmatrix} 2\mathbf{I} & -\mathbf{I} & \mathbf{I} & \mathbf{I} \\ -\mathbf{I} & 2\mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{I} & \mathbf{0} & 2\mathbf{I} & \mathbf{0} \\ \mathbf{I} & \mathbf{0} & \mathbf{0} & 2\mathbf{I} \end{bmatrix} \tag{10}$$

that is, the cross-quasicovariance matrices $\tilde{\mathcal{C}}_{\boldsymbol{\alpha\beta}}$ all vanish.

IV. The Quaternion Widely Linear Model

To account for the complete second order statistics of quaternion valued signals in mean-squared error (MSE) estimation, we need to introduce a filtering model corresponding to the widely linear model in the complex case [4]. Consider the MSE estimator of a signal y in terms of another observation x, that is, $\hat{y} = E[y|x]$. For zero mean, jointly normal real y and x, the solution is

$$\hat{y} = \mathbf{h}^T \mathbf{x} \tag{11}$$

In the quaternion domain, however, the real estimator (11) applies to each component (the real and the three imaginary parts) of quaternion variables, that is

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

and thus

$$\begin{aligned} \hat{y} \ =\ & E[y_a|x_a, x_b, x_c, x_d] + \imath E[y_b|x_a, x_b, x_c, x_d] \\ & + \jmath E[y_c|x_a, x_b, x_c, x_d] + \kappa E[y_d|x_a, x_b, x_c, x_d] \end{aligned} \tag{12}$$

Upon employing the identities (5), it is clear that the quaternion estimator can also be expressed[2] as

$$\begin{aligned} \hat{y} \ =\ & E[y|x, x^*, x^{\imath *}, x^{\jmath *}] + \imath E[y^\imath|x, x^*, x^{\imath *}, x^{\jmath *}] \\ & + \jmath E[y^\jmath|x, x^*, x^{\imath *}, x^{\jmath *}] + \kappa E[y^\kappa|x, x^*, x^{\imath *}, x^{\jmath *}] \end{aligned} \tag{13}$$

and thus arrive at the widely linear estimator for general quaternion signals

$$y \ =\ \mathbf{h}^H \mathbf{x} + \mathbf{g}^H \mathbf{x}^* + \mathbf{u}^H \mathbf{x}^{\imath *} + \mathbf{v}^H \mathbf{x}^{\jmath *} = \mathbf{w}^{aH} \mathbf{x}^a \tag{14}$$

where $\mathbf{w}^a = [\mathbf{h}^T \ \mathbf{g}^T \ \mathbf{u}^T \ \mathbf{v}^T]^T$ and $\mathbf{x}^a = [\mathbf{x} \ \mathbf{x}^H \ \mathbf{x}^{\imath H} \ \mathbf{x}^{\jmath H}]^T$. Following on the proposed quaternion widely linear model, the Wiener solution is now derived as the optimal Mean Square Estimator (MSE). Consider the standard real valued quadratic cost function, that is,

$$\begin{aligned} \mathcal{J} \ &=\ E\{ee^*\} = E\{[d - y][d^* - y^*]\} \\ &=\ E\{dd^*\} + E\{yy^*\} - E\{yd^*\} - E\{dy^*\} \end{aligned} \tag{15}$$

[1]Similarly, for a complex valued random vector $\mathbf{z}$, $\mathbb{C}$-properness means that $\mathbf{z}$ is uncorrelated with $\mathbf{z}^*$ in the 'complex sense', since $E\{\mathbf{z}(\mathbf{z}^*)^H\} = E\{\mathbf{z}\mathbf{z}^T\} = \mathbf{0}$, for more detail see [7].

[2]Any other augmented basis other than $\{x, x^*, x^{\imath *}, x^{\jmath *}\}$ can be used, as explained in the Section III.

The derivative of the cost function (15) can be expressed as (for full derivation, see Appendix IX-B)

$$\begin{aligned}\nabla_{\mathbf{w}^a}\mathcal{J} &= E\Big\{(\nabla_{\mathbf{w}^a}y)y^* + y(\nabla_{\mathbf{w}^a}y^*) - (\nabla_{\mathbf{w}^a}y)d^* - d(\nabla_{\mathbf{w}^a}y^*)\Big\} \\ &= \underbrace{E\Big\{4[\mathbf{x}^a y^* - \mathbf{x}^a d^*]\Big\}}_{\mathrm{I}} + \underbrace{E\Big\{2[d(n)\mathbf{x}^a - y\mathbf{x}^a]\Big\}}_{\mathrm{II}} \qquad (16)\end{aligned}$$

To obtain the Wiener solution, the expectations of I and II in (16) are set to zero. In the complex domain, we can sum up the terms I and II in (16); however, due to the noncommutativity of the quaternion product, we need to consider the terms in (16) individually, giving the solution[3]

$$\begin{aligned}\mathrm{I}: \ \mathbf{w}_o &= E\{\mathbf{x}^a\mathbf{x}^{aH}\}^{-1}E\{\mathbf{x}^a d^*\} \qquad (17)\\ \mathrm{II}: \ \mathbf{w}_o &= E\{\mathbf{x}^{a*}\mathbf{x}^{aH}\}^{-1}E\{\mathbf{x}^{a*} d^*\} \qquad (18)\end{aligned}$$

The first condition for the Wiener solution (17) requires the inversion of the augmented covariance $\mathcal{C}^a_{\mathbf{x}} = E\{\mathbf{x}^a\mathbf{x}^{aH}\}$. On the other hand, the second condition (18) also depends on the conjugate of pseudo-covariance matrix of the augmented vector $\mathbf{x}^a$, which conforms with the observation in [3] that the quaternion domain accounts inherently for the information contained in pseudocovariance.

V. The Widely Linear Quaternion Least Mean Square Algorithm

We now extend some recent results in quaternion adaptive filtering [3], and employ the quaternion widely linear model (14) within the stochastic gradient adaptive filtering framework in $\mathbb{H}$, to propose the Widely Linear Quaternion Least Mean Square (WL-QLMS) adaptive filtering algorithm. Within the stochastic gradient descent optimisation, the gradient of the instantaneous cost function (15) is

$$\begin{aligned}\nabla_{\mathbf{w}^a}\mathcal{J}(n) &= e(n)\big(\nabla_{\mathbf{w}^a}e^*(n)\big) + \big(\nabla_{\mathbf{w}^a}e(n)\big)e^*(n) \\ &= 2e(n)\mathbf{x}^a(n) - 4\mathbf{x}^a(n)e^*(n) \qquad (19)\end{aligned}$$

Notice that due to the non-commutativity of the quaternion product, the two error gradient terms in (19) need to be treated separately. Based on the generic stochastic gradient update $\Delta\mathbf{w}^a = -\mu\nabla_{\mathbf{w}^a}\mathcal{J}(n)$, the update of the weight vector of the WL-QLMS algorithm can be obtained as[4]

$$\mathbf{w}^a(n+1) = \mathbf{w}^a(n) + \mu\big(2\mathbf{x}^a(n)e^*(n) - e(n)\mathbf{x}^a(n)\big) \qquad (20)$$

where the scaling factor of two has been absorbed in the stepsize μ. Given $y(n) = \mathbf{w}^H(n)\mathbf{x}(n)$, observe the same form of the update in (20) as that within the QLMS in [3]. Although AQLMS outperformed QLMS, as shown in [3] in the context of $\mathbb{Q}$-improper signals, its second order information is derived from only the covariance and the pseudocovariance, and is therefore still suboptimal for $\mathbb{Q}$-improper data, as it cannot model the information contained in matrices such as $\mathcal{C}_{\mathbf{q}_c}$ (see Appendix IX-A). Also, observe that the real-valued multichannel LMS does not exploit the interchannel cross-correlation in the same elegant and intuitive way as the WL-QLMS. For more details, see p. 132 of [12] and the performance comparisons with AQLMS in [3]. Finally, a convergence analysis of WL-QLMS algorithm is provided in Appendix IX-C, and the upper bound on the stepsize is found to be $0 < \mu < 2/\lambda_{\max}$, where $\lambda_{\max}$ denotes the maximum right eigenvalue of $\mathcal{C}^\alpha_{\mathbf{x}} = 2\mathcal{C}^a_{\mathbf{x}} + \mathcal{P}^{a*}_{\mathbf{x}}$.

[3]Similarly to complex-valued case [7], there are several equivalent formulations for the quaternion-valued Wiener solution. For instance, if the filter output is considered as $y(n) = \mathbf{w}^{aT}(n)\mathbf{x}^a(n)$ instead of $y(n) = \mathbf{w}^{aH}\mathbf{x}^a$, an alternative solution is obtained.

[4]For more detail of the derivation, see [3] and the Appendix IX-B.

VI. Simulations

Three sets of simulations were conducted in an M-step ahead prediction setting in order to comprehensively assess the performance of the proposed WL-QLMS algorithm against QLMS and AQLMS. The datasets used were a $\mathbb{Q}$-proper synthetic AR(4) process [7], the $\mathbb{Q}$-improper four dimensional (4D) Saito's signal [2] and real-world 4D wind field signal [3]. The quantitative performance index was the prediction gain $R_p = 10\log\frac{\sigma^2_{\mathbf{x}}}{\sigma^2_e}$, where $\sigma^2_{\mathbf{x}}$ and σ^2_e denote respectively the estimated variances of the input and error; the filter length is denoted by L.

1) $\mathbb{Q}$-proper Autoregressive(4) Model: The autoregressive AR(4) process $x(n) = 1.79x(n-1) - 1.85x(n-2) + 1.27x(n-3) - 0.41x(n-4)$ was driven by quadruply white Gaussian noise, whose real and imaginary components were uncorrelated, but had equal variances. From Fig. 1, initially, QLMS converged faster than WL-QLMS and AQLMS. This is because 1) it operates based on the covariance $\mathcal{C}_{\mathbf{x}}$ only, which is adequate to describe the complete statistics of $\mathbb{Q}$-proper signals, due to the deterministic relationship $\mathcal{C}_{\mathbf{x}} = -\mathcal{P}_{\mathbf{x}}/2 = \mathcal{P}^{\imath}_{\mathbf{x}}/2 = \mathcal{P}^{\jmath}_{\mathbf{x}}/2$; 2) QLMS has fewer filter parameters to adapt. However, at steady state the prediction gain for all the algorithms considered was approximately 30 dB, as they are all suited to process $\mathbb{Q}$-proper data.

2) $\mathbb{Q}$-improper Four-Dimensional Saito's Circuit [2]: Fig. 2 compares the performance of the quaternion algorithms [3] over a range of filter parameters. Conforming with the analysis in Section IV, the WL-QLMS algorithm outperformed the QLMS and AQLMS on the $\mathbb{Q}$-improper Saito process, also exhibiting better convergence properties (see top plot of Fig. 4), as QLMS accounts for the second order information only from the covariance matrix $\mathcal{C}_{\mathbf{x}}$, and the AQLMS operates based on $\mathcal{C}_{\mathbf{x}}$ and $\mathcal{P}_{\mathbf{x}}$; both are not adequate to account for the complete statistics of $\mathbb{Q}$-improper signals.

3) $\mathbb{Q}$-improper Wind Forecasting: The 4D quaternion dataset comprised the 3D wind speed (North-South, East-West, and vertical directions[5]) as a vector part (pure quaternion) and air temperature as the scalar part. Fig. 3 shows that the prediction results over a range of filter parameters are in agreement with theoretical analysis, illustrating that WL-QLMS outperformed QLMS and AQLMS on $\mathbb{Q}$-improper nonstationary data. Fig. 4 (bottom plot) illustrates the improved convergence properties of WL-QLMS over QLMS, due to WL-QLMS using the complete second order statistical information available.

VII. Conclusion

We have introduced a quaternion widely linear model (QWL) for enhanced second order estimation of general quaternionic signals. We have demonstrated the efficacy of the QWL model, by incorporating it into the Quaternion Least Mean Square (QLMS) algorithm [3] to yield the widely linear QLMS (WL-QLMS) algorithm. For rigour, the convergence analysis includes the stepsize bound and learning curves for both second order circular and noncircular signals. Experiments have been conducted for a range of filter parameters and dataset, illustrating the WL-QLMS outperforming other algorithms of the kind.

VIII. Acknowledgement

We would like to thank Gill instruments for kindly providing us with the 3D WindMaster anemometer.

[5]The wind data were sampled at 32 Hz and recorded by the 3D WindMaster anemometer by Gill Instruments.

References

[1] 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.

[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. 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.

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

[5] H. Gerstacker, R. Schober, and A. Lampe, "Receivers with widely linear processing for frequency-selective channels," *IEEE Transactions on Communication*, vol. 51, no. 9, pp. 1512–1523, 2003.

[6] R. Schober, W. H. Gerstacker, and L. H.-J. Lampe, " Data-aided and blind stochastic gradient algorithms for widely linear MMSE MAI suppression for DS-CDMA," *IEEE Transactions on Signal Processing*, vol. 52, no. 3, pp. 746– 756, 2004.

[7] 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.

[8] 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.

[9] 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.

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

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

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

IX. Appendix

A. Complete Second Order Statistics

The real valued correlation matrices of single components $\mathbf{q}_a$, $\mathbf{q}_b$, $\mathbf{q}_c$ and $\mathbf{q}_d$ of the quaternion random vector $\mathbf{q}$ can be expressed in terms of the quaternion valued covariance matrices as

$$\begin{aligned}
\mathcal{C}_{\mathbf{q}_a} &= \frac{1}{2}\Re\{\mathcal{C}_{\mathbf{q}} + \mathcal{P}_{\mathbf{q}}\} & \mathcal{C}_{\mathbf{q}_b} &= \frac{1}{2}\Re\{\mathcal{C}_{\mathbf{q}} - \mathcal{P}^{\imath}_{\mathbf{q}}\}\\
\mathcal{C}_{\mathbf{q}_c} &= \frac{1}{2}\Re\{\mathcal{C}_{\mathbf{q}} - \mathcal{P}^{\jmath}_{\mathbf{q}}\} & \mathcal{C}_{\mathbf{q}_d} &= \Re\{\mathcal{C}_{\mathbf{q}}\} - \Big(\mathcal{C}_{\mathbf{q}_a} + \mathcal{C}_{\mathbf{q}_b} + \mathcal{C}_{\mathbf{q}_c}\Big)\\
\mathcal{C}_{\mathbf{q}_b\mathbf{q}_a} &= \frac{1}{2}\Im_i\{\mathcal{C}_{\mathbf{q}} + \mathcal{P}_{\mathbf{q}}\} & \mathcal{C}_{\mathbf{q}_c\mathbf{q}_a} &= \frac{1}{2}\Im_j\{\mathcal{C}_{\mathbf{q}} + \mathcal{P}_{\mathbf{q}}\}\\
\mathcal{C}_{\mathbf{q}_d\mathbf{q}_a} &= \frac{1}{2}\Im_k\{\mathcal{C}_{\mathbf{q}} + \mathcal{P}_{\mathbf{q}}\} & \mathcal{C}_{\mathbf{q}_c\mathbf{q}_b} &= \frac{1}{2}\Im_k\{\mathcal{C}_{\mathbf{q}} - \mathcal{P}^{\imath}_{\mathbf{q}}\}\\
\mathcal{C}_{\mathbf{q}_d\mathbf{q}_c} &= \frac{1}{2}\Im_i\{\mathcal{C}_{\mathbf{q}} - \mathcal{P}^{\jmath}_{\mathbf{q}}\} & \mathcal{C}_{\mathbf{q}_d\mathbf{q}_b} &= -\frac{1}{2}\Im_j\{\mathcal{C}_{\mathbf{q}} - \mathcal{P}^{\kappa}_{\mathbf{q}}\}
\end{aligned} \tag{21}$$

B. Derivation of the Stochastic Gradients

To calculate the derivatives of the output $y(n)$ and its conjugate with respect to the augmented weight vector, the terms $\mathbf{w}^{aH}(n)\mathbf{x}(n)$ and $\mathbf{x}^{aH}(n)\mathbf{w}$ need to be examined. Consider the first term in (14) $\mathbf{h}^H\mathbf{x}$ and its conjugate $\mathbf{x}^H\mathbf{h}$. The partial derivatives of the output $y(n) = \mathbf{w}^{aH}(n)\mathbf{x}(n)$ with respect to $\mathbf{h}$ are

$$\begin{aligned}
\frac{\partial y}{\partial \mathbf{h}_a} &= \mathbf{x}_a + \imath\mathbf{x}_b + \jmath\mathbf{x}_c + \kappa\mathbf{x}_d & \frac{\partial y^*}{\partial \mathbf{h}_a} &= \mathbf{x}_a - \imath\mathbf{x}_b - \jmath\mathbf{x}_c - \kappa\mathbf{x}_d\\
\imath\frac{\partial y}{\partial \mathbf{h}_b} &= \mathbf{x}_a + \imath\mathbf{x}_b + \jmath\mathbf{x}_c + \kappa\mathbf{x}_d & \imath\frac{\partial y^*}{\partial \mathbf{h}_b} &= -\mathbf{x}_a + \imath\mathbf{x}_b - \jmath\mathbf{x}_c - \kappa\mathbf{x}_d\\
\jmath\frac{\partial y}{\partial \mathbf{h}_c} &= \mathbf{x}_a + \imath\mathbf{x}_b + \jmath\mathbf{x}_c + \kappa\mathbf{x}_d & \jmath\frac{\partial y^*}{\partial \mathbf{h}_c} &= -\mathbf{x}_a - \imath\mathbf{x}_b + \jmath\mathbf{x}_c - \kappa\mathbf{x}_d\\
\kappa\frac{\partial y}{\partial \mathbf{h}_d} &= \mathbf{x}_a + \imath\mathbf{x}_b + \jmath\mathbf{x}_c + \kappa\mathbf{x}_d & \kappa\frac{\partial y^*}{\partial \mathbf{h}_d} &= -\mathbf{x}_a - \imath\mathbf{x}_b - \jmath\mathbf{x}_c + \kappa\mathbf{x}_d
\end{aligned} \tag{22}$$

Therefore, $\frac{\partial y}{\partial \mathbf{h}} = \frac{\partial y}{\partial \mathbf{h}_a} + \imath\frac{\partial y}{\partial \mathbf{h}_b} + \jmath\frac{\partial y}{\partial \mathbf{h}_c} + \kappa\frac{\partial y}{\partial \mathbf{h}_d} = 4\mathbf{x}$ and $\frac{\partial y^*}{\partial \mathbf{h}} = -2\mathbf{x}$. Similarly, the derivatives of the terms $\mathbf{g}^H\mathbf{x}^*$, $\mathbf{u}^H\mathbf{x}^{\imath*}$, $\mathbf{v}^H\mathbf{x}^{\jmath*}$ and their conjugates in (14) can also be computed respectively as $\frac{\partial y}{\partial \mathbf{g}} = 4\mathbf{x}^*$, $\frac{\partial y}{\partial \mathbf{u}} = 4\mathbf{x}^{\imath*}$, $\frac{\partial y}{\partial \mathbf{v}} = 4\mathbf{x}^{\jmath*}$; $\frac{\partial y^*}{\partial \mathbf{g}} = -2\mathbf{x}^*$, $\frac{\partial y^*}{\partial \mathbf{u}} = -2\mathbf{x}^{\imath*}$, and $\frac{\partial y^*}{\partial \mathbf{v}} = -2\mathbf{x}^{\jmath*}$. As a result, the augmented quaternion gradients can be expressed as $\frac{\partial y}{\partial \mathbf{w}^a} = 4\mathbf{x}^a$ and $\frac{\partial y^*}{\partial \mathbf{w}^a} = -2\mathbf{x}^a$.

C. Convergence Analysis

To factor out the conjugate error $e^*(n)$ on the right hand side, we start by replacing $e(n)\mathbf{x}^a(n) = \mathbf{x}^{a*}(n)e^*(n) - 2\Im\{\mathbf{x}^{a*}(n)e^*(n)\}$ into (20) to yield

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu\Big(\big[2\mathbf{x}^a(n) + \mathbf{x}^{a*}(n)\big]e^*(n) - 2\Im\{\mathbf{x}^{a*}(n)e^*(n)\}\Big) \tag{23}$$

Since $e(n) = -[\mathbf{w}(n) - \mathbf{w}_o]^H\mathbf{x}^a(n) = -\mathbf{v}^H(n)\mathbf{x}^a(n)$ [with $\mathbf{w}_o$ as the Wiener solution], (23) becomes

$$\begin{aligned}
\mathbf{v}(n+1) &= \mathbf{v}(n) - \mu\Big[\big(2E\{\mathbf{x}^a(n)\mathbf{x}^{aH}(n)\}\\
&\quad - E\{\mathbf{x}^{a*}(n)\mathbf{x}^{aH}(n)\}\big)\mathbf{v}(n)\\
&\quad + 2\Im\Big\{E\{\mathbf{x}^{a*}(n)\mathbf{x}^{aH}(n)\}\mathbf{v}(n)\Big\}\Big]\\
\mathbf{v}(n+1) &= \mathbf{v}(n) - \mu\Big[\big(2\mathcal{C}^a_{\mathbf{x}} - \mathcal{P}^{a*}_{\mathbf{u}}\big)\mathbf{v}(n) + 2\Im\{\mathcal{P}^{a*}_{\mathbf{x}}\mathbf{v}(n)\}\Big]
\end{aligned} \tag{24}$$

The upper bound on the stepsize can then be approximated, by considering the imaginary part $\Im\{\mathcal{P}^{a*}_{\mathbf{x}}\mathbf{v}(n)\}$ as a full quaternion, to give $\mathbf{v}(n+1) \approx \mathbf{v}(n) - \mu[2\mathcal{C}^a_{\mathbf{x}} + \mathcal{P}^{a*}_{\mathbf{x}}]\mathbf{v}(n)$. By letting $\mathcal{C}^{\alpha}_{\mathbf{x}} = 2\mathcal{C}^a_{\mathbf{x}} + \mathcal{P}^{a*}_{\mathbf{x}}$ and taking the right eigenvalue decomposition of $\mathcal{C}^{\alpha}_{\mathbf{x}}$, with its maximum eigenvalue $\lambda_{\max}$, we can obtain the stepsize bound for WL-QLMS as $0 < \mu < \frac{2}{\lambda_{\max}}$, a generic form which also applies to QLMS and AQLMS. For instance, in the case of AQLMS, the augmented covariance matrix $\mathcal{C}^a_{\mathbf{x}}$ degenerates into its 2×2 top-left submatrix in (7), whereas $\mathcal{C}^a_{\mathbf{x}}$ becomes $\mathcal{C}_{\mathbf{x}}$ for QLMS. This implies a larger stepsize is required for these algorithms to converge at the same rate as WL-QLMS in the case of $\mathbb{Q}$-improper signals (when every element of $\mathcal{C}^a_{\mathbf{x}}$ does not vanish).

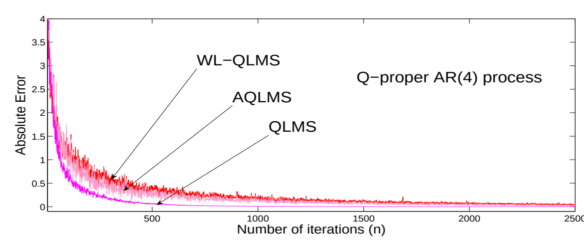

Fig. 1. Learning curves of WL-QLMS, QLMS and AQLMS on the prediction of the $\mathbb{Q}$-proper AR(4) process $x(n) = 1.79x(n-1) - 1.85x(n-2) + 1.27x(n-3) - 0.41x(n-4)$.

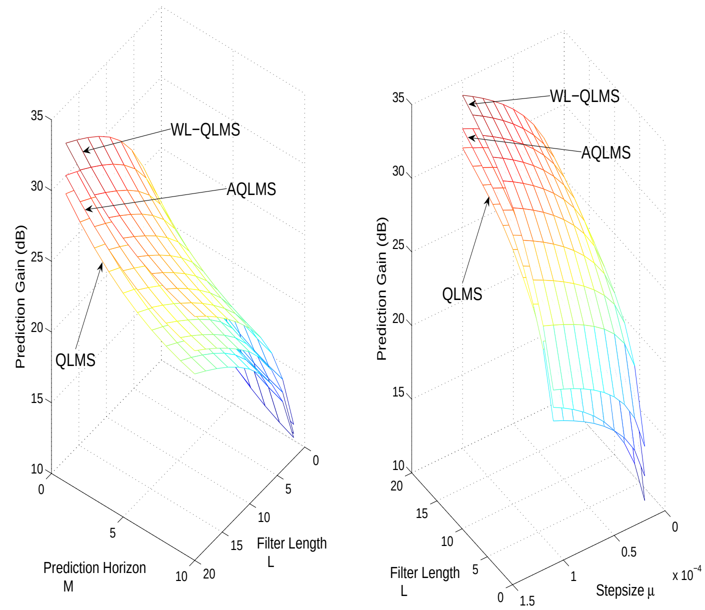

Fig. 2. The performance of WL-QLMS , QLMS and AQLMS on the prediction of 4D Saito's process.

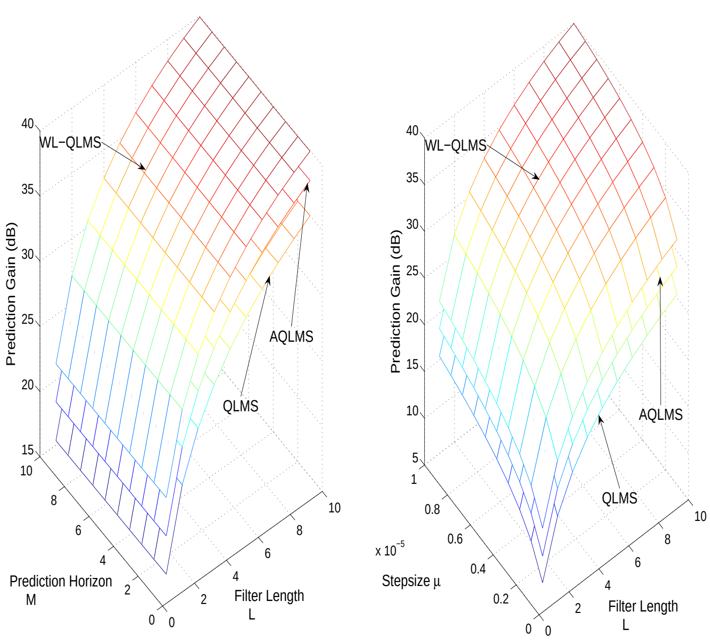

Fig. 3. The performance of WL-QLMS, QLMS and AQLMS on the prediction of a 4D wind field.

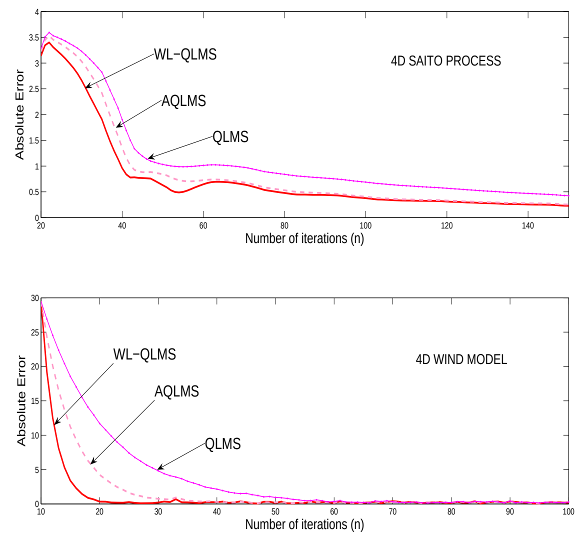

Fig. 4. Learning curves of WL-QLMS, QLMS and AQLMS on the prediction of $\mathbb{Q}$-improper processes, i.e. 4D Saito's process (top) and 4D wind field model (bottom).