

A Unifying Framework for the Analysis of Quaternion Valued Adaptive Filters

Clive Cheong Took, Cyrus Jahanchahi and Danilo P. Mandic

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

Email: {c.cheong-took, cyrus.jahanchahi, d.mandic}@imperial.ac.uk

Abstract—The recently proposed $\mathbb{H}\mathbb{R}$ -calculus has enabled rigorous derivation of quaternion-valued adaptive filtering algorithms, and has also introduced several equivalent forms of the quaternion least mean square (QLMS). This work aims to address the uniqueness of the solutions to the stochastic gradient optimisation problems, and to provide a unified framework for the derivation and analysis of quaternion least mean square algorithms. In doing so, we assess and compare the properties of the adaptive algorithms in the context of their convergence and steady state performances. For generality, the convergence properties of both QLMS and its widely linear extension, the WL-QLMS are illuminated.

I. INTRODUCTION

The advances in the technology of vector sensors has lead to the development of a new paradigm in signal processing: operating in the multidimensional quaternion domain. Recent advances in this direction have been in spectrum analysis [1], algebraic matrix decompositions [2], augmented statistics [3] and independent component analysis [4]. Similarly, the quaternion least mean square (QLMS) was proposed for the adaptive filtering of hypercomplex processes [5] and was used in various applications such as renewable energy [6] and gait recognition [7]; it has also undergone improvements in terms of computational complexity [8] and functional simplicity [9]. In contrast to the real and complex domains, applications of quaternions are not well understood. For instance, the fact that quaternion gradients are still not well defined means that the standard pseudogradient approach had to be borrowed from the complex domain to derive QLMS. As a result, the derivation of QLMS was lengthy and tedious; it involved the cumbersome task of differentiating separately with respect to the real and the three imaginary parts of the comprising quaternion variables, and solutions which are not generic extensions of the corresponding real- and complex-valued ones.

Only recently, theoretical developments in quaternion calculus have been made to facilitate the derivation of adaptive algorithms [9] [10]. In particular, the optimisation of real-valued cost functions of quaternion variables are commonly encountered in statistical signal processing; these functions are not analytic and hence cannot be differentiated in the ‘quaternion’ sense. The $\mathbb{H}\mathbb{R}$ -calculus was thus proposed to offer a unified and elegant way to differentiate directly in the hypercomplex division algebra [10]. More recently, the $\mathbb{H}\mathbb{R}$ -calculus was extended to cater specifically for functions

of quaternion involutions $q^\eta \forall \eta \in \{i, j, \kappa\}$ (analogous to the complex conjugate operator), resulting in the so-called i-gradient [9]. In the sequel, it is shown how these advances in quaternion calculus pave the way to the derivation of varying forms of QLMS that exhibit different properties.

Our aim is to introduce a unifying framework to analyse quaternion valued adaptive filters. This will not only allow us to shed light on this emerging topic, but will also enable us to make use of these adaptive filters in specific application driven contexts. For completeness, we also demonstrate the convergence properties of the QLMS based on the widely linear model, given by

$$y = \mathbf{u}^H \mathbf{x} + \mathbf{v}^H \mathbf{x}^i + \mathbf{g}^H \mathbf{x}^j + \mathbf{h}^H \mathbf{x}^\kappa \quad (1)$$

in terms of the eigenvalues of the quaternion-valued covariance matrices. These results confirm that the widely linear QLMS reduces into the QLMS when processing a proper signal, that is, a signal with equal powers in all the components of the quaternion.

II. QUATERNIONS AND ITS INVOLUTIONS

For a quaternion $q = q_a + iq_b + jq_c + \kappa q_d = \Re[q] + \Im[q]$, the real (scalar) part is denoted by $q_a = \Re[q]$, whereas the vector part (also called pure quaternion) $\Im[q] = iq_b + jq_c + \kappa q_d$ comprises the three imaginary parts. The noncommutative quaternion product is given by

$$\begin{aligned} \Re[q_1 q_2] &= \Re[q_1] \Re[q_2] - \Im[q_1] \cdot \Im[q_2] \\ \Im[q_1 q_2] &= \Re[q_2] \Im[q_1] + \Re[q_1] \Im[q_2] + \Im[q_1] \times \Im[q_2] \end{aligned} \quad (2)$$

where the symbols ‘ \cdot ’ and ‘ \times ’ denote respectively the inner and the outer product. The quaternion conjugate is given by $q^* = \Re[q] - \Im[q]$, the norm by $|q| = \sqrt{qq^*}$, and hence the inverse by $q^{-1} = \frac{q^*}{\|q\|^2}$, and $\eta^2 = -1$, $\eta \forall \{i, j, \kappa\}$.

A. Quaternion Involutions

An involution is a similarity relation and represents as a self-inverse mapping; a commonly encountered involution is the quaternion conjugate, that is, $(q^*)^* = q$. In the context of our work, however, similarity relations of particular interest are the three orthogonal (\perp) quaternion involutions, given by

$$\begin{aligned} q^i &= q_a + iq_b - jq_c - \kappa q_d & q^j &= q_a - iq_b + jq_c - \kappa q_d \\ q^\kappa &= q_a - iq_b - jq_c + \kappa q_d \end{aligned} \quad (3)$$

The properties of these \perp quaternion involutions can be summarised as

$$\begin{aligned} (p^\eta)^* &= (p^*)^\eta & (p+q)^\eta &= p^\eta + q^\eta \\ (pq)^\eta &= p^\eta q^\eta & (pq)^{\eta*} &= q^{\eta*} p^{\eta*} \\ (p^\delta)^\eta &= (p^\eta)^\delta = p^\alpha & \forall \eta \neq \delta \neq \alpha \in \{i, j, \kappa\} \end{aligned} \quad (4)$$

The relationship between the involutions and the conjugate of a quaternion variable is given by

$$q^* = \frac{1}{2}(q^i + q^j + q^\kappa - q) \quad (5)$$

$$q = \frac{1}{2}(q^{i*} + q^{j*} + q^{\kappa*} - q^*) \quad (6)$$

The identities in (5) - (6) are important for the analysis of quaternion gradients, as discussed below.

B. The $\mathbb{H}\mathbb{R}$ -Calculus and Quaternion Differentiability

The $\mathbb{H}\mathbb{R}$ -calculus provides a unifying framework for the differentiation of both analytic and non-analytic functions of quaternion variables. It was developed in [10] to overcome the stringent Cauchy-Riemann-Fueter (CRF) conditions, which are satisfied only for linear functions and constants. The $\mathbb{H}\mathbb{R}$ and $\mathbb{H}\mathbb{R}^*$ -derivatives within the $\mathbb{H}\mathbb{R}$ -calculus are given respectively as [10]

$$\frac{\partial f(q, q^i, q^j, q^\kappa)}{\partial q} = \frac{1}{4} \left[\frac{\partial f}{\partial q_a} - i \frac{\partial f}{\partial q_b} - j \frac{\partial f}{\partial q_c} - \kappa \frac{\partial f}{\partial q_d} \right] \quad (7)$$

$$\frac{\partial f(q^*, q^{i*}, q^{j*}, q^{\kappa*})}{\partial q^*} = \frac{1}{4} \left[\frac{\partial f}{\partial q_a} + i \frac{\partial f}{\partial q_b} + j \frac{\partial f}{\partial q_c} + \kappa \frac{\partial f}{\partial q_d} \right] \quad (8)$$

For instance, to use the $\mathbb{H}\mathbb{R}^*$ -derivative, we first express a quaternion valued the function $f(\cdot)$ in terms of the involutions $q^*, q^{i*}, q^{j*}, q^{\kappa*}$ using (6) and then differentiate with respect to q^* . Consider the function $f(\cdot) = q$, its $\mathbb{H}\mathbb{R}^*$ -derivative is $\frac{\partial f}{\partial q^*} = -1/2$. In addition to the $\mathbb{H}\mathbb{R}$ - and $\mathbb{H}\mathbb{R}^*$ -derivatives, a new definition of quaternion gradient based on involutions has been recently proposed [9] and is termed the involution- or i-gradient, given by

$$\nabla_{w^\eta} f(q, q^i, q^j, q^\kappa) = \frac{\partial f}{\partial q^i} + \frac{\partial f}{\partial q^j} + \frac{\partial f}{\partial q^\kappa} \quad (9)$$

$$\begin{aligned} \nabla_{w^\eta} f(q, q^i, q^j, q^\kappa) &= \frac{1}{4} \left[3 \frac{\partial}{\partial q_a} + i \frac{\partial f}{\partial q_b} + j \frac{\partial f}{\partial q_c} + \kappa \frac{\partial f}{\partial q_d} \right] \\ &= \frac{\partial f}{\partial q^*} + \frac{1}{2} \frac{\partial f}{\partial q_a} \end{aligned} \quad (10)$$

where, for simplicity, $\eta \in \{i, j, \kappa\}$. In addition to the standard $\mathbb{H}\mathbb{R}^*$ -derivative $\frac{\partial f}{\partial q^*}$, the additional gradient term in the direction of the real variable $\partial f / \partial q_a$ in (10) indicates a potential faster convergence of the i-gradient, as compared to the standard conjugate gradient using the $\mathbb{H}\mathbb{R}^*$ -derivative. For more details on the i-gradient, see [9].

III. QUATERNION-VALUED ADAPTIVE FILTERS REVISITED

For convenience, the derivations¹ of adaptive filtering algorithms based on the $\mathbb{H}\mathbb{R}$ -calculus are next revisited.

¹For continuity with our work [5], the filter output $y(k) = \mathbf{w}^T(k)\mathbf{x}(k)$.

A. iQuaternion Least Mean Square Algorithm

We first consider the derivation of the iQLMS, which is based on the i-gradient for $\eta \in \{i, j, \kappa\}$. The gradient of the cost function $J(k) = \frac{1}{2}e(k)e^*(k)$ is given in the form

$$\nabla_{\mathbf{w}^\eta} J(k) = \frac{1}{2} \sum_{\eta=\{i, j, \kappa\}} e(k) \frac{\partial e^*(k)}{\partial \mathbf{w}^\eta(k)} + \frac{\partial e(k)}{\partial \mathbf{w}^\eta(k)} e^*(k) \quad (11)$$

Using the rules of $\mathbb{H}\mathbb{R}$ -calculus, we have [10]

$$\frac{\partial e(k)}{\partial \mathbf{w}^\eta} = \mathbf{0} \quad \forall \eta \in \{i, j, \kappa\} \quad (12)$$

To derive $\frac{\partial e^*(k)}{\partial \mathbf{w}^\eta} \forall \eta \in \{i, j, \kappa\}$, use the identity in (5) to substitute \mathbf{w}^* into the expression for $e^*(k)$, to give

$$e^*(k) = d^*(k) - \frac{1}{2} \mathbf{x}^H(k) (\mathbf{w}^i(k) + \mathbf{w}^j(k) + \mathbf{w}^\kappa(k) - \mathbf{w}(k))$$

The gradients $\frac{\partial e(k)}{\partial \mathbf{w}^i}$, $\frac{\partial e(k)}{\partial \mathbf{w}^j}$ and $\frac{\partial e(k)}{\partial \mathbf{w}^\kappa}$ can now be calculated by direct differentiation, yielding

$$\frac{\partial e(k)}{\partial \mathbf{w}^\eta} = -\frac{1}{2} \mathbf{x}^*(k) \quad \forall \eta \in \{i, j, \kappa\} \quad (13)$$

Substituting (12) and (13) into (11) and substituting the result in the steepest descent $\mathbf{w}(k+1) = \mathbf{w}(k) - \mu \nabla_{\mathbf{w}^\eta} J(k)$ gives the weight update

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \frac{3}{4} \mu e(k) \mathbf{x}^*(k) \quad (14)$$

This stochastic gradient algorithm is termed the iQLMS (the i-gradient based QLMS). Observe that it has the same form as the LMS and complex LMS [11].

B. Quaternion Least Mean Square Algorithm

Similarly, the original QLMS algorithm [5] can now be rederived based on the $\mathbb{H}\mathbb{R}^*$ -derivative. Its cost function is given by $\mathcal{J}(k) = \frac{1}{2}e(k)e^*(k)$ and its weight update expressed as $\mathbf{w}(k+1) = \mathbf{w}(k) - \mu \nabla_{\mathbf{w}^*} \mathcal{J}(k)$, where

$$\nabla_{\mathbf{w}^*} \mathcal{J}(k) = e(k) \frac{\partial e^*(k)}{\partial \mathbf{w}^*(k)} + \frac{\partial e(k)}{\partial \mathbf{w}^*(k)} e^*(k) \quad (15)$$

Expanding the error expression, we obtain the terms $e(k) = d(k) - \mathbf{w}^T(k)\mathbf{x}(k)$ and $e^*(k) = d^*(k) - \mathbf{x}^H(k)\mathbf{w}^*(k)$. The relationship in (6) permits the use of $\mathbb{H}\mathbb{R}^*$ -calculus, yielding $\partial \mathbf{w}(k) / \partial \mathbf{w}^*(k) = -1/2$. The error gradients thus become

$$\frac{\partial e^*(k)}{\partial \mathbf{w}^*(k)} = -\mathbf{x}^*(k) \quad \frac{\partial e(k)}{\partial \mathbf{w}^*(k)} = \frac{1}{2} \mathbf{x}(k) \quad (16)$$

giving the update referred to as the “ $\mathbb{H}\mathbb{R}$ -QLMS” update

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \frac{1}{2} \mu \left(e(k) \mathbf{x}^*(k) - \frac{1}{2} \mathbf{x}(k) e^*(k) \right) \quad (17)$$

On the other hand, the QLMS originally proposed in [5] involves a lengthy derivation, as it was based on the componentwise pseudogradient; for brevity, it is given by

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \frac{1}{2} \mu \left(e(k) \mathbf{x}^*(k) - \frac{1}{2} \mathbf{x}^*(k) e^*(k) \right) \quad (18)$$

TABLE I
COMPARATIVE ANALYSIS OF QUATERNION ADAPTIVE FILTERS

Algorithms in their generic form (21)	ν	τ	ρ	ς
iQLMS in (14)	3/4	3/4	3/4	3/4
HIR-QLMS in (17)	1/4	1/4	3/4	3/4
QLMS in (18)	1/4	3/4	3/4	1/4

IV. UNIFIED ANALYSIS OF QUATERNION ADAPTIVE FILTERS

The three quaternion least mean square algorithms in (14), (17), and (18) have been derived based on different gradient definitions, leading to different updates. To have a unified framework to perform a fair and comparative analysis of their properties, we shall express all the updates in (14), (17), and (18) in the same generic form.

For clarity, we start with the simplest functional expression, the iQLMS, and separate its update into the real and imaginary parts by using the identity $\mathbf{x}^*(k) = \mathbf{x}(k) - 2\Im[\mathbf{x}(k)]$, as follows

$$\begin{aligned} \Re[\mathbf{w}(k+1)] &= \Re[\mathbf{w}(k)] + \frac{3}{4}\mu\Re[e(k)\mathbf{x}^*(k)] \\ &= \Re[\mathbf{w}(k)] + \frac{3}{4}\mu\Re[e(k)(\Re[\mathbf{x}(k)] - \Im[\mathbf{x}(k)])] \\ &= \Re[\mathbf{w}(k)] + \frac{3}{4}\mu\Re[e(k)\Re[\mathbf{x}(k)]] - \frac{3}{4}\mu\Re[e(k)\Im[\mathbf{x}(k)]] \quad (19) \end{aligned}$$

$$\begin{aligned} \Im[\mathbf{w}(k+1)] &= \Im[\mathbf{w}(k)] + \frac{3}{4}\mu\Im[e(k)\mathbf{x}^*(k)] \\ &= \Im[\mathbf{w}(k)] + \frac{3}{4}\mu\Im[e(k)(\Re[\mathbf{x}(k)] - \Im[\mathbf{x}(k)])] \\ &= \Im[\mathbf{w}(k)] + \frac{3}{4}\mu\Im[e(k)\Re[\mathbf{x}(k)]] - \frac{3}{4}\mu\Im[e(k)\Im[\mathbf{x}(k)]] \quad (20) \end{aligned}$$

A more generic form of these equations can be rewritten as

$$\begin{aligned} \Re[\Delta\mathbf{w}] &= \nu\mu\Re[e(k)\Re[\mathbf{x}(k)]] - \tau\mu\Re[e(k)\Im[\mathbf{x}(k)]] \\ \Im[\Delta\mathbf{w}] &= \rho\mu\Im[e(k)\Re[\mathbf{x}(k)]] - \varsigma\mu\Im[e(k)\Im[\mathbf{x}(k)]] \quad (21) \end{aligned}$$

where $\Delta\mathbf{w} = \mathbf{w}(k+1) - \mathbf{w}(k)$. The updates for the HIR-QLMS and QLMS in (17) – (18) can also be expressed in the same generic form as above. Observe that their weight updates only differ in the weighting coefficients ν , τ , ρ and ς , as shown in Table I.

Remark#1: The three QLMS algorithms in (14), (17), and (18) are expected to exhibit similar performances at steady state, as they essentially utilise the same information;

Remark#2: Each algorithm in (21) has different coefficients ν , τ , ρ and ς , suggesting different rates of convergence. For instance, the comparatively biggest coefficients of the iQLMS indicates its fastest convergence rate.

These remarks have been confirmed the simulation studies of our work in [9].

V. ON THE RELATIONSHIP BETWEEN QLMS AND WIDELY LINEAR QLMS

Now that the relationship between the variants of the QLMS has been established, we proceed to illuminate the relationship between the QLMS and the widely linear QLMS. For simplicity, we consider the iQLMS and its corresponding widely linear counterpart, WL-iQLMS given by²

$$\mathbf{w}^a(k+1) = \mathbf{w}^a(k) + \frac{3}{4}\mu\mathbf{x}^a(k)e^*(k) \quad (22)$$

where $\mathbf{x}^a = [\mathbf{x}^T \ \mathbf{x}^{iT} \ \mathbf{x}^{jT} \ \mathbf{x}^{\kappa T}]^T$ is the augmented input vector and $\mathbf{w}^a = [\mathbf{u}^T \ \mathbf{v}^T \ \mathbf{g}^T \ \mathbf{h}^T]^T$ in the WL model (1). Notice that the widely linear iQLMS degenerates into the iQLMS for $\mathbf{x}^a = \mathbf{x}$ and $\mathbf{w}^a = \mathbf{u}$, illustrating the obvious: the second order statistics of the widely linear iQLMS is augmented with the quaternion involutions as exemplified by its augmented covariance matrix

$$\mathcal{C}_a = E\{\mathbf{x}^a\mathbf{x}^{aH}\} = \begin{bmatrix} \mathcal{C} & \mathcal{C}_i & \mathcal{C}_j & \mathcal{C}_\kappa \\ (\mathcal{C}_i)^i & \mathcal{C}^i & (\mathcal{C}_\kappa)^i & (\mathcal{C}_j)^i \\ (\mathcal{C}_j)^j & (\mathcal{C}_\kappa)^j & \mathcal{C}^j & (\mathcal{C}_i)^j \\ (\mathcal{C}_\kappa)^\kappa & (\mathcal{C}_j)^\kappa & (\mathcal{C}_i)^\kappa & \mathcal{C}^\kappa \end{bmatrix} \quad (23)$$

where $\mathcal{C} = E\{\mathbf{x}\mathbf{x}^H\}$, $\mathcal{C}_i = E\{\mathbf{x}\mathbf{x}^{iH}\}$, $\mathcal{C}_j = E\{\mathbf{x}\mathbf{x}^{jH}\}$, $\mathcal{C}_\kappa = E\{\mathbf{x}\mathbf{x}^{\kappa H}\}$. Upon adopting an approximate diagonalisation procedure [12], holds, the approximation $\varepsilon = \mathbf{x}^H\mathbf{x} \approx (\mathbf{x}^H\mathbf{x})^\eta \ \forall \eta \in \{i, j, \kappa\}$ implying that one singular value decomposition is adequate to diagonalise all the quaternion-valued correlation matrices, that is, (see [12] and [2] for more details)

$$\begin{aligned} \mathcal{C} &\approx \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H & \mathcal{C}_i &\approx \mathbf{U}\mathbf{\Lambda}_i\mathbf{U}^{iH} \\ \mathcal{C}_j &\approx \mathbf{U}\mathbf{\Lambda}_j\mathbf{U}^{jH} & \mathcal{C}_\kappa &\approx \mathbf{U}\mathbf{\Lambda}_\kappa\mathbf{U}^{\kappa H} \end{aligned} \quad (24)$$

In this spirit, the augmented covariance matrix in (23) can be divided into four 2×2 submatrices

$$\mathcal{C}_a = \begin{bmatrix} \mathbf{\Gamma} & \mathbf{\Sigma} \\ \mathbf{\Sigma}^j & \mathbf{\Gamma}^j \end{bmatrix} \quad (25)$$

Notice that $\mathbf{\Gamma} = \mathbf{\Gamma}^H$ and that $\mathbf{\Sigma} = \mathbf{\Sigma}^{jH}$, and therefore they can be factorised as

$$\begin{aligned} \mathbf{\Gamma} &= \mathbf{Q}\mathbf{\Lambda}_\gamma\mathbf{Q}^H \\ \mathbf{\Sigma} &= \mathbf{Q}\mathbf{\Lambda}_\sigma\mathbf{Q}^{jH} \end{aligned} \quad (26)$$

These factorisations (26) allow us to express \mathcal{C}_a in terms of the eigenvalues γ and σ of $\mathbf{\Gamma}$ and $\mathbf{\Sigma}$, that is

$$\mathcal{C}_a = \begin{bmatrix} \mathbf{Q} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}^j \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda}_\gamma & \mathbf{\Lambda}_\sigma \\ \mathbf{\Lambda}_\sigma & \mathbf{\Lambda}_\gamma \end{bmatrix} \begin{bmatrix} \mathbf{Q}^H & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}^{jH} \end{bmatrix} \quad (27)$$

The middle term above can be therefore diagonalised as

$$\begin{bmatrix} \mathbf{\Lambda}_\gamma & \mathbf{\Lambda}_\sigma \\ \mathbf{\Lambda}_\sigma & \mathbf{\Lambda}_\gamma \end{bmatrix} = \frac{1}{2} \begin{bmatrix} \mathbf{I} & -\mathbf{I} \\ \mathbf{I} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda}_\gamma + \mathbf{\Lambda}_\sigma & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_\gamma - \mathbf{\Lambda}_\sigma \end{bmatrix} \begin{bmatrix} -\mathbf{I} & \mathbf{I} \\ \mathbf{I} & \mathbf{I} \end{bmatrix} \quad (28)$$

yielding the condition number of \mathcal{C}_a

$$\kappa(\mathcal{C}_a) = \frac{\max(\lambda_\gamma) + \max(\lambda_\sigma)}{\min(\lambda_\gamma) - \min(\lambda_\sigma)} \quad (29)$$

²For mathematical tractability, the filter output $y(k) = \mathbf{w}^H(k)\mathbf{x}(k)$.

However, we experience an obstacle since $\kappa(\mathcal{C}_a)$ cannot be expressed in terms of the covariance and complementary covariance matrices. To this end, the same diagonalisation procedure is recursively used to further decompose $\mathbf{\Gamma}$ and $\mathbf{\Sigma}$ as

$$\begin{aligned}\mathbf{\Gamma} &= \mathbf{Q}\mathbf{\Lambda}_\gamma\mathbf{Q}^H \\ &= \frac{1}{2}\begin{bmatrix} \mathbf{U} & -\mathbf{U} \\ \mathbf{U}^\iota & \mathbf{U}^\iota \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda} + \mathbf{\Lambda}_\iota & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda} - \mathbf{\Lambda}_\iota \end{bmatrix} \begin{bmatrix} \mathbf{U}^H & \mathbf{U}^{\iota H} \\ -\mathbf{U}^H & \mathbf{U}^{\iota H} \end{bmatrix}\end{aligned}\quad (30)$$

$$\begin{aligned}\mathbf{\Sigma} &= \mathbf{Q}\mathbf{\Lambda}_\sigma\mathbf{Q}^H \\ &= \frac{1}{2}\begin{bmatrix} \mathbf{U} & -\mathbf{U} \\ \mathbf{U}^\iota & \mathbf{U}^\iota \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda}_j + \mathbf{\Lambda}_\kappa & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_j - \mathbf{\Lambda}_\kappa \end{bmatrix} \begin{bmatrix} \mathbf{U}^{jH} & \mathbf{U}^{\kappa H} \\ -\mathbf{U}^{jH} & \mathbf{U}^{\kappa H} \end{bmatrix}\end{aligned}\quad (31)$$

This demonstrates that the augmented covariance in (25) can be decomposed into $\mathcal{C}_a = \mathbf{V}\mathbf{\Lambda}_a\mathbf{V}^H$, where $\mathbf{\Lambda}_a$ is given by

$$\mathbf{\Lambda}_a = \begin{bmatrix} \mathbf{\Lambda}_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{\Lambda}_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{\Lambda}_4 \end{bmatrix}\quad (32)$$

and

$$\begin{aligned}\mathbf{\Lambda}_1 &= \mathbf{\Lambda} + \mathbf{\Lambda}_\iota + \mathbf{\Lambda}_j + \mathbf{\Lambda}_\kappa & \mathbf{\Lambda}_2 &= \mathbf{\Lambda} - \mathbf{\Lambda}_\iota + \mathbf{\Lambda}_j - \mathbf{\Lambda}_\kappa \\ \mathbf{\Lambda}_3 &= \mathbf{\Lambda} + \mathbf{\Lambda}_\iota - \mathbf{\Lambda}_j - \mathbf{\Lambda}_\kappa & \mathbf{\Lambda}_4 &= \mathbf{\Lambda} - \mathbf{\Lambda}_\iota - \mathbf{\Lambda}_j + \mathbf{\Lambda}_\kappa\end{aligned}\quad (33)$$

A. Convergence analysis

We now demonstrate the usefulness of the diagonalisation in (32) in adaptive filtering applications. From (29), the convergence properties of the widely-linear iQLMS can be expressed through the condition number of augmented covariance matrix

$$\begin{aligned}\kappa(\mathcal{C}_a) &= \frac{\max(\lambda + \lambda_\iota) + \max(\lambda_j + \lambda_\kappa)}{\min(\lambda - \lambda_\iota) - \min(\lambda_j - \lambda_\kappa)} \\ &= \frac{\max(\lambda) + \max(\lambda_\iota) + \max(\lambda_j) + \max(\lambda_\kappa)}{\min(\lambda) - \min(\lambda_\iota) - \min(\lambda_j) + \min(\lambda_\kappa)}\end{aligned}\quad (34)$$

Remark#3: The condition number of \mathcal{C}_a can be expressed in terms of the true eigenvalues of the original covariance matrix \mathcal{C}_x and the complementary covariance matrices $\mathcal{C}_\eta \forall \eta \in \{\iota, j, \kappa\}$, instead of those of the augmented covariance matrix \mathcal{C}_a , which are less straightforward to physically interpret in terms of e.g. the degree of circularity or signal power.

Remark#4: The generality of the expression in (34) allows us to inspect special cases of component-wise impropreness. For instance, for a Q -proper signal the complementary covariance matrices vanish, implying that $\lambda_\eta = 0, \forall \eta \in \{\iota, j, \kappa\}$; in this case the condition number degenerates into that of the standard covariance matrix, i.e. $\kappa(\mathcal{C}_a) = \frac{\max(\lambda)}{\min(\lambda)}$ for the iQLMS.

B. Stability Bounds

As the stepsize parameter μ plays a prominent role in the performance of adaptive filters, we shall now derive the stability bounds as a function of the stepsize. Since the error can be modelled as $e(k) = (\mathbf{w}_{a,opt} - \mathbf{w}_a(k))^H \mathbf{x}_a(k)$, where

$\mathbf{w}_{a,opt}$ denotes the Wiener solution, the update of the widely linear iQLMS in (22) becomes³

$$\begin{aligned}\mathbf{w}_a(k+1) - \mathbf{w}_{a,opt} &= \mathbf{w}_a(k) - \mathbf{w}_{a,opt} \\ &\quad + \mu \mathbf{x}_a(k) \mathbf{x}_a^H(k) \left(\mathbf{w}_a(k) - \mathbf{w}_{a,opt} \right) \\ \check{\mathbf{w}}_a(k+1) &= \left(\mathbf{I} - \mu \mathcal{C}_a \right) \check{\mathbf{w}}_a(k)\end{aligned}\quad (35)$$

The evolution of (35) gives the stability bounds for the stepsize

$$0 < \mu < \frac{2}{\max(\lambda) + \max(\lambda_\iota) + \max(\lambda_j) + \max(\lambda_\kappa)}\quad (36)$$

Remark#5: The corresponding stability bounds for the iQLMS can be straightforwardly obtained as $0 < \mu < \frac{2}{\max(\lambda)}$.

VI. CONCLUSION

We have provided a unifying framework for the analysis of quaternion-valued adaptive filtering algorithms, achieved by expressing all the variants of Quaternion LMS in the same generic form. This has enabled for the convergence and steady state properties to be studied in a unifying way. Rigorous analysis by making use of the eigenvalues of the covariance and the complementary covariance matrices has shown how the degree of impropriety of a signal affects the convergence and stability of the widely linear QLMS. Similarly, the widely linear QLMS has been shown to degenerate into the QLMS when processing proper signals, illustrating the usefulness and rigour of the proposed analysis framework.

REFERENCES

- [1] S. Said, N. L. Bihan, and S. J. Sangwine, "Fast complexified quaternion Fourier transform," *IEEE Transactions on Signal Processing*, vol. 56, no. 4, pp. 1522–1531, 2008.
- [2] C. Cheong Took, D. P. Mandic, and F. Zhang, "On the Unitary Diagonalisation of a Special Class of Quaternion Matrices," *Applied Mathematics Letters*, vol. 24, no. 11, pp. 1806 – 1809, 2011.
- [3] J. Vía, D. Ramírez, and I. Santamaría, "Properness and Widely Linear Processing of Quaternion Random Vectors," *IEEE Transactions on Information Theory*, vol. 56, no. 7, pp. 3502–3515, 2010.
- [4] S. Javidi, C. Cheong Took, and D. Mandic, "Fast Independent Component Analysis Algorithm for Quaternion Valued Signals," *IEEE Transactions on Neural Networks in Press*, 2011.
- [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*, vol. 57, no. 4, pp. 1316–1327, 2009.
- [6] C. Cheong Took, G. Strbac, K. Aihara, and D. P. Mandic, "Quaternion-valued short term forecasting of three-dimensional wind and atmospheric parameters," *Renewable Energy*, vol. 36, no. 6, pp. 1754 – 1760, 2011.
- [7] A. Bravi and A. Sabatini, "A multidimensional approach to postural sway modelling," in *Proceedings of IEEE Workshop on Medical Measurements and Applications*, pp. 121 – 124, April 2010.
- [8] F. G. A. Neto and V. H. Nascimento, "A novel reduced-complexity widely linear QLMS algorithm," in *Proceedings of IEEE Statistical Signal Processing Workshop*, pp. 81–84, June 2011.
- [9] C. Jahanchahi, C. Cheong Took, and D. P. Mandic, "On gradient calculation in quaternion adaptive filtering," *submitted to ICASSP*, 2012.
- [10] D. P. Mandic, C. Jahanchahi, and C. Cheong Took, "A quaternion gradient operator and its applications," *IEEE Signal Processing Letters*, vol. 18, no. 1, pp. 47 –50, 2011.
- [11] 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.
- [12] C. Cheong Took, S. C. Douglas, and D. P. Mandic, "On Approximate Diagonalisation of Correlation Matrices in Widely Linear Signal Processing," *IEEE Transactions on Signal Processing*, *submitted*, 2011.

³For simplicity, the term 3/4 in (22) can be absorbed into the stepsize μ .